



Published in final edited form as:

Clin Trials. 2021 April ; 18(2): 207–214. doi:10.1177/1740774520980070.

A Case Study of Ascertainment Bias for the Primary Outcome in the Strategies to Reduce Injuries and Develop Confidence in Elders (STRIDE) Trial

Denise A. Esserman¹, Thomas M. Gill², Michael E. Miller³, Erich J. Greene¹, James D. Dziura¹, Thomas G. Trivison⁴, Can Meng¹, Peter N. Peduzzi¹

¹Yale Center for Analytical Sciences; Yale University; New Haven, CT, USA

²Yale School of Medicine, Department of Internal Medicine, New Haven, CT, USA

³Wake Forest University School of Medicine, Winston-Salem, NC, USA

⁴Marcus Institute for Aging Research, Hebrew SeniorLife, Harvard Medical School, Boston, MA, USA

Abstract

Background/Aims: In clinical trials, there is potential for bias from unblinded observers that may influence ascertainment of outcomes. This issue arose in the Strategies to Reduce Injuries and Develop Confidence in Elders (STRIDE) Trial, a cluster randomized trial to test a multicomponent intervention vs. enhanced usual care (control) to prevent serious fall injuries, originally defined as a fall injury leading to medical attention. An unblinded nurse falls care manager administered the intervention, while the usual care arm did not involve contact with a falls care manager. Thus, there was an opportunity for falls care managers to refer participants reporting falls to seek medical attention. Since this type of observer bias could not occur in the usual care arm, there was potential for additional falls to be reported in the intervention arm, leading to dilution of the intervention effect and a reduction in study power. We describe the clinical basis for ascertainment bias, the statistical approach used to assess it, and its effect on study power.

Methods: The prespecified interim monitoring plan included a decision algorithm for assessing ascertainment bias and adapting (revising) the primary outcome definition, if necessary. The original definition categorized serious fall injuries requiring medical attention into Type 1 (fracture other than thoracic/lumbar vertebral; joint dislocation; cut requiring closure) and Type 2 (head injury; sprain or strain; bruising or swelling; other). The revised definition, proposed by the monitoring plan, excluded Type 2 injuries that did not necessarily require an overnight hospitalization since these would be most subject to bias. These injuries were categorized into those with (Type 2b) and without (Type 2c) medical attention. The remaining Type 2a injuries required medical attention and an overnight hospitalization. We used the ratio of $2b/(2b + 2c)$ in

Corresponding Author: Denise A Esserman, Yale Center for Analytical Sciences, Suite 511, 300 George Street, New Haven, CT, USA 06510, denise.esserman@yale.edu, Phone: (203) 785-4297.

Trial registration number at clinicaltrials.gov: NCT02475850

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

intervention vs. control as a measure of ascertainment bias; ratios > 1 indicated the likelihood of falls care manager bias. We determined the effect of ascertainment bias on study power for the revised (Types 1 and 2a) vs. original definition (Types 1, 2a and 2b).

Results: The estimate of ascertainment bias was 1.14 (95% CI: 0.98, 1.30), providing evidence of the likelihood of falls care manager bias. We estimated that this bias diluted the hazard ratio from the hypothesized 0.80 to 0.86 and reduced power to under 80% for the original primary outcome definition. In contrast, adapting the revised definition maintained study power at nearly 90%.

Conclusion: There was evidence of ascertainment bias in STRIDE. The decision to adapt the primary outcome definition reduced the likelihood of this bias while preserving the intervention effect and study power.

Keywords

Cluster randomized trial; ascertainment bias; adjudication; hazard ratio; power

Background

The Cochrane Handbook for Systematic Reviews of Interventions defines detection bias as “systematic differences between groups in how outcomes are determined.”¹ This bias is also termed *ascertainment bias* and occurs when knowledge of a trial participant’s treatment assignment influences outcome assessment. When ascertainment bias results from participants, it is often referred to as *response bias*, and when it results from researchers, it is referred to as *observer bias*. Blinding or masking of participants and/or researchers is a standard approach to minimize ascertainment bias; however, in many trials it is not possible to blind these individuals. Thus, measures to mitigate this type of bias are needed to obtain valid comparisons between randomized groups.

Concerns about ascertainment bias arose during the planning of the Strategies to Reduce Injuries and Develop Confidence in Elders (STRIDE) Trial.² STRIDE was designed to compare a multicomponent intervention to enhanced usual care (control) for the prevention of serious fall injuries. Enrollment and follow-up were conducted by blinded staff at the central recruitment and assessment center, and the primary outcome was adjudicated by a blinded committee; however, treatment assignment could not be blinded. A nurse falls care manager administered the intervention. Because the study definition of a serious fall injury included seeking any medical attention, the study team was concerned that falls care managers might recommend participants reporting a fall to them to seek medical attention. Since this type of observer bias could not occur in the control arm, there was the possibility for additional fall injuries to be reported in the intervention arm leading to a dilution of the intervention effect. Although there was a suspicion of bias, there was insufficient information upon which to base a decision to change the primary outcome definition. Thus, STRIDE was designed with several adaptive features, one of which was monitoring the potential for ascertainment bias so that an informed decision could be made about changing the primary outcome definition if necessary. In this paper, we describe the design of

STRIDE, the clinical basis for ascertainment bias, and the interim monitoring approach used to assess the likelihood of ascertainment bias and its effect on study power.

Methods

Design of STRIDE

Full details of the STRIDE design are provided by Bhasin et al.³ Briefly, STRIDE was a pragmatic cluster randomized, parallel group, superiority trial that enrolled and randomized 86 clinical practices stratified by healthcare system (10 across US) with participants nested within practices. Participants were community-living persons, 70 years or older, at increased risk for serious fall injuries. The trial enrolled 5,451 individuals over a 20-month period (August 1, 2015 to March 31, 2017). The intervention was delivered for a minimum of 20 months and a maximum of 40 months. Follow-up ended on March 31, 2019; maximum follow-up was 44 months.

The intervention, described in detail elsewhere,⁴ was an individually tailored set of recommendations/interventions based on a multifactorial risk assessment that utilized a primary care co-management model implemented by a nurse falls care manager. The participants in the control practices received a falls informational booklet⁵ and were encouraged to discuss fall prevention with their primary care provider at their next clinic visit. Their physicians received the results of the three screening questions³ and were referred to a training webinar about fall prevention.

The primary outcome in STRIDE was first serious fall injury and was adjudicated by an Adjudication Committee blinded to randomized group. Two sources of data were required to confirm a serious fall injury. The primary data source was participants' self-report of a fall leading to medical attention identified during a 4-monthly telephone interview, while confirmatory sources could be (1) administrative data (electronic health records; Medicare) or (2) the full text of selected medical records.⁶

STRIDE was designed as an event driven trial to detect a hazard ratio for intervention relative to control of 0.80 with a type I error of 5% (2-sided) and 90% power. The required number of events was determined by the method of Schoenfeld,⁷ which yielded 844 events. The target number of participants required to achieve this number of events in STRIDE was 6000 and was determined using PASS⁸ software (version 12) assuming a control event rate of 14% per year, a loss to follow-up rate of 3% per year, and a competing risk of death among those at risk of first fall of 7% per year for a 36-month trial with 18 months of uniform enrollment. With these parameters, the needed 844 events would be comprised of 462 control and 382 intervention events. The sample size was inflated (1) for a cluster design effect of 1.53 based on an intracluster correlation coefficient of 0.0076 [estimated from an analysis of serious fall injuries in the LIFE Study,⁹ the only data available in a similar population at the time] and 86 clusters; and (2) 3% for interim monitoring for efficacy and futility. (Per decision of the Data and Safety Monitoring Board [DSMB], interim monitoring for efficacy/futility was not done.) The trial duration was ultimately extended to 44 months to generate more events to maintain study power.

Clinical basis for ascertainment bias and adapting the primary outcome

The original protocol definition of the primary outcome was a fall injury leading to medical attention, including non-vertebral fractures, joint dislocation, head injury, lacerations, and other major sequelae (e.g., rhabdomyolysis, internal injuries, hypothermia). Under this definition, seeking medical attention was intended to affirm the seriousness of the injury. Fall-related injuries were classified into two types:

Type 1: fracture other than thoracic/lumbar vertebral; joint dislocation; or cut requiring closure;

Type 2: head injury; sprain or strain; bruising or swelling; or other.

The trial leadership recognized during the planning of the study that, as part of the intervention, participant interactions with the falls care managers, which were primarily done in person, could lead to differential medical attention, particularly for the subset of “Type 2” injuries – head injury, sprain or strain, bruising or swelling, or other – that were insufficiently severe to require a hospitalization, and may not have led the participant to seek medical attention in the absence of interactions with the falls care manager. Because receipt of medical care was an inherent part of the primary outcome definition, any factor that differentially caused participants to obtain medical care in one arm versus the other could bias estimates of outcome rates between the arms. This bias could dilute the overall intervention effect, reduce the trial’s power, and threaten the validity of the trial’s findings.

To address the concern about potential ascertainment bias, the interim monitoring plan included a decision algorithm for assessing ascertainment bias and adapting the primary outcome, if necessary (Figure 1). The adaptation would revise the definition of the Type 2 fall injuries to require an overnight hospitalization rather than any medical attention. Because Type 2 injuries can differ greatly in their severity, they were considered to be much more susceptible to observer or response bias. Requiring hospitalization for these Type 2 injuries would reduce any bias since the decision to hospitalize participants would not be influenced by participants’ interactions with the falls care manager. Requiring hospitalization would also increase the likelihood that all injuries were serious rather than minor.

Despite the advantages of revising the primary outcome definition, it would lower the rate of first serious fall injury and diminish the statistical power of the trial. Because of these tradeoffs, the study team was supportive of collecting more data to make an informed decision about adapting the primary outcome.

Interim monitoring plan for assessing ascertainment bias and adapting the primary outcome

The final interim monitoring plan was approved by the STRIDE DSMB at its May 31, 2017 meeting before any unblinding of outcomes. Figure 1 displays Stage 1 of the interim monitoring plan with the decision algorithm for assessing ascertainment bias and adapting the primary outcome, if necessary. The first step determined whether there was evidence for ascertainment bias. If there was evidence, the second step assessed whether the annual control event rate under the revised definition was $\geq 10\%$ (the minimum level to maintain

80% power under the hypothesized hazard ratio of 0.80 for intervention relative to control). If the event rate was 10%, the primary outcome definition was adapted to control bias and dilution of the intervention effect. If the control event rate was < 10%, the third step compared the power under both the original protocol definition (accounting for the diluted intervention effect due to ascertainment bias) and the revised definition and selected the definition giving the greater power. The interim monitoring plan was executed by the unblinded trial statisticians using the observed control event rates based on patient-reported events and the hypothesized hazard ratio of 0.80. Patient-reported events were used because they were the primary data source for ascertaining fall injuries, adjudication of patient-reported events was not completed until the end of the trial, with a considerable lag between reported and adjudicated events, and not all events under consideration for this study would have been adjudicated (see below).

At the January 4, 2018 DSMB meeting, the DSMB evaluated the evidence for ascertainment bias based on data lock as of November 30, 2017. After determining there was evidence of possible ascertainment bias, the DSMB recommended that the decision on adapting the primary outcome definition be made by an independent, blinded panel. (This recommendation was made because the DSMB was also reviewing interim safety data that included information about fall-related hospitalizations by randomized group, designated as A vs. B, and was not comfortable addressing the adaption issue, which needed a decision made blinded to outcomes by randomized groups.) The sponsors (National Institute on Aging and Patient Centered Outcomes Research Institute) convened an independent Ad Hoc Expert Panel, which reviewed data provided by the unblinded trial statisticians on March 12, 2018 based on data lock as of February 22, 2018. These data included **only** control event rates. Based on the data, the panel recommended that the primary outcome should be adapted, and the sponsors accepted this recommendation.

Approach to assessing ascertainment bias and its effect on power

To assess ascertainment bias, we classified Type 2 fall-related injuries (head injury; sprain or strain; bruising or swelling; or other) into the following three types:

Type 2a: requires medical attention **and an** overnight hospitalization.

Type 2b: requires medical attention **but not an** overnight hospitalization.

Type 2c: did not require medical attention. Although not part of the protocol definition of a serious fall injury, these were the types of injuries that the falls care manager might refer participants to seek medical attention for and thus convert to Type 2b injuries. (Type 2c injuries were not adjudicated.)

The original protocol definition for the primary outcome included Types 1, 2a and 2b injuries.

The proposed adapted definition included only Types 1 and 2a injuries.

We hypothesized that ascertainment bias would arise if patients with injuries that did not require medical attention (Type 2c) were referred to medical attention by the falls care manager. To determine the expected inflation in Type 2b events caused by falls care manager

referral, we compared the proportion of Type 2 injuries not requiring hospitalization (2b and 2c) that were Type 2b – i.e., the fraction $2b/(2b+2c)$ – in the intervention and control groups. [Note, other ratios could have been used to detect bias, but the ratio based on $2b/(2b+2c)$ is particularly useful for calculating the impact of bias on observed events and study power, as it includes only the bias-able events and the potential excess events.] We assumed the intervention affected all events equally and that in the absence of ascertainment bias, these proportions should be the same, yielding a ratio (intervention vs. control) of 1. Therefore, ratios greater than 1 would suggest that the falls care managers were referring participants to seek medical attention at greater rates than patients would otherwise have sought on their own. Because there was no corresponding mechanism for this type of referral bias in the control arm, we assumed that ascertainment bias operated only in the intervention arm. This bias would increase the number of observed events in the intervention arm, increasing the intervention event rate, but would have no effect on the event rate in the control arm. The net result would be an increase in the hazard ratio towards 1 and a decrease in the power of the trial. Our approach to addressing this problem was four-fold:

1. Estimated the degree of ascertainment bias in Type 2b intervention events. We assumed that the mechanism of ascertainment bias was true 2c events being observed as 2b events. We calculated the fraction of the combined 2b and 2c events in each arm that were actually 2b. In the absence of ascertainment bias (and with a few additional assumptions), this fraction should be the same in both arms. Therefore, we used the ratio (1 indicating no bias) of these fractions to construct a measure of bias.
2. Estimated how this bias affects the overall number of intervention events. We estimated the excess number of events observed in the intervention arm using the expected number of events without any bias ("true" events) and a proportionality factor that took into account the estimate of the ascertainment bias.
3. Used the bias-adjusted number of intervention events to estimate the bias-adjusted intervention hazard rate and the bias-adjusted hazard ratio comparing the randomized groups. This was done by expressing the expected observed number of intervention events both in terms of the hypothesized hazard ratio and estimated bias and in terms of an effective (bias-adjusted) hazard ratio (ratio of the intervention to control hazards in the presence of ascertainment bias) and solving for the effective hazard ratio.
4. Used the bias-adjusted event total and hazard ratio to estimate study power. Plugging these into the standard power formula for time-to-event data,⁷ we estimated the projected power factoring in the impact of ascertainment bias.

The technical details for the above estimates are derived in Greene et al.¹⁰

Results

Estimating ascertainment bias

Based on patient-reported events (Table 1) ascertained as of February 22, 2018, the proportion of Type 2b events ($2b/[2b+2c]$) was 1.14 (95% CI: 0.98, 1.30) times higher in the

intervention compared with the control arm, supporting the possibility of ascertainment bias. We estimated that this bias would inflate the total number of intervention events by about 23 (6%) from 382 to 405 and inflate the hazard ratio from the hypothesized 0.80 to 0.86 (see Greene et al.¹⁰ for details). Thus, the net effect of ascertainment bias would be to dilute the hazard ratio and inflate the number of events (and event rate) in the intervention arm (original protocol definition) or preserve the hazard ratio but reduce the total number of events due to the omission of Type 2b events (revised definition).

Projected power based on accumulated STRIDE trial data

We calculated projected power¹⁰ using the accumulated study data as of February 22, 2018, including the aggregate annual loss to follow-up and death rates, annual control patient-reported event rate, estimated adjudication confirmation rate, and estimated cluster design effect (Table 2).

The cluster design effect (i.e., variance inflation factor) measures the inflation of the variance of the intervention effect that results from correlated observations in the cluster randomized design. Because there were no closed form estimates of the cluster design effect based on an intracluster correlation coefficient for clustered time-to-event data with a competing risk, we estimated the variance inflation factor for the STRIDE trial by comparing the variances of the intervention effect from a lognormal frailty model accounting for clustering to the standard Cox model without adjustment for clustering. Estimates of variance inflation at the time of data lock were small for both the protocol (1.00) and revised (1.05) outcome definitions (Table 2). The respective upper 95% confidence intervals estimated by the bootstrap were slightly greater than 1.00 and 1.18, respectively.

In the calculations, we approximated the adjudicated control event rate by adjusting the observed patient-reported event rate by the proportion of patient-reported events that were adjudicated as primary outcome events. The confirmation rate for the original protocol primary outcome definition was 84.7% vs. 90.3% for the adapted definition. The adjudication confirmation rate was slightly lower for the original primary outcome definition because Type 2b events, which are only part of the original definition, were less likely to be adjudicated as serious and often required ascertainment of medical records for out-of-system events and so were more difficult to confirm.

Table 2 shows the projected power for the original protocol and revised definitions and the parameter estimates used for these projections. Power for the revised definition (89%) was greater than the power for the original protocol definition (78%), meaning that the change in the hazard ratio (from the original hypothesized 0.80 to the bias-adjusted 0.86) had a greater effect on power than the reduction in the annual control event rate under the revised definition (0.148 vs. 0.089).

In a sensitivity analysis, we determined the effect on power under both outcome definitions for a range of hypothesized hazard ratios, ascertainment biases, variance inflation factors, and adjudication confirmation fractions. In all cases, the original protocol definition was more sensitive than the adapted definition to inflation of the hazard ratio, variance inflation

factor and ascertainment bias, and to lower adjudication confirmation rates (data not shown, see Greene et al.¹⁰)

Discussion

In unblinded randomized controlled trials, special care is needed to ensure that the outcome is determined in an unbiased manner. Ascertainment bias can manifest in many ways and can influence the differential ascertainment of outcomes in randomized arms, potentially leading to dilution of intervention effects. Potential for ascertainment bias arose in the STRIDE trial.

During the design phase of STRIDE, a broad definition for serious fall injury was adopted, with seeking medical care a marker of a definitively ‘serious’ injury. However, as specific details about the intervention became apparent, concerns were raised about ascertainment bias because the falls care managers could refer participants to seek medical care who might not have otherwise done so, potentially inflating the number of events in the intervention arm and causing spurious dilution of the intervention effect. This concern was not based on examination of trial data, but rather on the definition of the primary outcome. Concern focused specifically on fall injuries for which seeking medical attention might be discretionary. Because the control arm was not subject to this referral bias, the trial leadership considered adapting the primary outcome definition to exclude less severe injuries that were susceptible to this type of bias, namely those injuries not requiring an overnight hospitalization. Because of the broad definition of a serious fall injury, the interim monitoring plan approved by the DSMB included provisions for adapting the primary outcome definition, if necessary, blinded to primary outcome data by randomized group.

During the execution of the interim monitoring plan, the DSMB became uncomfortable with addressing the adaptation issue because even though they had not reviewed primary outcome data by intervention, they had been reviewing safety data on fall-related hospitalizations. Thus, they felt compromised (unblinded) and recommended to the sponsors that an independent ad hoc committee be convened to review the data and make a recommendation to the sponsors about adapting the primary outcome definition. The sponsors concurred with this recommendation and convened an independent committee. Based on the accumulated trial data, the committee recommended changing the primary outcome definition, and the sponsors accepted the recommendation.

Because there was no direct way to determine the presence and magnitude of ascertainment bias in STRIDE, we developed an indirect approach. We presumed that the falls care managers could refer certain types of less severe injuries (Type 2c) for medical attention, inflating the number of Type 2b injuries requiring medical attention in the intervention arm. Thus, we inferred that referral bias could be detected as an increase in the proportion of Type 2b injuries relative to Type 2b and 2c injuries combined in the intervention versus control arm. We had to rely on patient-reported events for this determination because Type 2c injuries were not being adjudicated and because there was a considerable lag in the adjudication process. We recognized, however, that study power would be reduced with an adapted primary outcome definition since the overall number of events in the trial would be

diminished. Thus, there was a trade-off between dilution of the intervention effect under the original protocol definition and a reduced number of events under the revised definition. It was also unclear how much ascertainment bias was needed to reduce the power of the trial, which led to examining trial power under both definitions. Based on the interim monitoring results, we concluded that the adapted primary outcome provided more assurance that power would be maintained despite generating fewer total events.

When designing pragmatic clinical trials, we recommend that investigators carefully consider the possibility of outcome ascertainment bias. This could be addressed by developing a more stringent definition of the outcome, or as was done in STRIDE, by adapting the outcome definition if interim data indicate that ascertainment bias may threaten the trial. Although the revised definition has better power in the majority of the situations considered for STRIDE,¹⁰ we recommend evaluating the impact of the bias, as well as the impact of revising the definition (Figure 1) and using an independent body to make the decision on which outcome definition to use in the trial blinded to treatment. This is a critical issue since one of the major concerns raised about pragmatic clinical trials is that treatment effects are often smaller than hypothesized and, thus, require larger sample sizes to detect smaller effects.¹¹ Ascertainment bias could play a role in the dilution of treatment effects and may be an issue for trials in which participants by nature of the intervention have increased interactions with the healthcare system in ways that could introduce bias in the ascertainment of outcomes (i.e., behavioral interventions).

Although we presumed a one-directional ascertainment bias because the falls care managers could refer participants to seek medical attention if they reported a fall, it is possible for the bias to go in the opposite direction. Because of their knowledge of intervention assignment, the falls care managers could have advised participants not to seek medical attention for fall-related injuries. However, this is not consistent with good clinical practice, our measure of ascertainment bias supported this presumption.

There are several limitations to this study. First, we used patient-reported events because of the lag in adjudication and because Type 2c events were not adjudicated (did not lead to medical attention). Whether we would have obtained the same results if all events had been adjudicated is uncertain. However, because STRIDE was a pragmatic trial, adjudicating all events would have been impractical. Second, we assumed the intervention affected all events equally. It is possible that the intervention operated selectively on the different types of events, i. e., the effect could be greater for the more severe events. Third, the power calculations were based on many assumptions and approximations (i.e., the event rate, the hypothesized intervention effect, the impact of the ascertainment bias, the rate of conversion of self-reported events to adjudicated events and variance inflation factor), and the estimates based on the interim data could prove unstable.¹²

In summary, we used a prespecified, interim analysis to adapt the primary outcome in STRIDE based on the presumption of ascertainment bias that was supported by the available data. Although we can never be completely certain about the magnitude of the bias, adapting the primary outcome mitigated concerns about this type of bias. The issues in STRIDE may arise in other unblinded pragmatic trials that use care managers (e.g., dementia care

managers) to deliver an intervention in which patient contact may increase the likelihood of referral for medical attention that is directly linked to the outcome. This problem is well known in the psychiatric/psychological literature and is a reason that controlling for contact time in the usual care arm is often done as part of the study design. However, this is not a practical solution for pragmatic trials with a usual care comparator arm that aims to mimic clinical practice and provide generalizable results that can be readily implemented into practice.

Acknowledgments

Funding source: The STRIDE study was funded primarily by the Patient Centered Outcomes Research Institute (PCORI), with additional support from the National Institute on Aging (NIA) at NIH (5U01AG048270). Dr. Gill received support from P30AG021342 and K07AG043587.

References

1. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from <http://handbook.cochrane.org>.
2. Bhasin S, Gill TM, Reuben DB, et al. A randomized trial of a multifactorial fall injury prevention strategy. *N Engl J Med* 2020; 383(2):129–140. [PubMed: 32640131]
3. Bhasin S, Gill T, Reuben D, et al. Strategies to Reduce Injuries and Develop Confidence in Elders (STRIDE): A cluster-randomized pragmatic trial of a multifactorial fall injury prevention strategy: Design and methods. *J Gerontol A Biol Sci Med Sci* 2018; 73(8):1053–1061. [PubMed: 29045582]
4. Reuben DB, Gazarian P, Alexander N, et al. The Strategies to Reduce Injuries and Develop Confidence in Elders Intervention: Falls risk factor assessment and management, patient engagement, and nurse co-management. *J Am Geriatr Soc* 2017; 65(12), 2733–2739. [PubMed: 29044479]
5. Centers for Disease Control and Prevention’s STEADI toolkit: <https://www.cdc.gov/steady/training.html>
6. Ganz DA, Siu AL, Magaziner J, et al. Protocol for serious fall injury adjudication in the Strategies to Reduce Injuries and Develop Confidence in Elders (STRIDE) Study. *Inj Epidemiol* 2019; 6:14. Doi: 10.1186/s40621-019-0190-2. [PubMed: 31245263]
7. Schoenfeld D Sample-size for the proportional hazards regression model. *Biometrics* 1983; 39:499–503. [PubMed: 6354290]
8. Hintze J (2013). PASS 12. NCSS, LLC. Kaysville, Utah, USA. www.ncss.com.
9. Pahor M, Guralnik JM, Ambrosius WT, et al. Effect of structured physical activity on prevention of major mobility disability in older adults: the LIFE study randomized clinical trial. *J Am Med Assoc* 2014; 311(23):2387–2396.
10. Greene EJ, Peduzzi P, Dzuira J, et al. Estimation of ascertainment bias and its effect on power in clinical trials with time-to-event outcomes. arXiv:2004.13775 [stat.ME] (2020). <https://arxiv.org/abs/2004.13775>
11. Patsopoulos NA. A pragmatic view on pragmatic trials. *Dialogues in clinical neuroscience*. 2011;13(2):217–224. [PubMed: 21842619]
12. Esserman D From screening to ascertainment of the primary outcome using electronic records: challenges in the STRIDE Trial. *Clin Trials* 2020; 17(4): 346–350. [PubMed: 32408769]

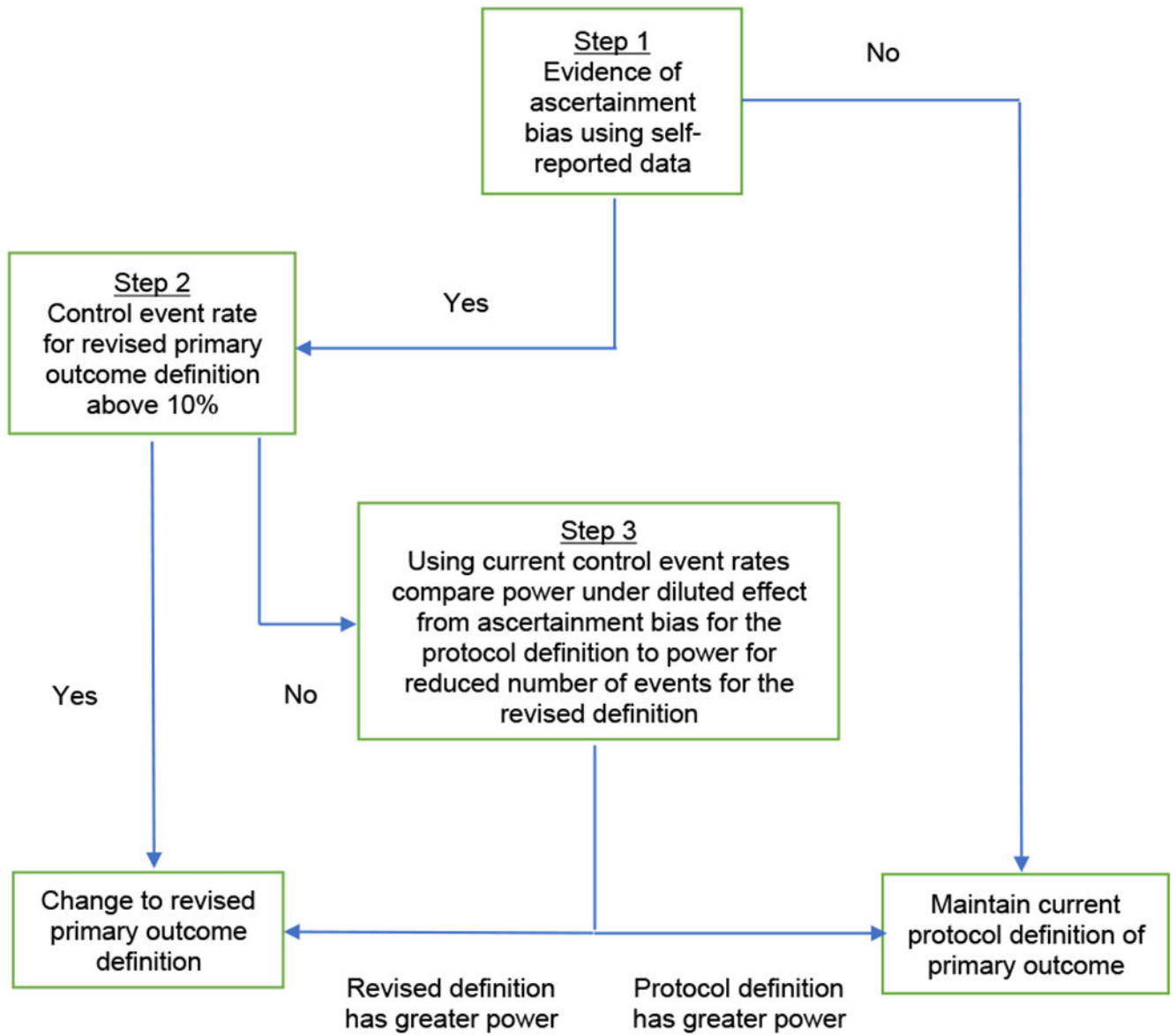


Figure 1. Interim monitoring plan decision algorithm for assessing ascertainment bias and adapting the primary outcome.

Table 1:

Number of self-reported fall injuries in STRIDE trial by fall classification and intervention category at the time of the outcome adaptation (February 22, 2018)

Time	Treatment Arm	Type 1 ^a	Type 2a ^b	Type 2b ^c	Type 2c ^d	Total
February 22, 2018	Intervention	227	65	263	526	1081
	Usual Care	269	76	253	613	1211

^aType 1: fracture other than thoracic/lumbar vertebral; joint dislocation; or cut requiring closure;

^bType 2a: head injury; sprain or strain; bruising or swelling; or other requiring medical attention *and an* overnight hospitalization.

^cType 2b: head injury; sprain or strain; bruising or swelling; or other requiring medical attention *but not an* overnight hospitalization.

^dType 2c: head injury; sprain or strain; bruising or swelling; or other not requiring medical attention.

Table 2.

Projected power under original protocol definition vs. adapted (revised) definition based on data as of February 22, 2018 for a type I error of 5% (2-sided).*

Parameters	Original Protocol Definition	Revised Definition
Annual enhanced usual care event rate based on patient-reported events	0.148	0.089
Annual aggregate death rate	0.022	0.022
Annual aggregate lost to follow-up rate	0.019	0.019
Adjudication confirmation fraction	0.847	0.903
Cluster design effect (VIF)	1.00	1.05
Hazard ratio	0.86	0.80
Projected power	78%	89%

* Rates were calculated as number of events divided by person years of follow-up. Hazard ratio for original definition is adjusted for ascertainment bias vs. hypothesized hazard ratio for revised definition. Adjudication confirmation fraction is the proportion of patient-reported events that were adjudicated as confirmed serious fall injuries. VIF, variance inflation factor.