Published in final edited form as:

Clin Trials. 2008 ; 5(1): 23–30. doi:10.1177/1740774507087781.

Use of dose modification schedules is effective for blinding trials of warfarin: evidence from the WASID study

Vicki Hertzberg^a, Marc Chimowitz^b, Michael Lynn^a, Cristen Chester^c, William Asbury^c, and George Cotsonis^a

^aDepartment of Biostatistics, Rollins School of Public Health, Emory University, Atlanta, GA

^bDepartment of Neurology, School of Medicine, Emory University, Atlanta, GA

^cEmory University Hospital, Atlanta, GA

Abstract

Background—Randomized clinical trials are blinded to prevent knowledge of treatment assignment from influencing outcomes and their assessments, thus protecting the trial’s scientific integrity. Trials involving a warfarin treatment arm are difficult to blind due to the need to continuously adjust dose.

Purpose—We sought to examine the effectiveness of blinding secondary stroke prevention trials with a warfarin treatment arm in which the blinding system incorporates use of placebo warfarin dose modification schedules for patients in the placebo warfarin arm.

Methods—We examined treatment assignment guesses of 569 patients or their next of kin as well as study coordinators and principal neurologists at the clinical sites in a multicenter, randomized, double-dummy, double-blinded clinical trial of warfarin and aspirin using dose adjustment schedules for management of placebo warfarin.

Results—Overall, the crude rates of correct responses are 60% for patient/proxy, 66% for study coordinator, and 56% for principal neurologist. Several indices were used to assess the consistency of guesses with what would be expected if the guessing were done completely at random, and all measures indicate adequate blinding.

Limitations—Comparison to other trials using warfarin is difficult due to limited data and differences in assessment of blinding. However, results compared favorably to one existing trial.

Conclusions—Placebo warfarin dose adjustment schedules can protect blinding adequately in trials involving warfarin.

Introduction

Good practice in design and conduct of randomized clinical trials (RCTs) calls for blinding of treatment assignment whenever possible. Results regarding success of blinding are often not reported [1,2]. The CONSORT group [3] has included blinding among elements included in reporting randomized trial results. In particular they call for documenting if study subjects, interventionalists, and outcome assessors were blinded to treatment assignment and how success of blinding was ascertained.

Blinding of RCTs in which warfarin is an active treatment poses special challenges due to risk of hemorrhagic complications and need for frequent blood draws to measure International Normalized Ratio (INR) levels. Thus, many trials with a warfarin arm have not been blinded [4–6], however, a few stroke prevention trials evaluating warfarin have used blinding strategies. Two trials of warfarin in atrial fibrillation patients, the Veteran's Administration Stroke Prevention in Nonrheumatic Atrial Fibrillation (SPINAF) study [7] and the Canadian Atrial Fibrillation Anticoagulation (CAFA) study [8] used a team of blinded and unblinded investigators at each site. In the Warfarin-Aspirin Recurrent Stroke Study (WARSS) [9], all patients were held to the same schedule for clinic visits to draw blood, measure INR levels, and adjust placebo or active warfarin dose. INR values were transmitted daily from central blood lab to statistical analysis center, where they were processed such that placebo patient INRs were replaced by INR values consistent with a similar dose and duration of treatment with active warfarin. Results were subsequently sent to each clinic. To preserve blinding, some emergency notifications for (falsely) extremely elevated INR values for placebo warfarin patients were included, akin to emergency notification for (truly) extremely elevated INR values for active warfarin patients.

Of these three blinded warfarin trials, only one (SPINAF) [7] has reported on the success of blinding. The Warfarin-Aspirin for Symptomatic Intracranial Disease (WASID) study, funded by the National Institute of Neurological Diseases and Stroke (NINDS), was recently completed [10]. We report here on the success of the WASID blinding system.

Methods

Study patients and primary endpoint

Design of WASID is described in detail elsewhere [11]. Briefly, patients with transient ischemic attack (TIA) or nondisabling stroke attributed to 50–99% stenosis of a major intracranial artery were randomized centrally to receive either warfarin (target INR 2–3) or aspirin (1300 mg daily). The primary endpoint was stroke (ischemic or hemorrhagic) or vascular death.

Blinding system

The blinding system in WASID was similar to that used in the SPINAF [7] and CAFA [8] trials and is detailed elsewhere [11]. This system involved both blinded and unblinded investigators at each site. The blinded team consisted of a principal site neurologist, study coordinator, and treating neurologist. The principal neurologist was responsible for initial patient enrollment as well as routine care, including risk factor management, and follow-up examinations, including endpoint assessment. The study coordinator was responsible for monthly patient contact. The treating neurologist was called to care for the patient when a major medical problem occurred.

The unblinded investigator was a pharmacist or physician at that site who did not share office space with the blinded personnel and who had experience managing warfarin. The major role of the unblinded investigator was to adjust active or placebo warfarin dose for all patients at that site and decide when each patient's next INR should be done. Active warfarin dose was adjusted according to an algorithm that was adapted from the Stroke Prevention in Atrial Fibrillation (SPAF) III trial anticoagulation protocol for maintaining the INR in the range of 2–3 [5]. Placebo warfarin dose was adjusted according to a predetermined dose adjustment schedule chosen randomly from 20 schedules. Quest Clinical Trials Laboratory provided central determination of INRs.

Efforts to maintain the blind

Concealing INR results from the principal blinded neurologist, study coordinator, and patient was critical to maintaining blinding in WASID. The methods used are described in detail elsewhere [11]. INRs processed at Quest were faxed directly to the unblinded investigator and did not appear in the patient's chart. If an urgent medical condition arose, INRs were usually performed locally by other physicians and entered into the patient's chart, which we asked the study coordinator or principal neurologist not to review during the trial.

When a study patient developed a serious medical problem (e.g., a stroke), care was entrusted to the treating neurologist to ensure that the blinded neurologist and coordinator did not become unblinded. The treating neurologist could have access to local INRs needed for patient management (e.g., a major hemorrhage) and thus could become unblinded as a result, but this was necessary for patient safety.

Adjudication of endpoints

Endpoints were adjudicated centrally by blinded panels of expert neurologists and cardiologists who did not participate in WASID. At the time of suspected endpoint, the principal neurologist performed necessary patient evaluations. The unblinded investigator then sent all supporting documentation to the statistical center, blackening out any reference that might give away treatment assignment. Statistical center personnel again carefully reviewed this material prior to transmission to the adjudicators.

Method for establishing suspected treatment assignment

The study protocol called for the patient or her/his next of kin (if patient unable to do so) to guess her/his treatment assignment at the time of primary endpoint or study close out, whichever occurred first. In addition, the study coordinator and principal neurologist were instructed to make treatment assignment guesses (TAGs) for these patients. Each of the individuals (patient/proxy, study coordinator, principal neurologist) was required to choose from aspirin or warfarin; 'don't know' (DK) was not an option. Next each respondent was asked to indicate certainty of answer, with options being uncertain, slightly suspicious, highly suspicious, and certain. After this information was obtained, actual treatment assignment (ATA) was revealed to the study coordinator, principal neurologist, and patient/proxy at study close out.

Statistical methods

WASID was designed to enroll 403 patients per group to detect a difference between primary endpoint rates of 22% per 3 years for patients receiving warfarin versus 33% per 3 years for patients receiving aspirin. Randomization began in February, 1999. Upon recommendation of the monitoring committee, NINDS stopped WASID on July 18, 2003 because of concerns about safety of patients assigned to warfarin. At that point 569 patients had been randomized. Patients were seen for close out visits by September 1, 2003 and events occurring up to close out were included. The average length of follow-up time was 1.8 years.

Several indices were used to investigate the extent of disagreement between TAG and ATA. First we calculated Cohen's kappa, κ , [12] defined as:

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

where $P_O = (N_{11} + N_{22})/N$ defines the observed amount of agreement and $P_E = (N_{1.}N_{.1} + N_{2.}N_{.2})/N$ describes the expected amount of agreement and the N_{ij} are defined according to the top portion of Table 1. We interpret success of blinding as a function of κ using commonly accepted guidelines [13]. We initially calculated κ on all responses, and denote this as κ_{ALL} . To address the issue of uncertainty of guess, we created a new TAG, which took on values for aspirin or warfarin as the original TAG except when the certainty response was ‘uncertain,’ giving the configuration shown in the lower portion of Table 1. We then recalculated κ , this time based on only TAGs that were aspirin or warfarin, excluding the ‘uncertain’ values, denoting this κ_{CERT} .

Since κ can exhibit paradoxical behavior [14,15], we also calculated alternative blinding indices. We calculated BI, defined by James *et al.* [16] as:

$$BI = [1 + P_{DK} + (1 - P_{DK})\kappa_D]/2,$$

with

$$\kappa_D = \frac{P_{D_o} - P_{D_e}}{P_{D_e}}$$

where $p_{ij} = n_{ij}/N$ as in the lower portion of Table 1,

$$P_{DK} = n_{0.}/N,$$

$$P_{D_o} = \sum_{i=1}^k \sum_{j=1}^k w_{ij} p_{ij} / (1 - P_{DK})$$

when $P_{DK} \neq 1$ and

$$P_{D_e} = \sum_{i=1}^k \sum_{j=1}^k w_{ij} p_{i.}(p_{.j} - p_{0j}) / (1 - P_{DK})^2.$$

P_{DK} denotes the proportion of DK responses. BI takes on values between 0 and 1, equaling 1 when there is complete blinding (defined as all ‘don’t know’), 0 when there is total lack of blinding (i.e., everyone guesses correctly), and 0.5 when agreement between guess and assignment appears to be purely due to chance. We also calculated newBI_j [17] where $j = 1$ or 2 indicating aspirin or warfarin. The index newBI_j is defined as:

$$\text{newBI}_j = 2(\widehat{r}_{jj}) \times (n_{1j} + n_{2j}) / (n_{1j} + n_{2j} + n_{0j})$$

where

$$\widehat{r}_{jj} = \frac{n_{jj}}{(n_{1j} + n_{2j})}.$$

The index newBI_j takes on values between -1 and 1 , equaling 1 when there is complete lack of blinding of the j th arm (all guesses are correct), -1 when there is opposite blinding of the j th arm (all guesses are incorrect), and 0 when there is perfect blinding, that is, agreement between guess and assignment appears to be purely due to chance for the j th arm.

We suspected *a priori* that some factors associated with treatment (number of warfarin dose changes, occurrence of hemorrhage, and occurrence of gastrointestinal (GI) upset) might be related to TAG and that such associations might vary according to ATA. Contingency tables and logistic regression with TAG as the dependent variable were used to investigate factors associated with TAGs while controlling for ATA.

Results

Study population

Patient characteristics have been previously reported [10]. Aspirin and warfarin groups were comparable with respect to demographic characteristics, medical history, concurrent risk factors, and intracranial disease characteristics at baseline. Patients were followed for a mean of 1.8 years. Only 2.3% (13) of patients did not complete follow-up, with 6 lost and 7 withdrawing consent. Study medications were discontinued permanently in 128 patients (22%) after an average of 0.9 years, with a significantly higher proportion among warfarin patients (28%) versus aspirin patients (16%) ($p = 0.0006$).

Overall assessment of agreement

The vast majority (89% for patient/proxy, 91% for study coordinator, 91% for blinded neurologist) of blinding surveys were completed during the study close out period. Patient/proxy, study coordinator, and principal neurologist TAG and ATA cross-classifications are shown in Table 2. Overall crude rates of 47% of patient guesses, 64% of study coordinator guesses, and 53% of principal neurologist guesses were correct when guesses were made. Patients were more likely to guess warfarin, while study coordinators and principal neurologists were more likely to guess aspirin.

Assessment of blind

Blinding indices for patient/proxy, study coordinator, and principal neurologist are shown in Table 2. Values of κ_{ALL} are well below 0.4, commonly accepted as consistent with agreement purely due to chance for patients, study coordinator, and principal neurologist. Although values for κ_{CERT} are higher than values for κ_{ALL} , they still fall short of 0.7, commonly accepted as consistent with good agreement.

Examination of BI gives a slightly different picture. The 95% confidence intervals (CIs) exceed the value of 0.5 for BI values for all respondents, with principal neurologist having the highest BI value. However, the 95% CIs do not approach 1, the BI value indicating complete blinding.

Inspection of newBI_j reveals that patients/proxies in the aspirin group appear to be guessing the assignment randomly. For all respondents, newBI_A is consistent with random guessing for aspirin patients, while newBI_W suggests a modest amount of unblinding for warfarin patients.

Other factors associated with treatment assignment guesses

Associations of various patient characteristics with patient/proxy, study coordinator, and principal neurologist TAGs were investigated. Occurrence of hemorrhage and number of warfarin dose changes were most consistently associated with TAG, particularly for study

coordinator (Table 3). In general, experience of hemorrhage was not associated with patient/proxy, site coordinator, or principal neurologist TAG among aspirin patients, but it was positively associated with a TAG of warfarin by all respondents among warfarin patients. A TAG of warfarin was positively associated with number of warfarin dose changes for patients in both treatment groups and for TAGs made by all three respondents.

Discussion

The blinding system used in WASID appears to be moderately successful, although not completely so. Since newBI_j can be interpreted as percent of unblinding, we have minimal (<10%) in the aspirin arm, with more (ranging from 12 to 39%) in the warfarin arm. Of the individual roles assessed, study coordinator guesses tended to be more correct than either patient/proxy or principal neurologist values, with principal neurologist guesses being least correct. This ordering might be anticipated. Study coordinators are routinely dealing with a variety of patient management issues for several patients. Principal neurologists, however, are less directly involved in day-to-day care delivery than study coordinators. Lack of correctness of their guesses is heartening, since they provide blinded determinations of patient endpoint.

Number of warfarin dose changes, occurrence of patient hemorrhage, and occurrence of potential unblinding events are most closely related to corruption of blinding. There is little that can be done to prevent unblinding due to occurrence of hemorrhage and potential unblinding events. However, number of warfarin dose changes can be tracked and group imbalances corrected by manipulation of the placebo warfarin dose adjustment schedules in a prospective manner during the trial. Such a correction was made in WASID after an interim analysis of number of warfarin dose changes in each group. Unfortunately, the correction did not eliminate the imbalance. Thus we recommend frequent close scrutiny of the number of warfarin dose adjustments in future trials using this system.

Only three stroke prevention trials using warfarin have been blinded [5–7] and only one (SPINAF) [7] has briefly reported success of blinding. Using information regarding frequency of treatment assignment guesses from the manuscript along with the group sizes ($n = 265$ for placebo; $n = 260$ for warfarin) we were able to reconstruct the 3×2 tables of TAGs versus ATAs and derive blinding indices for patients and study coordinators (guess data for principal neurologists were not given) for comparison to WASID (Table 4).

Although the percentages of patients/proxies and study coordinators correctly guessing treatment assignment were higher in WASID than SPINAF, the WASID blinding indices lie closer to the values indicating complete randomness in guessing (0.5 for BI, 0 for κ , 0 for each newBI_j). However, when the WASID results are corrected for uncertainty in guessing, the only substantial difference is that WASID aspirin patients were more blinded than SPINAF aspirin patients, while the reverse is true for warfarin patients. These differences were not sizeable enough to conclude that one trial was successful in blinding while the other was not. Thus we conclude that both trials experienced the same success in blinding, a reassuring result.

In order to make the comparison of WASID to SPINAF, we must assume that our ‘uncertain’ is equivalent to DK in SPINAF. This suggests that elicitation of TAG is important. Unfortunately we only have results from these two trials out of the class of blinded secondary stroke prevention trials using warfarin; hence we do not have sufficient data to draw on for inference. However, since TAG was elicited prior to uncertainty and DK was not a valid response to the TAG question in WASID, we can ascertain values of κ (95% CI) for patient/proxy (0.06 (−0.07, 0.19)), study coordinator (0.16 (0.06, 0.26)), and

principal neurologist (-0.01 ($-0.10, 0.08$)) when responses were uncertain. Thus we are assured that the certainty expressed by these respondents is true, since the values of κ expressed are consistent with random guessing.

The timing of TAG elicitation is also a controversial issue [18–24]. Some argue that TAG elicitation is only valid before onset of treatment, theorizing that efficacy corresponds with correct guessing [20,21]. In the same vein, Sackett [22,23] relates an anecdote in which opposite blinding occurred when lack of efficacy corresponded to incorrect guessing, thus purporting that assessment of TAG at trial end is really an assessment of correctness of hunches about efficacy. On a different tack, Senn [24] argues that unblinding should happen only through efficacy. Admittedly, disentangling hunches about efficacy from blinding may be nearly impossible [19]. Nevertheless, if a trial is to be fully blinded, and knowledge of the success of blinding is desirable, inquiry about TAG only makes sense when the study is nearly over. To do otherwise leaves doubt as to if subjects or caregivers have become unblinded later in the study, after TAG inquiry. Thus an inference about blinding success based on guesses elicited early in the study may be biased in that they do not reflect the entire subject's experience in the study. Under equipoise, hunches should not favor one treatment. However, there may be physiologic effects of treatment other than efficacy that may influence guess. In WASID, these effects were number of warfarin dose changes, occurrence of hemorrhage, and occurrence of GI upsets. Elicitation of TAG early in the trial would not allow for development of such effects. We propose that the relationship between TAG and such factors in each treatment group be evaluated as part of the assessment of blinding success.

Little formal research has been done on the assessment of blinding success from the statistical perspective. James *et al.* [16] proposed a formal estimator (BI) in which the DK responses are incorporated as well as the formula for its asymptotic variance. BI uses weights of 0 for correct responses, 1 for DK responses, and values between 0 and 1 for incorrect responses, arguing that the truly blinded subjects are those who respond DK. Close inspection of the formula for BI reveals that it is not a function of the weights chosen for the DK responses.

For the proper application of BI, one must assume that the most desirable response to TAG is DK. Although our survey did not directly allow for the DK response, we could reconstruct our data to allow for 'uncertain' responses to the certainty question to equate to DK. The critical assumption of this index is the validity of the DK response. Bang *et al.* [17] demonstrated that when $P_{DK} > 0.3$, BI does a poor job at ascertaining moderate to low levels of unblinding. Moreover, since BI is a single index, it combines different reasons of unblinding which may differ by treatment arms and which may cancel each other out. The newBI_{*i*} indices allow for assessment of unblinding in each arm, and thus allow for different reasons of unblinding to affect each arm differentially. We note that newBI identifies a trend toward more unblinding for warfarin than for aspirin, a trend not seen with the other indices. This trend might be explained on the basis of the differences between groups in other occurrences associated with warfarin (i.e., number of warfarin dose changes, occurrence of hemorrhage). Moreover, as for BI, the critical assumption for newBI_{*i*} is the validity of the DK response.

In summary we conclude that adequate blinding was achieved in WASID for both patient and principal neurologist, although study coordinators were less than perfectly blinded. Similarly, for both patient and study coordinator, adequate blinding was achieved for the SPINAF trial. Use of placebo warfarin dosing schedules, such as those used in WASID and SPINAF may be useful for the design of future blinded trials comparing warfarin with novel antithrombotic agents (e.g., direct thrombin inhibitors) for patients with other conditions

e.g., atrial fibrillation or deep vein thrombosis. Use of this blinding system for future trials will require vigilant tracking and adaptive modifications in placebo warfarin dosing schedules as the trials progress.

References

1. Hróbjartsson A, Forfang E, Haahr M, et al. Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. *International Journal of Epidemiology*. 2007; 36:654–63. [PubMed: 17440024]
2. Fergusson D, Glass K, Waring D, Shapiro S. Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. *BMJ*. 2004; 328:1–5. [PubMed: 14703521]
3. Altman D, Schulz F, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*. 2001; 134:663–94. [PubMed: 11304107]
4. The EPIC Investigators. Use of monoclonal antibody directed against the platelet glycoprotein IIb/IIIa receptor in high risk coronary angioplasty. *N Engl J Med*. 1994; 330:956–61. [PubMed: 8121459]
5. Stroke Prevention in Atrial Fibrillation Investigators. Adjusted-dose warfarin versus low-intensity, fixed-dose warfarin plus aspirin for high-risk patients with atrial fibrillation: Stroke Prevention in Atrial Fibrillation III randomised clinical trial. *Lancet*. 1996; 348:633–38. [PubMed: 8782752]
6. Stroke Prevention in reversible Ischemia trial (SPIRIT) study Group. A randomized trial of anticoagulants versus aspirin after cerebral ischemia of presumed arterial origin. *Annals of Neurology*. 1997; 42:857–65. [PubMed: 9403477]
7. Ezekowitz M, Bridgers S, James K, et al. Warfarin in the prevention of stroke associated with nonrheumatic atrial fibrillation. *N Engl J Med*. 1992; 327:1406–12. [PubMed: 1406859]
8. Connolly S, Laupacis A, Gent M, et al. Canadian atrial fibrillation anticoagulation (CAFA) study. *Journal of the American College of Cardiology*. 1991; 18:349–55. [PubMed: 1856403]
9. Mohr J, Thompson JRML, Levin B, et al. A comparison of warfarin and aspirin for the prevention of recurrent ischemic stroke. *New England Journal of Medicine*. 2001; 345:1444–51. [PubMed: 11794192]
10. Chimowitz M, Lynn M, Howlett-Smith H, et al. Comparison of warfarin and aspirin for symptomatic intracranial arterial stenosis. *New England Journal of Medicine*. 2005; 352:1305–16. [PubMed: 15800226]
11. The Warforin-Asprin symptomatic Intracranial Disease (WASID) Trial Investigators. Design, progress, and challenges of a double-blind trial of warfarin versus aspirin for symptomatic intracranial arterial stenosis. *Neuroepidemiology*. 2003; 22:106–17. [PubMed: 12656117]
12. Agresti, A. *Categorical Data Analysis*. 2. John Wiley & Sons; New York: 1990.
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33:159–74. [PubMed: 843571]
14. Cicchetti D, Feinstein A. High agreement by low kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology*. 1990; 43:551–58. [PubMed: 2189948]
15. Feinstein A, Cicchetti D. High agreement by low kappa: the problem of two paradoxes. *Journal of Clinical Epidemiology*. 1990; 43:543–9. [PubMed: 2348207]
16. James K, Bloch D, Lee K, et al. An index for assessing blindness in a multicentre clinical trial: disulfiram for alcohol cessation – a VA Cooperative study. *Stat in Med*. 1996; 15:1421–34. [PubMed: 8841652]
17. Bang H, Ni L, Davis C. Assessment of blinding in clinical trials. *Controlled Clinical Trials*. 2004; 25:143–56. [PubMed: 15020033]
18. Fergusson D, Glass K, Waring D, Shapiro S. Letter: Turning a blind eye. *Brit Med J*. 2004; 328:1136.
19. Bang H, Ni L, Davis C. Response to Henneickevon-Zepelina and Harri Hemila's comment. *Contemporary Clinical Trials*. 2005; 26:514–5.

20. Hemila H. Assessment of blinding may be inappropriate after the trial. *Contemporary Clinical Trials*. 2005; 26:512–4. [PubMed: 15951244]
21. Henneicke-von Zepelin H-H. Letter to the editor. *Contemporary Clinical Trials*. 2005; 26:512. [PubMed: 15936250]
22. Sackett D. Letter: Turning a blind eye; Why we don't test for blindness at the end of our trials. *Brit Med J*. 2004; 328:1136. [PubMed: 15130997]
23. Sackett D. Commentary: Measuring the success of blinding in RCTs: don't, must, can't or needn't? *International Journal of Epidemiology*. 2007; 36:664–5. [PubMed: 17675306]
24. Senn S. Letter: Turning a blind eye; Authors have a blinkered view of blinding. *Brit Med J*. 2004; 328:1135–6. [PubMed: 15130996]

Table 1

Theoretical responses for measuring relationship between guesses and actual treatment

<u>Before accounting for uncertainty</u>	<u>Actual treatment</u>		
Guessed treatment	Aspirin	Warfarin	Total
Aspirin	N_{11}	N_{12}	$N_{1\cdot}$
Warfarin	N_{21}	N_{22}	$N_{2\cdot}$
Total	$N_{\cdot 1}$	$N_{\cdot 2}$	N
<u>After accounting for uncertainty</u>	<u>Actual treatment</u>		
Guessed treatment	Aspirin	Warfarin	Total
Aspirin	n_{11}	n_{12}	$n_{1\cdot}$
Warfarin	n_{21}	n_{22}	$n_{2\cdot}$
Don't know	n_{01}	n_{02}	$n_{0\cdot}$
Total	$n_{\cdot 1}$	$n_{\cdot 2}$	N

Table 2

Guess versus assignment*

Treatment assignment	Patient/proxy		Study coordinator		Principal neurologist	
	Aspirin	Warfarin	Aspirin	Warfarin	Aspirin	Warfarin
Guess						
Aspirin	109	63	202	116	190	161
Warfarin	117	156	69	161	81	114
Missing	54	70	9	12	9	14
κ_{ALL} (95% CI ^{**})	0.19 (0.11, 0.28)		0.33 (0.25, 0.4)		0.12 (0.04, 0.2)	
After treating ‘uncertain’ as ‘Don’t Know’						
Guess						
Aspirin	41	14	42	23	32	18
Warfarin	55	99	16	97	10	52
Don’t Know	130 (58%)	106 (48%)	213 (79%)	157 (57%)	229 (85%)	205 (75%)
κ_{CERT} (95% CI)	0.31 (0.19, 0.43)		0.52 (0.38, 0.65)		0.49 (0.32, 0.65)	
BI (95% CI)	0.69 (0.65, 0.73)		0.75 (0.72, 0.79)		0.85 (0.82, 0.88)	
newBI _j (95% CI)						
Aspirin	−0.06 (−0.15, 0.02)		0.10 (0.04, 0.15)		0.08 (0.04, 0.13)	
Warfarin	0.39 (0.31, 0.47)		0.27 (0.20, 0.34)		0.12 (0.07, 0.18)	

* indices are calculated on the basis of complete cases only;

** denotes 95% confidence interval.

Table 3
Association between patient characteristics and treatment assignment guess stratified by treatment group

Assignment	Aspirin		Warfarin		Total	
	Aspirin	Warfarin	Aspirin	Warfarin	Aspirin	Warfarin
Guess						
For patient/proxy guess						
Patient characteristic	<i>n</i> = 109	<i>n</i> = 117	<i>n</i> = 63	<i>n</i> = 156	<i>n</i> = 172	<i>n</i> = 273
# warfarin dose changes (mean $\pm$ se) ^a	4.78 $\pm$ 0.3	5.2 $\pm$ 0.3	13.2 $\pm$ 1.4	15.2 $\pm$ 1	7.8 $\pm$ 0.6	10.9 $\pm$ 0.7^a
On therapy: Ever hemorrhage? (%) ^b	4	1	5	10	4	6
On therapy						
Major GI problem? (%)	0	2	0	0	0	1
Any GI problem? (%)	12	15	10	7	11	11
For study coordinator guess						
Patient characteristic	<i>n</i> = 202	<i>n</i> = 69	<i>n</i> = 116	<i>n</i> = 161	<i>n</i> = 318	<i>n</i> = 230
# warfarin dose changes (mean $\pm$ se)	4.6 $\pm$ 0.2	5.6 $\pm$ 0.4	11.3 $\pm$ 0.9	15 $\pm$ 1.1	7 $\pm$ 0.4	12.2 $\pm$ 0.8^c
On therapy: Ever hemorrhage? ^d (%)	3	3	3	13^e	3	10
On therapy						
Major GI problem? (%)	1	1	0	1	14	1
Any GI problem? (%)	13	8	13	5^f	13	6^g
For principal neurologist guess						
Patient characteristic	<i>n</i> = 190	<i>n</i> = 81	<i>n</i> = 161	<i>n</i> = 114	<i>n</i> = 351	<i>n</i> = 195
# warfarin dose changes (mean $\pm$ se)	4.6 $\pm$ 0.2	5.5 $\pm$ 0.4^h	12.5 $\pm$ 0.8	15.3 $\pm$ 1.3	8.3 $\pm$ 0.5	11.2 $\pm$ 0.9ⁱ
On therapy: Ever hemorrhage? (%)	3	5	5	14^j	4	10^k
On therapy						
Major GI problem? (%)	1	2	1	0	1	1
Any GI problem? (%)	13	11	12	4^l	12	7^m

* denotes standard error

^a *p* = 0.0003

^b *p* = 0.05 for Breslow–Day test for homogeneity of association

^c *p* < 0.0001

\$watermark-text

\$watermark-text

\$watermark-text

d $p = 0.05$ for Breslow–Day test for homogeneity of association
e $p = 0.002$
f $p = 0.03$
g $p = 0.0007$
h $p = 0.03$
i $p = 0.002$
j $p = 0.02$
k $p = 0.002$
l $p = 0.02$
m $p = 0.004$.

Table 4

Reconstructed guess versus assignment and summary statistics for blinding indices for SPINAF [7] trial

Treatment Assignment	<u>Patient</u>		<u>Study coordinator</u>	
	Aspirin	Warfarin	Aspirin	Warfarin
Guess				
Aspirin	42	42	26	10
Warfarin	16	74	10	47
Don't Know	207	144	229	203
κ (95% CI)	0.33 (0.19, 0.46)		0.55 (0.37, 0.72)	
BI (95% CI)	0.78 (0.76, 0.80)		0.86 (0.83, 0.89)	
newBI _j (95% CI)				
Aspirin	0.10 (0.04, 0.15)		0.06 (0.02, 0.10)	
Warfarin	0.12 (0.04, 0.20)		0.14 (0.09, 0.20)	

* denotes 95% confidence interval.