Published in final edited form as:

Clin Trials. 2016 October ; 13(5): 537–544. doi:10.1177/1740774516643689.

A Post-Hoc Evaluation of a Sample Size Re-estimation in the SPS3 Study Running head: Evaluating the SPS3 Sample Size Re-Estimation

Leslie A McClure¹, Jeff M Szychowski¹, Oscar Benavente², Robert G Hart³, and Christopher S Coffey⁴

¹Department of Biostatistics, University of Alabama at Birmingham, Birmingham, Alabama, USA

²Department of Neurology, University of British Columbia, Vancouver, British Columbia, Canada

³Department of Medicine, McMaster University, Hamilton, Ontario, Canada

⁴Department of Biostatistics, University of Iowa, Iowa City, Iowa, USA

Abstract

Background/Aims—The use of adaptive designs has been increasing in randomized clinical trials. Sample size re-estimation is a type of adaptation in which nuisance parameters are estimated at an interim point in the trial and the sample size re-computed based on these estimates. The Secondary Prevention of Small Subcortical Strokes (SPS3) study was a randomized clinical trial assessing the impact of single vs. dual antiplatelet therapy and control of systolic blood pressure to a higher (130-149 mmHg) vs. lower (<130 mmHg) on recurrent stroke risk in a two-by-two factorial design. A sample size re-estimation was performed during the SPS3 study resulting in an increase from the planned sample size of 2500 to 3020, and we sought to determine the impact of the sample size re-estimation on the study results.

Methods—We assessed the results of the primary efficacy and safety analyses with the full 3020 patients and compared them to the results that would have been observed had randomization ended with 2500 patients. The primary efficacy outcome considered was recurrent stroke, and the primary safety outcomes were major bleeds and death. We computed incidence rates for the efficacy and safety outcomes and used Cox proportional hazards models to examine the hazard ratios for each of the two treatment interventions (i.e. the antiplatelet and blood pressure interventions).

Results—In the antiplatelet intervention, the hazard ratio was not materially modified by increasing the sample size, nor did the conclusions regarding the efficacy of mono vs. dual-therapy change: there was no difference in the effect of dual vs. mono therapy on the risk of recurrent

Corresponding Author: Leslie McClure, 1720 2nd Ave South, Birmingham, AL 35294-0022, lmcclure@uab.edu, Phone: (205) 934-5924.

Clinical Trials Registration Information

Clinical Trials Registration — <http://clinicaltrials.gov/show/NCT00059306>.

Unique identifier: NCT00059306.

Declaration of competing interests

The authors have no conflicting interests to declare.

stroke hazard ratios (n=3020 (95% CI): 0.92 (0.72, 1.2) p=0.48; n=2500 (95% CI): 1.0 (0.78, 1.3), p=0.85). With respect to the blood pressure intervention, increasing the sample size resulted in less certainty in the results, as the hazard ratio for higher vs. lower systolic blood pressure target approached, but did not achieve, statistical significance with the larger sample (n=3020 (95% CI): 0.81 (0.63, 1.), p=0.089; n=2500 (95% CI): 0.89 (0.68, 1.17), p=0.40). The results from the safety analyses were similar with 3020 and 2500 patients for both study interventions. Other trial-related factors, such as contracts, finances and study management were impacted as well.

Conclusions—Adaptive designs can have benefits in randomized clinical trials, but do not always result in significant findings. The impact of adaptive designs should be measured in terms of both trial results, as well as practical issues related to trial management. More post-hoc analyses of study adaptations will lead to better understanding of the balance between the benefits and the costs.

Keywords

Adaptive designs; sample size re-estimation; SPS3; randomized clinical trials

Introduction

Adaptive designs address uncertainty during study planning by allowing a review of accumulating information during an ongoing clinical trial to possibly modify one or more trial design characteristics. The greater flexibility provided by adaptive designs may lead to improved efficiency through the use of fewer patients within trials and an increased chance of a “successful” trial that answers the primary question of interest. For these reasons, the use of adaptive designs has been growing in randomized clinical trials, with implementation ranging as broadly as biomarker adaptive designs in earlier phase cancer trials¹ to later phase trials of chronic diseases and surgical trials,²⁻⁴ as well as for comparative effectiveness trials.⁵ In fact, the recent Ebola outbreak in Africa has raised the issue of the ethical dilemma of randomized trials, and whether there are ethical reasons to use adaptive trial designs in situations such as this.^{6, 7} Further, the April 2015 edition of *Clinical Trials* contains an editorial about the ethical concerns of adaptive randomization,⁸ and lengthy commentary on the ethics of outcome-adaptive allocation.⁹⁻¹⁶ A PubMed search of “adaptive designs” yields over 800 manuscripts, many of which have been published within the last 10 years, indicating the recent increasing attention to adaptive trial designs. The increased interest in adaptive designs, and wide range of potential adaptations, has generated a lot of confusion regarding what types of adaptations are acceptable and unacceptable. Some recent published definitions that have appeared in the literature have helped to clarify this issue. An Adaptive Design Scientific Working Group published a white paper that provided the first published definition of an adaptive design,¹⁷ and provided three important aspects that must be considered: 1) the adaptations must be made “J.without undermining the integrity and validity of the trial”, 2) “changes are made by design, and not on an ad-hoc basis”, and 3) adaptive designs are “Jnot a remedy for inadequate planning”. Two recent FDA draft guidelines^{18, 19} have provided similar definitions that support the notion of “Adaptive By Design”. This is an important point because properly designed simulations are often needed to confirm whether an adaptation preserves the integrity and validity of the

overall study. In order to properly define the required simulation study, the adaptation rules must be clearly specified in advance. Thus, only planned adaptations can be guaranteed to avoid any unknown bias due to the adaptation. While the Working Group definition requires adaptations to be planned prior to the initiation of the trial, both FDA guidance documents allow a bit more flexibility and allow for planned adaptation to be implemented after the initiation of the trial provided that key trial decision makers have had no access to unblinded data (i.e., the proposed adaptations are not driven by unblinded outcome data observed in the study).

A sample size re-estimation design refers to an adaptive design that allows for an adjustment of sample size based on a review of the interim data.²⁰⁻²² This allows for some error in the original specification of parameters utilized in the a priori sample size determination, as often little data are available upon which to base the estimates of these parameters. Even small amounts of misspecification can result in sample sizes that are too small or too large, leading to under- or over-powered studies.²⁰ Sample size re-estimation allows adjusting the final sample size up or down in order to ensure sufficient power. There are several subtypes of sample size re-estimation, depending on whether the estimated treatment effect only, estimated nuisance parameters (e.g., variance, overall event rate, accrual rate) only, or both are used for re-estimating sample size. A great deal of controversy exists with respect to designs that utilize treatment effect estimates at the time of re-estimation. On the contrary, internal pilot designs re-estimate sample size for a fixed treatment effect based on estimates of nuisance parameters derived from accumulating trial data.²²⁻²⁴ With moderate to large sample sizes, internal pilot designs can be used to make appropriate modifications with minimal (if any) inflation of type I error rate. Levin *et al* (2013) explore different approaches to designing clinical trials with pre-specified rules for modifying the sample size during the course of the study under theoretical conditions,²⁵ while others have investigated whether the statistical properties of a study are preserved when an unplanned study modification is implemented.^{17, 26} However, most of the literature examining the use of sample size re-estimation has addressed the prospective properties of the design. There have been few post-hoc examinations of studies utilizing sample size re-estimation in order to determine whether or not the changes to sample size practically affected the study results. In fact, Levin *et al* indicate “There is a need for studies that exactly quantify and discuss the relative costs and benefits of simple and easily implemented pre-specified adaptive designs compared with alternative designs in realistic settings.”²⁵ In this paper, we attempt to do just that, although in the context of an adaptive design that falls into a gray area with respect to whether or not it meets the FDA definition of being pre-specified. Herein, we compare the results of a large randomized clinical trial which implemented a sample size re-estimation to what “would have been” should the sample size not have been extended.

The SPS3 study

The Secondary Prevention of Small Subcortical Strokes (SPS3) study is a two-by-two factorial trial designed to examine the relationship between systolic blood pressure control and recurrent stroke and between antiplatelet therapy and recurrent stroke. Details of the trial have been published elsewhere.^{27, 28} In brief, patients suffering from an magnetic resonance imaging-verified lacunar stroke were randomized to receive either aspirin or aspirin plus

clopidogrel (double blind), and to a blood pressure lowering target of 130-149 mmHg vs. <130 mmHg (open label). Table 1 depicts the randomization scheme, including the number randomized into each study intervention. The original study design parameters were based on data from the WARSS study,²⁹ suggesting that the recurrent stroke rate among patients on aspirin with lacunar stroke was 7% annually. SPS3 was designed to detect a 25% relative risk reduction in recurrent stroke attributable to combination antiplatelet therapy (5.25% vs. 7%), with 90% power, 5% significance level, assuming a loss-to-follow-up rate of 3% per year, and adjustment for 2 planned efficacy and futility interim analyses. Given these assumptions, it was determined that a total of 417 events were required, and based on recruitment assumptions (35 sites, with patients accrued at a rate of 2 per site per month over 3 years) and follow-up assumptions (patients followed for 1 year after the last patient was enrolled), a total sample size of 2500 patients was necessary. The planned primary analysis was a time-to-event analysis using a Cox proportional hazards model to determine whether differences existed in either the antiplatelet treatments, or the blood pressure targets. SPS3 was approved by the institutional review board at each participating site, and each patient enrolled provided written informed consent.

As previously described,²⁶ slow rates of accrual and published recurrent stroke rates in other clinical trials well below those planned in SPS3 prompted a sample size re-estimation. Based on the results of extensive simulations examining the operating characteristics of the study under several difference scenarios, the study investigators proposed, and the funding agency approved, extending the sample size to 3000 patients, with follow-up concluding 1 year after the last patient was enrolled. The Data and Safety Monitoring Board and study sponsor accepted this proposal, and the study modification was implemented after lengthy discussions among study investigators. Thus, the internal pilot design utilized in this study falls into a gray area with respect to whether or not it meets the FDA definitions for being a prospective adaptation. On one hand, the study Principal Investigators requested the sample size re-estimation based on lower observed stroke rates from published studies that raised concerns regarding whether the original assumptions used to determine the sample size remained valid. Since this was not driven by any unblinded study results, this aspect of the design meets the FDA definition of a prospective adaptive design. However, since the Data Safety Monitoring Board (which had seen unblinded results) was asked to approve the design, this aspect of the design does not appear to meet the FDA definition. Since different individuals become unblinded at different time points during the conduct of a clinical trial, some clarification is needed to address situations like this in the FDA definitions.

Recruitment into SPS3 concluded in April 2011 with 3020 patients (20 greater than the actual target from the sample size re-estimation). Late in 2011, after all 3020 patients were recruited, the Data Safety Monitoring Board made the decision to terminate the antiplatelet intervention, due to safety concerns coupled with lack of efficacy, reducing the length of follow-up for this comparison. The blood pressure intervention concluded in April, 2012. The results of the SPS3 study, including all 3020 patients recruited, have been published.^{30, 31} Briefly, no significant differences were observed for the comparison of antiplatelet or blood pressure interventions, and minor safety concerns were found due to an increased risk of major bleeds and death among those on combination antiplatelet therapy versus aspirin alone. Since neither primary hypothesis provided positive results, the benefits

of increasing the sample size through the use of sample size re-estimation here could be questioned. However, because of the design of the study in which patients are enrolled on a rolling basis, we have the benefit of being able to “look back” to see what “might have been” should we have stopped recruitment after the planned sample size was attained (2500), and only followed those patients for 1 year after the 2500th patient was enrolled. This allows an examination of what was gained or lost by implementing the sample size re-estimation in the SPS3 study.

Methods

The SPS3 analysis plan has been described in detail previously.²⁷ In brief, both the antiplatelet and blood pressure effects were assessed using Cox proportional hazards models to determine whether there were differences in the event rates over time between those randomized to dual antiplatelet therapy vs. monotherapy, and those randomized to higher vs. lower blood pressure targets (see Table 1). During the antiplatelet analyses, we assessed an interaction with blood pressure treatments; in the absence of a significant interaction, analysis of each intervention (i.e. the blood pressure and the antiplatelet interventions) proceeded by assessing the marginal effects.

We compared the results from assessing both the antiplatelet and the blood pressure effects of the SPS3 study under originally planned sample size (2500) to those observed with the re-estimated sample size (3020). For the former, we exclude patients 2501-3020, and allowed follow-up to continue for one year after patient 2500 was enrolled for the blood pressure intervention, and through the early termination of the antiplatelet intervention in August 2011 (range of end-of-follow-up: September 2011 to November 2011). Patient 2500 was enrolled in November 11, 2010; thus, any event that occurred up to and including November 11, 2011 was included in the analyses. However, any events that occurred after that date were censored. For the latter, we included all events as described in the primary results, with the caveat that the early stopping for the antiplatelet intervention shortened the follow-up time planned during the sample size re-estimation for this comparison.^{30, 31}

We examined both efficacy and safety outcomes, as well as secondary outcomes. As described above, the primary efficacy outcome was recurrent stroke, defined as either ischemic or hemorrhagic stroke. The primary safety outcomes were major bleeds and death. Secondary outcomes examined were myocardial infarction and major vascular events. In addition, we estimated the additional costs of extending the study based on patient costs.

Results

Table 2 and Figure 1 provide the results from the original design (n=2500) and the modified design (n=3020) when assessing the antiplatelet effect. The table provides the total number of events, average follow-up, event rate, hazard ratio and 95% confidence interval for the hazard ratio, and the p-value for the difference in the rates across the antiplatelet groups, whereas the Figure includes the hazard ratio and 95% confidence interval for the hazard ratio. As would be expected, the number of events increases in each group as more participants are followed for a longer duration. However, the average follow-up time is not

impacted by the increase in sample size, likely due to multiple factors: (1) the antiplatelet intervention was terminated early due to safety concerns, and (2) the addition of more participants resulted in adding some with very short follow-up times, thus the average changed little. No substantive differences were observed in the event rates, and the hazard ratio and 95% confidence intervals suggest that the conclusions drawn would be the same regardless of whether we increased the sample size.

Table 3 (Figure 1) presents the same information, only now for the blood pressure intervention. We similarly see that the number of events increased with the larger sample; however, in this case we also see an increase in the average follow-up time. This differs from above in that follow-up continued for a full year in this intervention after enrolling the last participant. The event rate decreased in the <130 mmHg group; whereas it increased in the 130-149 mmHg group, indicating that extending the sample size may have increased the likelihood of detecting a statistically significant difference. Comparing the hazard ratios suggests that the magnitude of the association did not change substantively; however, the bounds of the confidence interval changed with the addition of 520 participants. Had the study stopped after 2500, the conclusions would have clearly been made that there was no association between blood pressure control and risk of recurrent stroke ($p=0.40$). With the additional recruitment of 520 patients, the results were more supportive of benefit ($p=0.089$) although with a smaller estimated effect than that upon which the study was powered to detect.

Table 4 presents the safety results for each of the antiplatelet and blood pressure interventions, for each of the original design ($n=2500$) and the modified design ($n=3020$), also depicted in Figure 1. The table contains the hazard ratios and 95% confidence intervals for the major bleeds and the deaths. The statistical impact of extending the study was minimal on the safety outcomes, both for the antiplatelet and blood pressure interventions.

Supplementary tables S1a (antiplatelet intervention) and S1b (blood pressure intervention) provide the results for myocardial infarctions for the original design ($n=2500$) and the modified design ($n=3020$), and supplementary tables S2a (antiplatelet intervention) and S2b (blood pressure intervention) provide the same for major vascular events. As with both the primary outcomes and the safety outcomes, while we saw some shifts in the magnitude of the associations, the conclusions did not change substantively.

With respect to costs, each site received \$4000 per baseline visit, and then \$400 for each follow-up visit. Given that we recruited an additional 520 patients, and the sample size re-estimation led to an additional 2583 follow-ups, we can estimate that the additional patient costs to the study were approximately to be \$3,113,200. When examined based on the total cost of the study rather than just the patient care costs (including the medications, the cost of the study drugs distribution center, statistical center and coordinating center), we estimate that the cost per patient was approximately \$21,000. Thus we can, roughly determine that the additional costs for extending the recruitment was \$10,920,000.

Discussion

In this paper, we compared the primary results, results from the secondary outcomes and the safety results from the SPS3 trial for the observed sample (after increasing the sample size to 3020 due to a sample size re-estimation) to those that would have been observed had we recruited only the planned sample of 2500 patients. We found that in this case the extension of the sample made little difference with respect to the antiplatelet therapy and to the safety results (for either intervention), but that the primary results from the blood pressure intervention approached significance with the larger sample size whereas they were clearly non-significant under the original assumptions. No substantive differences were observed in the secondary outcomes. This dramatically affected the writing and interpretation of the primary blood pressure manuscript. On one hand, it could be argued that stopping at the original planned sample size might have missed an interesting result since the blood pressure intervention results would have been interpreted as a clearly 'negative' trial. On the other hand, it could be argued that the marginally significant result was observed for an estimated effect less than the 25% target effect. Thus, both situations led to a non-significant result for the blood pressure intervention and provide pretty clear evidence against the existence of a 25% or larger reduction in the risk of recurrent stroke associated with more intensive blood pressure treatment. Whether the ability to determine the potential existence of an effect less than 25% was worth the extra cost and complexity of the larger sample size is an interesting question. These are the types of questions that future clinical trialists will need to address when considering the use of sample size re-estimation designs. By modifying sample size based on estimated nuisance effects for pre-specified effects of interest, internal pilot designs can clearly achieve their primary benefit or providing greater power for detecting these effects when they truly exist. However, this may lead to unnecessary increases in sample size for studies where the pre-specified effect is not obtained (i.e., the study leads to a non-significant result). Unfortunately, as much as we would like to assume otherwise, experience suggests that a large number of clinical trials will not show significant results. Hence, since it is impossible to know whether or not a trial will eventually lead to a positive result or not, further discussion regarding the utility of sample size re-estimation designs in trials is needed to address this concern and avoid the unnecessary expenditure of vital resources.

While the conclusions from the study did not substantively change after the implementation of the sample size re-estimation, we did see some shift in the magnitude of the effects, for both the antiplatelet intervention and the blood pressure intervention. We examined trends over time in each intervention in order to assess the proportional hazards model, and the results were not statistically significant ($p=0.49$ for antiplatelet intervention; $p=0.72$ for the blood pressure intervention). Additionally, we examined the Kaplan-Meier curves for the first 2500 vs. the additional 520 and observed no differences in trends. Thus, it is likely that the shifts in the magnitude of the effects were random noise.

Levin *et al* (2013) describes the optimal adaptive designs under theoretical conditions for different types of adaptations of the maximal sample size. They show that there are small gains in efficiency in average sample number when a pre-specified adaptation is implemented in lieu of the group sequential design, and suggest that the gains in efficiency may not be worth the trade-off in terms of the logistical and ethical considerations of a trial,

relative to group sequential designs, but do suggest that simple adaptations may have merit. Our experience indicated that the overall conclusions did not change after the adaptation was implemented in SPS3, and while the evidence in support of a difference in the effect of blood pressure target range on recurrent stroke was marginally significant, it also did not attain the anticipated 25% reduction in risk that the study was powered to detect. It is possible that had the decision been made to increase the sample size beyond what was recommended, or that the follow-up time was increased beyond what was planned, the conclusions of the blood pressure trial may have been different; however, it is also possible that the shift in the magnitude of effect towards statistical significance was due to chance.

Recently, Bauer *et al* provided a discussion of the use of adaptive designs over the last 25 years,³² in which they described a use of a sample size re-estimation procedure that included an unblinded interim analysis and resulted in a reduction in sample size.³³ Although the authors of this paper were not able to compare results with and without the adaptation, this is a setting in which it is likely that use of the modified sample size resulted in similar results more quickly than would have occurred otherwise, due to the reduction in sample size as per the sample size re-estimation.

Further, the Bauer *et al* paper provides real-life experiences from other types of studies implementing adaptive designs (e.g. adaptive treatment selection), and what considerations need to be made other than statistical properties.³² Modification of a study mid-stream potentially could require changes to the protocol, informed consent, staffing, and numerous other logistical items. Additionally, changes to the study mid-stream may have implications for the randomization scheme, the data entry system, the case report forms, and other aspects of data management. Thus, consideration of these issues should be taken together with the statistical results. In the context of the SPS3 trial, increasing the sample size impacted each of these issues. For example, informed consent at some study sites expired after 5 years, and thus there were some issues with having to re-consent patients. Forms which recorded the study visit had to be modified to allow for additional visits, requiring commensurate changes to the data entry system. Staff contracts had to be extended and for some sites, additional institutional review board approval was needed. Negotiations were necessary with the company providing study drug and additional funding was required to continue distributing medications through the drug distribution center at the Veterans Affairs Cooperative Studies Center. Thus, in the SPS3 study, the operational impacts of increasing sample size were fairly substantial. Given the fact that the primary study results were not substantially altered with the increased sample size, one could question whether this added complexity was necessary. Therefore, when assessing the benefit of adaptation to a study, it is also essential to assess the additional operational complexity, additional regulatory issues, and cost.

Adaptive designs are playing larger and larger roles in clinical trials; however, there is little description of the impact of adaptations on study results, largely because of the difficulty in ascertaining what might have happened had the adaptation not been made. Post-hoc evaluations of adaptive designs are important to assessing the costs and benefits of the adaptation. In this paper, we have shown that when extending the recruitment in the SPS3 trial, the conclusions derived from the study did not change with the additional 520 patients although the strength of the association for the blood pressure intervention was increased;

whereas there was little impact on the magnitude of association for the antiplatelet intervention or on the safety analyses. While adaptive designs have great benefits in many situations, when using adaptive designs one should consider the statistical impact on the trial, as well as the impact on other trial aspects in order to assess the costs, together with the benefits, of the adaptation.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding

This work is supported by the National Institutes of Health/National Institute of Neurologic Diseases and Stroke (NIH/NINDS) U01 NS0385929.

References

1. Simon R. Biomarker based clinical trial design. *Chin Clin Oncol*. 2014; 3:39. [PubMed: 25841465]
2. Connor JT, Broglio KR, Durkalski V, et al. The Stroke Hyperglycemia Insulin Network Effort (SHINE) trial: an adaptive trial design case study. *Trials*. 2015; 16:72. [PubMed: 25885963]
3. Grieve AP, Chow SC, Curram J, et al. Advancing clinical trial design in pulmonary hypertension. *Pulm Circ*. 2013; 3:217–225. [PubMed: 23662200]
4. Pullenayegum EM. Adaptive Bayesian randomized trials: realizing their potential. *J Bone Joint Surg Am*. 2012; 94(Suppl 1):29–33. [PubMed: 22810444]
5. Gajewski BJ, Berry SM, Quintana M, et al. Building efficient comparative effectiveness trials through adaptive designs, utility functions, and accrual rate optimization: finding the sweet spot. *Stat Med*. 2015; 34:1134–1149. [PubMed: 25640114]
6. Caplan AL, Plunkett C, Levin B. Selecting the right tool for the job. *Am J Bioeth*. 2015; 15:4–10. [PubMed: 25856591]
7. Lanini S, Zumla A, Ioannidis JP, et al. Are adaptive randomised trials or non-randomised studies the best way to address the Ebola outbreak in west Africa? *Lancet Infect Dis*. 2015; 15:738–745. [PubMed: 25881871]
8. Begg CB. Ethical concerns about adaptive randomization. *Clin Trials*. 2015; 12:101. [PubMed: 25755226]
9. Hey SP, Kimmelman J. Are outcome-adaptive allocation trials ethical? *Clin Trials*. 2015; 12:102–106. [PubMed: 25649106]
10. Berry DA. Commentary on Hey and Kimmelman. *Clin Trials*. 2015; 12:107–109. [PubMed: 25649111]
11. Lee JJ. Commentary on Hey and Kimmelman. *Clin Trials*. 2015; 12:110–112. [PubMed: 25649110]
12. Saxman SB. Commentary on Hey and Kimmelman. *Clin Trials*. 2015; 12:113–115. [PubMed: 25649109]
13. Joffe S, Ellenberg SS. Commentary on Hey and Kimmelman. *Clin Trials*. 2015; 12:116–118. [PubMed: 25649113]
14. Buyse M. Commentary on Hey and Kimmelman. *Clin Trials*. 2015; 12:119–121. [PubMed: 25649112]
15. Korn EL, Freidlin B. Commentary on Hey and Kimmelman. *Clin Trials*. 2015; 12:122–124. [PubMed: 25649108]
16. Hey SP, Kimmelman J. Rejoinder. *Clin Trials*. 2015; 12:125–127. [PubMed: 25649107]

17. Gallo P, Chuang-Stein C, Dragalin V, et al. Adaptive designs in clinical drug development--an Executive Summary of the PhRMA Working Group. *J Biopharm Stat.* 2006; 16:275–283. discussion 285-291, 293-298, 311-312. [PubMed: 16724485]
18. US Food and Drug Administration. US Food and Drug Administration Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics (draft). FDA. , editor. FDA;
19. US Food and Drug Administration. US Food and Drug Administration Guidance for Industry: Adaptive Designs for Medical Device Clinical Studies. FDA;
20. Lake S, Kammann E, Klar N, et al. Sample size re-estimation in cluster randomization trials. *Stat Med.* 2002; 21:1337–1350. [PubMed: 12185888]
21. Gould AL. Planning and revising the sample size for a trial. *Stat Med.* 1995; 14:1039–1051. discussion 1053-1055. [PubMed: 7569499]
22. Proschan MA. Sample size re-estimation in clinical trials. *Biom J.* 2009; 51:348–35. [PubMed: 19358221]
23. Wittes J, Brittain E. The role of internal pilot studies in increasing the efficiency of clinical trials. *Stat Med.* 1990; 9:65–71. discussion 71-72. [PubMed: 2345839]
24. Proschan MA. Two-stage sample size re-estimation based on a nuisance parameter: a review. *J Biopharm Stat.* 2005; 15:559–574. [PubMed: 16022163]
25. Levin GP, Emerson SC, Emerson SS. Adaptive clinical trial designs with pre-specified rules for modifying the sample size: understanding efficient types of adaptation. *Stat Med.* 2013; 32:1259–1275. discussion 1280-1282. [PubMed: 23081665]
26. McClure LA, Szychowski JM, Benavente O, et al. Sample size re-estimation in an on-going NIH-sponsored clinical trial: the secondary prevention of small subcortical strokes experience. *Contemp Clin Trials.* 2012; 33:1088–1093. [PubMed: 22750086]
27. Benavente OR, White CL, Pearce L, et al. The Secondary Prevention of Small Subcortical Strokes (SPS3) study. *Int J Stroke.* 2011; 6:164–175. [PubMed: 21371282]
28. White CL, Szychowski JM, Roldan A, et al. Clinical features and racial/ethnic differences among the 3020 participants in the Secondary Prevention of Small Subcortical Strokes (SPS3) trial. *J Stroke Cerebrovasc Dis.* 2013; 22:764–774. [PubMed: 22516427]
29. Mohr JP, Thompson JL, Lazar RM, et al. A comparison of warfarin and aspirin for the prevention of recurrent ischemic stroke. *N Engl J Med.* 2001; 345:1444–1451. [PubMed: 11794192]
30. Benavente OR, Hart RG, McClure LA, et al. Effects of clopidogrel added to aspirin in patients with recent lacunar stroke. *N Engl J Med.* 2012; 367:817–825. [PubMed: 22931315]
31. Group SS, Benavente OR, Coffey CS, et al. Blood-pressure targets in patients with recent lacunar stroke: the SPS3 randomised trial. *Lancet.* 2013; 382:507–515. [PubMed: 23726159]
32. Bauer P, Bretz F, Dragalin V, et al. Twenty-five years of confirmatory adaptive designs: opportunities and pitfalls. *Stat Med.* 2016; 35:325–347. [PubMed: 25778935]
33. Zajicek JP, Hobart JC, Slade A, et al. Multiple sclerosis and extract of cannabis: results of the MUSEC trial. *J Neurol Neurosurg Psychiatry.* 2012; 83:1125–1132. [PubMed: 22791906]

Figure 1.
Comparing results from the primary outcomes, major bleeds and death across the two sample sizes, for each intervention

Table 1

Sample size by treatment allocations for the full sample (n=3020)

		Blood Pressure Control		Total N
		Higher (130-149 mmHg)	Lower (<130 mmHg)	
Combination Antiplatelet Therapy?	Yes	764	753	1517
	No	755	748	1513
	Total N	1519	1501	3020

Table 2

Results from the analyses of primary events from the observed sample (n=3020) and the planned sample (n=2500): Antiplatelet Intervention

	3020 Randomized Patients		First 2500 Randomized Patients	
	Aspirin + Clopidogrel	Aspirin + Placebo	Aspirin + Clopidogrel	Aspirin + Placebo
N	1517	1503	1253	1247
# Events	125	138	108	107
Average Follow-up Time (years)	3.3 (2.1)	3.4 (2.1)	3.3 (1.8)	3.4 (1.8)
Event Rate	2.5	2.7	2.6	2.5
HR (95% CI)	0.92 (0.72, 1.2)		1.0 (0.78, 1.3)	
p-value	0.48		0.86	

Table 3

Results from the analyses of primary events from the observed sample (n=3020) and the planned sample (n=2500): Blood Pressure Intervention

	3020 Randomized Patients		First 2500 Randomized Patients	
	< 130	130-149	< 130	130-149
N	1501	1519	1249	1251
# Events	125	152	102	113
Average Follow-up Time (years)	3.7 (2.1)	3.6 (2.1)	3.4 (1.8)	3.3 (1.8)
Event Rate	2.2	2.8	2.4	2.7
HR (95% CI)	0.81 (0.63, 1.0)		0.89 (0.68, 1.17)	
p-value	0.089		0.40	

Table 4

Results from the analyses of safety events from the observed sample (n=3020) and the planned sample (n=2500)

	3020 Randomized Patients HR ^a (95% CI)	First 2500 Randomized Patients HR ^a (95% CI)
Antiplatelet Intervention		
Major Bleeds	1.96 (1.41, 2.71)	2.08 (1.46, 2.96)
Death	1.52 (1.14, 2.04)	1.41 (1.03, 1.94)
Blood Pressure Intervention		
Major Bleeds	1.08 (0.800, 1.47)	1.17 (0.837, 1.64)
Death	1.03 (0.785, 1.35)	1.15 (0.841, 1.58)

^aHR represents the combination therapy relative to control for the antiplatelet intervention, and the lower vs. higher blood pressure target for the blood pressure intervention