ORIGINAL ARTICLE

Effect of Home Testing of International Normalized Ratio on Clinical Events

David B. Matchar, M.D., Alan Jacobson, M.D., Rowena Dolor, M.D., M.H.S., Robert Edson, M.A., Lauren Uyeda, M.A., Ciaran S. Phibbs, Ph.D., Julia E. Vertrees, Pharm.D., Mei-Chiung Shih, Ph.D., Mark Holodniy, M.D., and Philip Lavori, Ph.D., for the THINRS Executive Committee and Site Investigators*

ABSTRACT

BACKGROUND

Warfarin anticoagulation reduces thromboembolic complications in patients with atrial fibrillation or mechanical heart valves, but effective management is complex, and the international normalized ratio (INR) is often outside the target range. As compared with venous plasma testing, point-of-care INR measuring devices allow greater testing frequency and patient involvement and may improve clinical outcomes.

METHODS

We randomly assigned 2922 patients who were taking warfarin because of mechanical heart valves or atrial fibrillation and who were competent in the use of point-of-care INR devices to either weekly self-testing at home or monthly high-quality testing in a clinic. The primary end point was the time to a first major event (stroke, major bleeding episode, or death).

RESULTS

The patients were followed for 2.0 to 4.75 years, for a total of 8730 patient-years of follow-up. The time to the first primary event was not significantly longer in the self-testing group than in the clinic-testing group (hazard ratio, 0.88; 95% confidence interval, 0.75 to 1.04; P=0.14). The two groups had similar rates of clinical outcomes except that the self-testing group reported more minor bleeding episodes. Over the entire follow-up period, the self-testing group had a small but significant improvement in the percentage of time during which the INR was within the target range (absolute difference between groups, 3.8 percentage points; P<0.001). At 2 years of follow-up, the self-testing group also had a small but significant improvement in patient satisfaction with anticoagulation therapy (P=0.002) and quality of life (P<0.001).

CONCLUSIONS

As compared with monthly high-quality clinic testing, weekly self-testing did not delay the time to a first stroke, major bleeding episode, or death to the extent suggested by prior studies. These results do not support the superiority of self-testing over clinic testing in reducing the risk of stroke, major bleeding episode, and death among patients taking warfarin therapy. (Funded by the Department of Veterans Affairs Cooperative Studies Program; ClinicalTrials.gov number, NCT00032591.)

From the Health Services Research Field Program, Veterans Affairs (VA) Medical Center, and the Division of General Medicine, Department of Medicine, and Center for Clinical Health Policy Research, Duke University Medical Center — both in Durham, NC (D.B.M., R.D.); Duke–National University of Singapore Graduate Medical School, Singapore (D.B.M.); Jerry L. Pettis VA Medical Center, Research and Development Service (151) and the Department of Internal Medicine, Loma Linda University — both in Loma Linda, CA (A.J.); the Cooperative Studies Program Coordinating Center (151K) (R.E., L.U., M.-C.S., M.H.), the Health Economics Resource Center (C.S.P.), and the Medical Service (M.H.), VA Palo Alto Health Care System, Palo Alto, CA; the Department of Pediatrics and Center for Primary Care and Outcomes Research, Stanford University School of Medicine, Stanford, CA (C.S.P.); the Department of Health Research and Policy, Division of Biostatistics (M.-C.S., P.L.), and the Department of Medicine, Division of Infectious Diseases (M.H.), Stanford University, Stanford, CA; and the VA Cooperative Studies Program Clinical Research Pharmacy Coordinating Center, Albuquerque, NM (J.E.V.). Address reprint requests to Dr. Matchar at North Pavilion, 2400 Pratt St., Durham, NC 27705, or at david.matchar@duke-nus.edu.sg.

*Members of the Home International Normalized Ratio Study (THINRS) are listed in the Supplementary Appendix, available at NEJM.org.

This article (10.1056/NEJMoa1002617) was updated on January 5, 2011, at NEJM.org.

N Engl J Med 2010;363:1608-20.

Anticoagulation with warfarin, if managed well, is effective in reducing thromboembolic complications such as stroke in patients with atrial fibrillation or a mechanical heart valve.[1] Unfortunately, in clinical practice, warfarin is typically underutilized and the quality of anticoagulation management can be poor, resulting in decreased effectiveness and increased complications.[2] High-quality management, such as that provided by anticoagulation clinics, can be an effective way to improve care but may require that patients travel to a centralized location, limiting the frequency of testing and in some cases access to anticoagulation treatment.

Frequent home monitoring of the internationalized normalized ratio (INR) by means of weekly self-testing is a promising strategy for improving outcomes, as suggested by one meta-analysis[3] and individual studies. An increase in the frequency of testing allows INR values that are outside the target range to be identified and addressed more quickly. Furthermore, self-testing promotes patients' engagement in their own care. Several handheld fingerstick devices have been approved by the Food and Drug Administration for home use. Coverage under Medicare was approved in 2002 for patients with mechanical heart valves and in 2008 for patients with other indications for long-term oral anticoagulation, including atrial fibrillation.

The Home INR Study (THINRS) was initiated under the Veterans Affairs (VA) Cooperative Studies Program to address the question of whether self-testing offers an advantage over high-quality clinic testing, the currently recommended practice, in reducing the risk of a major event (stroke, major bleeding episode, or death).

METHODS

STUDY DESIGN AND PATIENTS

This prospective, randomized, nonblinded trial (described previously[4]) was designed by the THINRS planning committee. Data were collected at 28 VA medical centers with anticoagulation clinics that met guidelines for high-quality anticoagulation management, as defined in the Managing Anticoagulation Services Trial (MAST),[5] and that were treating at least 400 patients. Eligible patients had atrial fibrillation, a mechanical heart valve, or both and required long-term warfarin therapy initiated by a clinical care provider for an indeterminate period. Also, unlike other trials, the patient (with or without assistance from a caregiver) needed to be deemed competent in performing self-tests using the ProTime Microcoagulation System (International Technidyne Corporation) on the basis of training by the study staff, 2 to 4 weeks of testing, and a formal competency evaluation. Patients gave separate, written informed consent for this evaluation and for participation in the randomized clinical trial. The appropriate institutional review boards approved the research protocol. The study was conducted in accordance with the amended protocol (available with the full text of this article at NEJM.org). The results of the training, testing, and competency phase of the study are described elsewhere.[6]

Randomization was carried out after the eligible patients were trained and deemed to be competent in self-testing of INR. Random assignments were made with the use of an adaptive-allocation stratification method[7] according to the duration of anticoagulation (<3 or ≥3 months) and indication for warfarin (atrial fibrillation with or without mechanical heart valve) within each site.

INR TESTING AND FOLLOW-UP

High-quality clinic testing of the INR was defined as testing that met three criteria: a designated, trained staff responsible for patients' visits and follow-up; the use of a standard local procedure at each site for anticoagulation management; and the performance of regular INR testing about once a month.[8] For self-testing, patients used an approved device at home once a week and recorded the results by means of an interactive voice-response reporting system with Web-based local monitoring. If the patient reported a measurement outside the assigned INR range or reported having been hospitalized, the reporting system instructed the patient to contact the study staff for dosing instructions or to provide additional information. Otherwise, the reporting system recorded the result and instructed the patient to continue testing at the assigned frequency.

Follow-up study visits were scheduled for every 3 months after randomization to collect data on medical events, concomitant medications, adverse events, and changes in the target INR range in the interval since the previous visit. Patients in the self-testing group were checked quarterly for meter-use competency (see the Supplementary Appendix, available at NEJM.org) and were given additional testing supplies.

An independent data and safety monitoring

committee oversaw the trial for safety and scientific integrity. Before enrollment was started, the committee established rules for interim reviews of efficacy data after accumulation of one third and two thirds of the total expected patient-years of follow-up, with the use of the Haybittle–Peto boundary[9,10] as a guide for possible early cessation of the study for efficacy. On the recommendation of the committee, the trial continued to the end of the planned follow-up period (a minimum of 2 years).

END POINTS

The prespecified primary end point was the time to a first major event (stroke, major bleeding episode, or death). Without knowledge of the treatment assignments, the Duke Clinical Research Institute adjudicated all primary end-point events, as well as most secondary end points (including myocardial infarctions and nonstroke infarction events (e.g., deep-vein thrombosis, pulmonary embolism, and other visceral infarctions [see the protocol]) but excluding minor bleeding episodes), for which there was sufficient supporting documentation, using standardized operational definitions[4] (although different criteria were used for major bleeding episodes; see the Supplementary Appendix). Prespecified secondary end points included time within the INR target range, calculated on the basis of linear interpolation[11]; patient satisfaction with anticoagulation, measured with the Duke Anticoagulation Satisfaction Scale (DASS)[12]; and quality of life, measured with the Health Utilities Index Mark 3.[13] Data regarding utilization and costs of VA services were obtained from centralized VA files. Case-report forms captured data regarding utilization of non-VA health care services, training and monitoring costs, and patients' travel expenses.

STATISTICAL ANALYSIS

On the basis of a sample size of 3200 patients, with 1 year of enrollment and a minimum follow-up period of 2 years, we calculated that a total of 363 patients with primary events would be required to discern a 32% relative reduction in the annual rate of primary events (from 5.5% for the clinic testing group to 3.75% for the self-testing group) with 90% power. Enrollment took longer than expected, however, and it took 2.75 years to complete the randomization process for the 2922 eligible patients. The mean follow-up period was 3 years (range, 2.0 to 4.75).

All analyses were based on the intention-to-treat principle unless otherwise noted. We report continuous variables as means ±SD and categorical variables as proportions. For the time to the first primary event, we estimated the survival curves according to group with the use of the Kaplan–Meier method[14] and compared the results by means of the log-rank test. Using Cox proportional-hazards regression, we estimated the effect of the intervention, adjusting for the factors used to stratify randomization.[15,16] We also compared the time-to-event distribution separately according to primary-event type, using the log-rank test and Cox regression for the time-to-death analysis and the competing-risks analysis[17] for time to stroke and time to major bleeding episode, with death as the competing risk. Regression models were used to evaluate the effect of the intervention on primary outcomes in prespecified subgroups for the stratification factors. In addition, we tested treatment by means of subgroup interactions for homogeneity of treatment effect. For patients with atrial fibrillation, we examined subgroups based on the $CHADS_2$ score, an index of stroke risk.[18] (The $CHADS_2$ score is determined by assigning 1 point each for congestive heart failure, hypertension, an age of 75 years or older, and diabetes mellitus and 2 points for prior stroke or transient ischemic attack; the sum of the points is then used to calculate the score for a given patient, with higher scores indicating a greater risk of stroke.) P values less than or equal to 0.05 were considered to indicate statistical significance. We used SAS, version 9.1.3 (SAS Institute), and R, version 2.9.1, statistical software for the analyses.

RESULTS

CHARACTERISTICS OF THE PATIENTS

The study took place from August 2003 through May 2008. Of the 3745 patients who were screened, 3643 were trained in self-testing, and 3057 underwent competency assessment; 2922 patients (80%) were randomly assigned to either the group that performed self-testing (1465) or the group that underwent high-quality clinic testing (1457) (Fig. 1). At randomization, 237 of the 2922 patients (8%) required caregiver support to perform INR testing.

Of the 2922 study patients, 684 had a mechanical heart valve and 2236 had atrial fibrillation without such a valve. Table 1 shows the baseline characteristics of the patients; 98% were men and 92% were white. The mean (±SD) age for the study

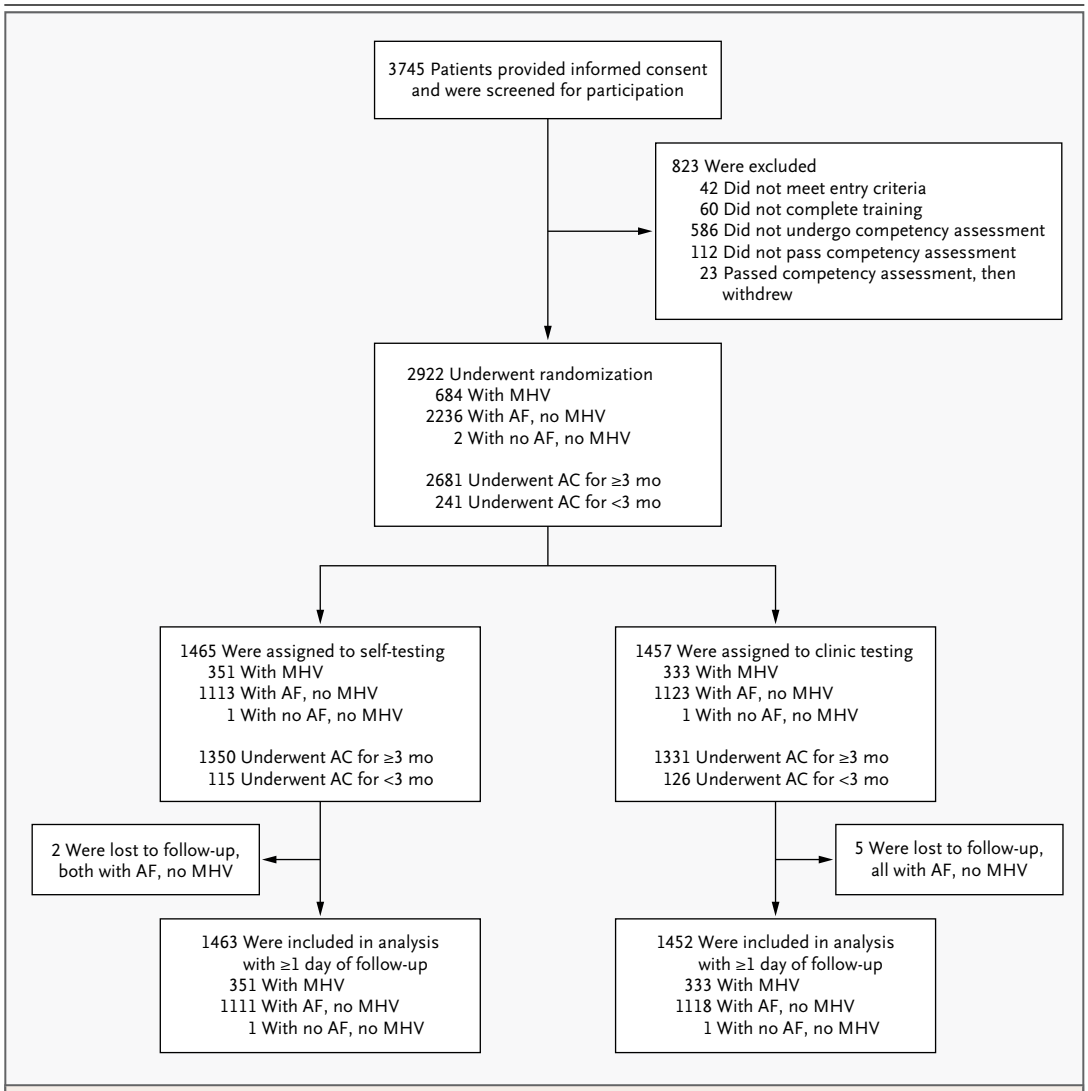


Figure 1. Enrollment and Randomization.
AC denotes anticoagulation, AF atrial fibrillation, and MHV mechanical heart valve.

population as a whole was 67.0±9.6 years, with the patients in the self-testing group being significantly younger than those assigned to the clinic-testing group (P=0.047). Patients with a mechanical heart valve were somewhat younger than those without such a valve, and among patients with mechanical heart valves, those in the self-testing group were significantly younger than those in the clinic-testing group. Atrial fibrillation was a more common indication for anticoagulation than was the presence of a mechanical heart valve (in 83% of patients vs. 23%); for patients with atrial fibrillation who did not have a mechanical heart valve, the mean $CHADS_2$ score was 1.9, after accounting for transient ischemic attacks. About 27% of patients were taking aspirin in addition to warfarin, and 8% were taking amiodarone.

PRIMARY END POINT

Figure 2 shows the Kaplan–Meier estimates for the time to the first primary event, based on 8730 patient-years of follow-up. The difference between the two study groups was not significant (P=0.14 by the log-rank test). The unadjusted hazard ratio for the primary end point in the self-testing group was 0.88 (95% confidence interval [CI], 0.75 to 1.04); after adjustment for duration of anticoagulation, presence or absence of a mechanical heart valve, and the interaction of these two factors, as well as for age, the adjusted hazard

ratio was 0.90 (95% CI, 0.76 to 1.06). Table 2 shows the number of patients with at least one primary event according to the type of first event. No significant difference was found for death (unadjusted hazard ratio, 0.91; 95% CI, 0.73 to 1.14), major bleeding episode with death as a competing risk (unadjusted hazard ratio, 0.98; 95% CI, 0.78 to 1.23), or stroke with death as a competing risk (unadjusted hazard ratio, 0.95; 95% CI, 0.58 to 1.56).

The time to the first primary event also did not differ significantly between the two study groups in any of the subgroups examined (Fig. 3), including two prespecified subgroups (indication for warfarin, mechanical heart valve vs. atrial fibrillation without such a valve; duration of antico-

Table 1. Baseline Characteristics of the Study Patients, According to Treatment Group.*

Characteristic	Self-Testing Group (N=1465)	Clinic-Testing Group (N=1457)	P Value
Male sex — no. of patients (%)	1440 (98)	1431 (98)	0.87
Age — yr			
Total study population	66.6±9.7	67.4±9.4	<0.05
Patients with MHV†	62.4±10.4	64.2±9.7	0.02
Patients with AF, without MHV†	67.9±9.1	68.3±9.1	0.30
Range	23–89	33–99	
Race or ethnic group — no. of patients (%)‡			
Hispanic	108 (7)	90 (6)	0.20
White	1347 (92)	1347 (92)	0.61
Black	94 (6)	77 (5)	0.19
Highest level of education — no. of patients (%)			0.10
Grades 1–8	57 (4)	62 (4)	
Grades 9–11	96 (7)	123 (8)	
High-school graduate	450 (31)	412 (28)	
Some college but no degree	485 (33)	447 (31)	
Undergraduate degree	261 (18)	298 (20)	
Postgraduate degree	116 (8)	115 (8)	
No. of household members	2.1±1.0	2.1±1.0	0.62
Transported to clinic (did not drive self) — no. of patients (%)	228 (16)	229 (16)	0.91
Months of anticoagulation treatment — no. of patients (%)			0.39
<3	115 (8)	126 (9)	
3–6	85 (6)	99 (7)	
>6–12	97 (7)	108 (7)	
>12	1168 (80)	1124 (77)	
Cardiac disorders — no. of patients (%)			
AF	1201 (82)	1221 (84)	0.19
AF without MHV	1113 (76)	1123 (77)	
MHV	351 (24)	333 (23)	0.48
Aortic valve	278 (19)	256 (18)	0.33
Mitral valve	91 (6)	89 (6)	0.91
Arrhythmia (other than AF)	158 (11)	160 (11)	0.87
Congestive heart failure	404 (28)	434 (30)	0.19
Angina	241 (16)	256 (18)	0.42

Table 1. (Continued.)

Characteristic	Self-Testing Group (N=1465)	Clinic-Testing Group (N=1457)	P Value
$CHADS_2$ score for patients who had AF without MHV — no./total no. (%)			0.42
0	128/1113 (12)	110/1123 (10)	
1	323/1113 (29)	328/1123 (29)	
2	327/1113 (29)	357/1123 (32)	
3	199/1113 (18)	208/1123 (19)	
4	96/1113 (9)	82/1123 (7)	
5	34/1113 (3)	36/1123 (3)	
6	6/1113 (<1)	2/1123 (<1)	
Mean $CHAD_2$ score	1.94	1.95	
Diabetes mellitus — no. of patients (%)	472 (32)	495 (34)	0.31
Hypertension — no. of patients (%)	1041 (71)	1010 (69)	0.31
Previous stroke — no. of patients (%)	136 (9)	140 (10)	0.76
Weekly warfarin dose — mg			
Mean	37.1 (16.3)	36.1 (15.9)	0.16
Median	35	35	
Range	3–135	5–112	
Antiplatelet medication — no. of patients (%)			
Aspirin	397 (27)	391 (27)	0.87
Clopidogrel	24 (2)	20 (1)	0.56
Ticlopidine	0	1 (<1)	0.50
Amiodarone	119 (8)	112 (8)	0.66

* Plus–minus values are means ±SD. AF denotes atrial fibrillation, and MHV mechanical heart valve.
† One patient in each of the randomly assigned groups had neither atrial fibrillation nor an MHV. Since randomization was stratified according to the presence or absence of an MHV, these two patients were included in the subgroup of patients who had atrial fibrillation without an MHV.
‡ The percentages total more than 100 because data on Hispanic ancestry and race were reported by the patients on separate questions; they could report being of Hispanic ancestry or more than one race. Two patients in the self-testing group and three in the clinic-testing group (<1% in each group) did not respond to the question about Hispanic ancestry.

agulation therapy, <3 months vs. ≥3 months) and two post hoc subgroups (age, <65 years vs. ≥65 years; $CHADS_2$ score, <2 vs. ≥2 among patients with atrial fibrillation but without a mechanical heart valve). With four subgroup analyses, no more than one significant interaction due to chance alone would be expected.

SECONDARY END POINTS

Table 2 summarizes the secondary end points in this study. During the period of the trial, the percentage of time during which the INR was within the therapeutic range was modestly higher in the self-testing group than in the clinic-testing group (absolute difference, 3.8 percentage points; 95% CI, 2.7 to 5.0; P<0.001). At 2 years (the minimum duration of follow-up), patient satisfaction with anticoagulation, as measured by the DASS (in which scores range from 25 to 225, with lower scores indicating better satisfaction), was greater in the self-testing group than in the clinic-testing group (difference, –2.4 points; 95% CI, –3.9 to –1.0; P=0.002), and a cumulative gain in health utilities according to the Health Utilities Index Mark 3[19] was noted in the self-testing group as compared with the clinic-testing group (difference, 0.155 points; 95% CI, 0.111 to 0.198; P<0.001). Costs were higher in the self-testing group but not significantly different from those in the clinic-testing group (difference, $1,249; 95% CI, –1,205 to 3,703; P=0.32).

The rate of loss to follow-up was the same in

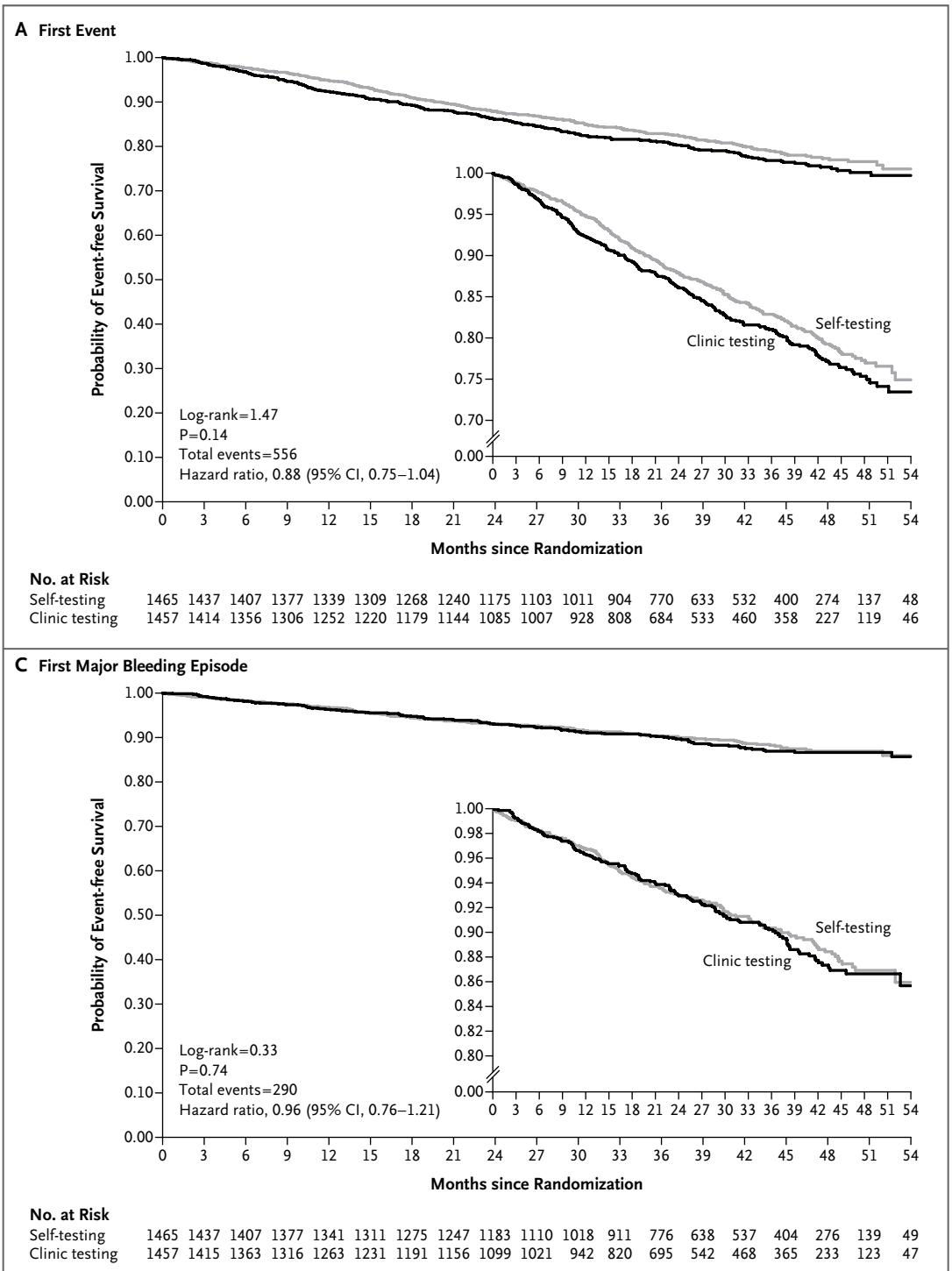
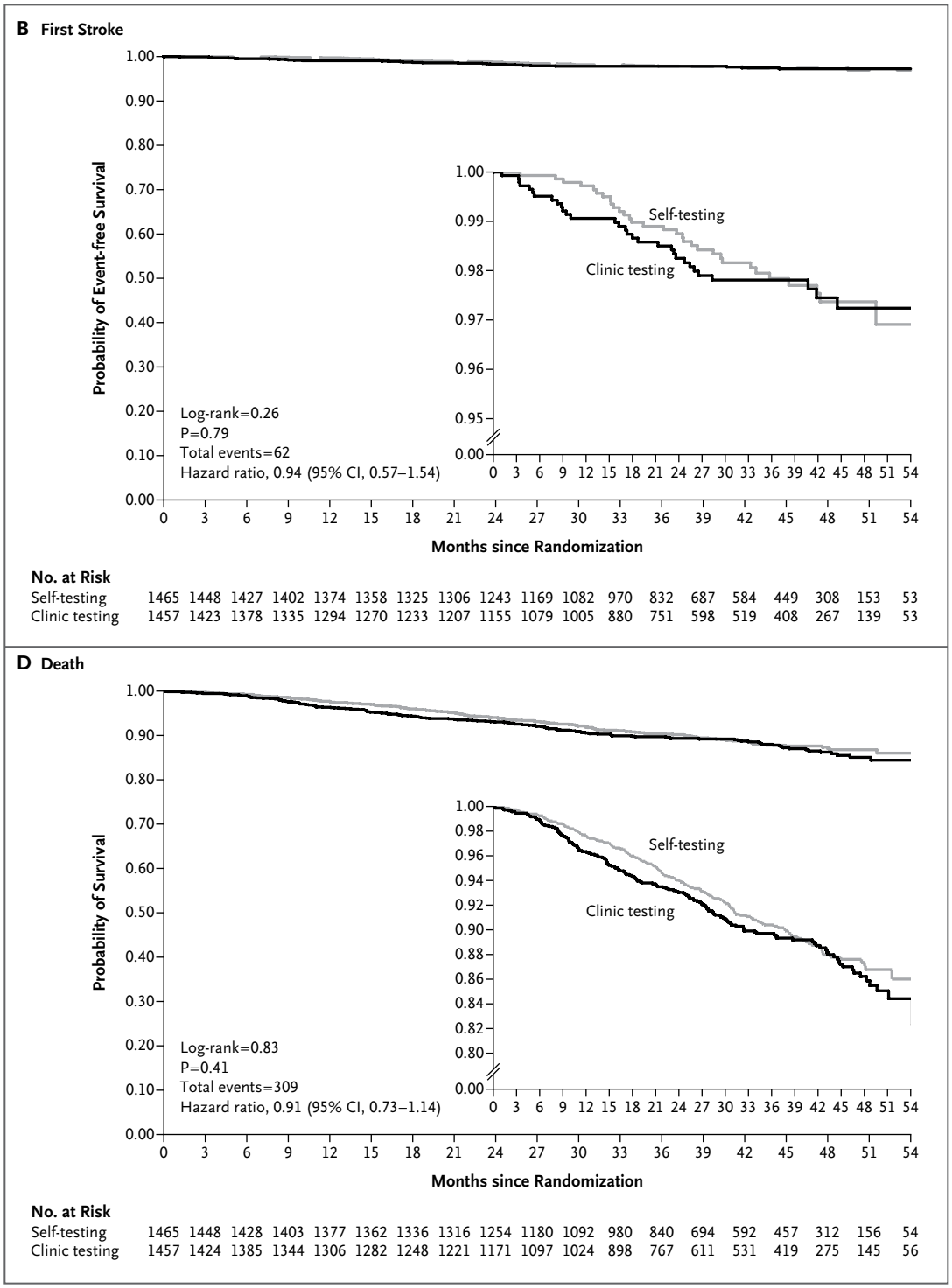


Figure 2. Kaplan–Meier Estimates of Event-free Survival According to Treatment Group.

Event-free survival rates are shown for any first event (Panel A), first stroke (Panel B), first major bleeding episode (Panel C), and death (Panel D). CI denotes confidence interval.



the two study groups (1%), as was the percentage of patients who discontinued warfarin therapy (7%). However, it was less common for patients assigned to the clinic-testing group to switch to the self-testing group (<1%) than it was for the patients assigned to the self-testing group to switch to the clinic-testing group (16% switched at some time during the study: 13% switched permanently, and 3% switched temporarily; 5% switched permanently owing to an inability to perform the INR test).

For patients in the self-testing group who were

assigned to test their INR once a week, we measured adherence to the test-frequency assignment by looking at the number of days between consecutive tests (ignoring multiple tests per day) and considered tests that took place 5 to 9 days apart to be adherent. Of a total of 164,626 tests performed by patients, about 87% were adherent; 5% were performed fewer than 5 days apart, and 8% were performed 10 or more days apart, with a mean interval between tests of 7.6±5.4 days (median, 7; interquartile range, 7 to 7). The test frequency in the clinic-testing group was decided by the care provider, but most sites tested every 4 or 6 weeks. With adherence for the clinic-testing group defined as tests performed 21 to 49 days apart, of a total of 63,673 clinic-performed tests, 52% were adherent; 44% of tests were performed fewer than 21 days apart, and about 4% were performed 50 or more days apart, with a mean interval between tests of 23.1±18.1 days (median, 22; interquartile range, 13 to 30).

ADVERSE EVENTS

Table 2 also provides tallies of patients with clinical events and the number of events by type. The total primary event rate (the number of primary events divided by the total number of patient-years of follow-up) was 9.2% (389 events in 4235 patient-years) for the clinic-testing group and 8.1% (365 in 4495) for the self-testing group. A significantly higher percentage of patients in the self-testing group reported minor bleeding episodes.

Table 2. Primary and Secondary End Points, Clinical Events, and Adherence to Test Frequency, According to Treatment Group.*

End Point	Self-Testing Group (N=1463)†	Clinic-Testing Group (N=1452)†	Unadjusted Hazard Ratio (95% CI)	P Value
Primary end point‡				
Stroke, major bleeding, or death	271 (19)	285 (20)	0.88 (0.75 to 1.04)	0.14
Stroke§	31 (2)	31 (2)	0.95 (0.58 to 1.56)	0.83
Major bleeding§	147 (10)	143 (10)	0.98 (0.78 to 1.23)	0.83
Death	152 (10)	157 (11)	0.91 (0.73 to 1.12)	0.41
Secondary end points¶			**Difference between Self-Testing Group and Clinic-Testing Group (95% CI)‖**	
Time within target therapeutic range over entire follow-up — %	66.2±14.2	62.4±17.1	3.8 (2.7 to 5.0)	<0.001
DASS score at 2 yr**	46.8±16.3	49.2±18.0	−2.4 (−3.9 to −1.0)	0.002
Cumulative gain in health utilities at 2 yr — yr††	1.204±0.619	1.049±0.575	0.155 (0.111 to 0.198)	<0.001
Health care costs at 2 yr — $	25,754±35,673	24,505±31,827	1,249 (−1,205 to 3,703)	0.32
Clinical events‡‡				
Patients with 1 or more events — no. (%)				
Nonstroke thrombotic event§§	34 (2)	41 (3)		0.40
Myocardial infarction	27 (2)	18 (1)		0.18
Minor bleeding, reported	315 (22)	254 (17)		<0.01
Primary events — no. of events (annual rate)¶¶				
Total	365 (8.1)	389 (9.2)		
Stroke‖‖	33 (0.7)	31 (0.7)		0.61
Major bleeding‖‖	180 (4.0)	199 (4.7)		0.60
Death	152 (3.4)	157 (3.7)		0.73
Secondary events — no.				
Nonstroke thrombotic event§§	38	52		0.22
Myocardial infarction	29	18		0.06
Minor bleeding episode, reported	540	401		<0.001

Table 2. (Continued.)

End Point	Self-Testing Group (N=1463)†	Clinic-Testing Group (N=1452)†	Difference between Self-Testing Group and Clinic-Testing Group (95% CI)‖	P Value
Adherence to assigned test frequency				
Weekly self-testing***				
No. of patients	1,224			
Interval since prior test — no. of tests (%)				
<5 days	7,608 (5)			
5–9 days	143,131 (87)			
10–21 days	12,318 (7)			
>21 days	1,569 (1)			
Mean interval — days	7.6±5.4			
Clinic testing every 4 or 6 wk†††				
No. of patients		1,408		
Interval since prior test — no. of tests (%)				
<21 days		28,018 (44)		
21–49 days		33,352 (52)		
50–84 days		1,847 (4)		
>84 days		456 (<1)		
Mean interval — days		23.1±18.1		

* Plus–minus values are means ±SD. CI denotes confidence interval.
† Totals exclude patients who dropped out at the time of randomization.
‡ Values are based on the first occurrence of the specified end point.
§ Death constituted a competing risk.
¶ The numbers of patients in the analysis of time within target therapeutic range over the entire follow-up period were 1462 for the self-testing group and 1408 for the clinic-testing group; at 2 years of follow-up, the numbers of patients in the analysis of the Duke Anticoagulation Satisfaction Scale (DASS) score were 1120 and 929, respectively; and the numbers of patients in the analysis of the cumulative gain in health utilities according to the Health Utilities Index (HUI) Mark 3 were 1464 and 1456, respectively.
‖ Confidence intervals were calculated on the basis of t-test comparisons.
** Scores on the DASS range from 25 to 225, with lower scores indicating higher satisfaction.
†† Scores for the HUI Mark 3 range from –0.36 to 1.00, with a negative score indicating a state worse than being dead and a score of 1.00 indicating perfect health. The range for cumulative gain in years of health utilities is the HUI Mark 3 range times the number of years after randomization when the gain is measured, so the range in this case is –0.72 to 2 yr.
‡‡ Tallies are based on adjudicated results for all end points except for minor bleeding episodes.
§§ The events included transient ischemic attack, pulmonary embolism, and deep venous thrombosis.
¶¶ The annual event rates were determined by dividing the number of primary events by the total number of patient-years of follow-up (4495 in the self-testing group and 4235 in the clinic-testing group).
‖‖ Intracranial hemorrhage and hemorrhagic stroke counted as both a major bleeding episode and a stroke, with 11 occurring in the patient self-testing group and 8 in the clinic-testing group.
*** The totals exclude patients at six sites who were assigned to perform self-testing either twice a week or once every 4 weeks for a test-frequency substudy.
††† Test frequency for the clinic-testing group was not standardized, but most sites tested every 4 or 6 weeks.

No patient reported an adverse event as a result of physical interaction with the testing device (e.g., infection, burn, or electrical shock).

DISCUSSION

THINRS was designed to test the hypothesis that weekly home INR testing would be superior to monthly clinic INR testing for improving the aggregate outcome of stroke, major bleeding, or death. However, the rates of these major events did not differ significantly between the groups; thus, other findings in this study must be interpreted in this context.

We can rule out with a high degree of confidence any important negative effect of self-testing on the primary end point — the time to a first major event. Furthermore, to a modest extent, home monitoring did improve some secondary outcomes (time in target INR range, general

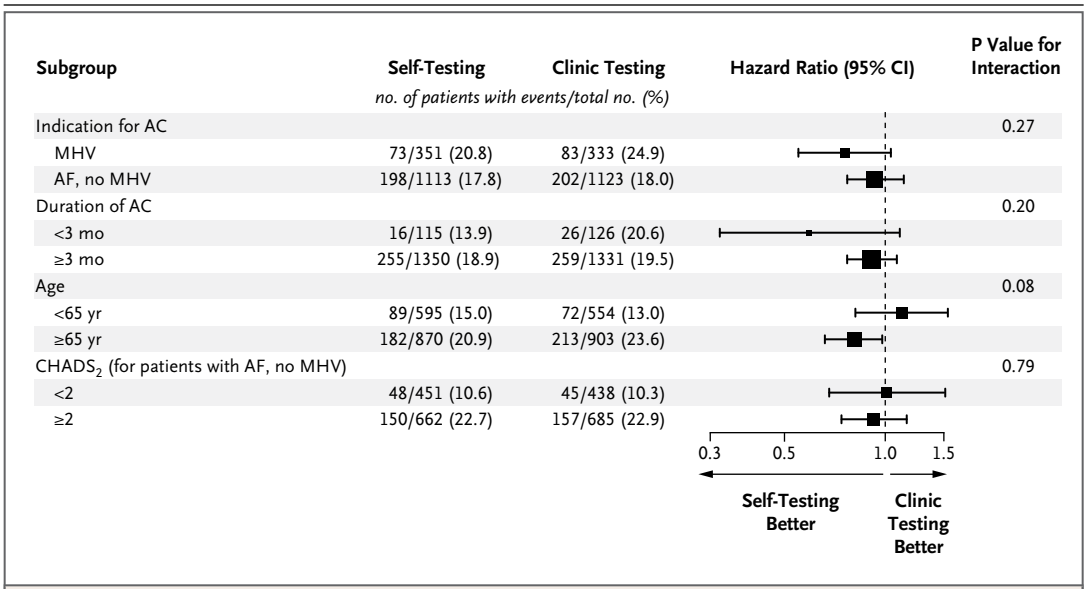


Figure 3. Hazard Ratio for the Primary End Point for Patient Self-Testing versus High-Quality Clinic Testing, According to Subgroup.

Events include death, stroke, and major bleeding episodes. The sizes of the blocks are proportional to the sizes of the subgroups. The $CHADS_2$ score is an index of stroke risk. One point each is assigned for congestive heart failure, hypertension, an age of 75 years or older, and diabetes mellitus and 2 points are assigned for prior stroke or transient ischemic attack; the sum of points is used to calculate the score for a given patient, with higher scores indicating a higher risk of stroke. AC denotes anticoagulation, AF atrial fibrillation, and MHV mechanical heart valve.

quality of life, and patient satisfaction with anticoagulation therapy, although the lack of blinding may have affected the latter two results). A high proportion (80%) of a diverse group of patients achieved competency in INR self-testing — either on their own or with the aid of a care provider.

We can also rule out a positive effect as large as that suggested by previous studies[3]; with a high degree of confidence, we can state that the hazard reduction for the primary event was no greater than 25%, as compared with the 30% figure estimated from the literature[3] on which we based our study design. The precision of our estimates of the hazard ratio (95% CI, 0.75 to 1.04) is attributable to the longer follow-up, which yielded a larger total number of events (556) than was required by the power calculation (363), despite the small shortfall in enrollment.

The patients in our study were at unusually low risk for stroke (nearly 40% of the patients with atrial fibrillation who did not have a mechanical heart valve had a $CHADS_2$ score below 2), and although the two study groups had similar but low stroke rates, we can say with a high degree of confidence that the stroke hazard ratio was above 0.58. In addition, the $CHADS_2$ score did not moderate this effect. Thus, there is no evidence that patient self-testing had an effect on the rate of stroke and there is no reason to attribute the lack of an effect to the low stroke rate.

Potential threats to external validity include unrealistically high or low quality of either treatment. In particular, it is possible that our control treatment (INR testing in the clinic), which reflects the standard of practice currently recommended,[1] exceeded the actual standard of community practice, where in some cases the INR of patients taking warfarin is in the target range less than half the time.[20] Unlike prior studies, we explicitly sought to estimate the effects of self-testing as compared with the current recommended standard of practice, and even the patients who were randomly assigned to the high-quality clinic testing had to show competency in self-testing. This requirement allowed us to assess the value of self-testing itself; moreover, randomly assigning one group to an intervention known to be substandard would, in our view, have been unethical.

The effect of self-testing may also have been less than that seen in more positive studies if the

intervention had been of lower intensity or quality, or if the devices had been less than optimal. In THINRS, all procedures were consistent with standards established by manufacturers and recommended by experts,[21] including formal patient training and evaluation, weekly testing confirmed through an interactive voice-response reporting system with prompt provider contact, confirmation of patient-reported results with data downloads from the devices at selected sites, and quarterly assessment of continued competency. Our efforts at quality control for self-testing were rewarded by good indicators, including a time in the target INR range that exceeded 66% overall and performance of more than 87% of all tests within 2 days of the goal, so we do not think failure to provide excellent self-testing was a source of bias. In addition, we obtained the devices under a competitive contract based on defined performance specifications.

Lack of a more substantial effect was not due to differences in rates of loss to follow-up or rates of discontinuation of warfarin, since these rates were similar in the two study groups. However, a higher percentage of patients in the self-testing group switched to the clinic-testing group than did the reverse; this is probably because the study did not provide INR meters to the patients in the clinic-testing group, so they had to incur the cost and inconvenience of obtaining their own devices and supplies. Even if more events were to occur after a switch from self-testing to clinic testing, such a switch would be expected to occur in clinical practice, and since our analysis was based on the intention-to-treat principle, events that occurred after a patient switched from one group to the other were included in the analysis as having occurred in the group assigned at randomization.

Although weekly self-testing did not reduce primary events to the extent suggested by the results of earlier studies, our findings may be useful in considering whether to initiate anticoagulation for a specific patient. First, a large proportion of the patients we approached to consider enrolling in THINRS demonstrated competency with both the device and the interactive voice-response reporting system, either by themselves or with help from a caregiver. Second, there was no evidence that self-testing was harmful, and patient satisfaction was modestly better with this intervention. It is noteworthy that the interactive voice-response reporting system provided a safety net to prevent events resulting from delays in follow-up, which is likely to be important in minimizing operational failures of anticoagulation monitoring (e.g., delayed reporting of INR results, lack of notification about overdue tests). Third, the time in the therapeutic INR range was moderately better for the patients in the self-testing group than for those in the clinic-testing group. Although this result did not translate into substantial reductions in event rates, it may offer some assurance to clinicians who are concerned that with self-testing the INR might not be properly monitored. As for the higher percentage of patients with reported minor bleeding episodes in the self-testing group, this may be due to the more frequent contact with the THINRS study staff in this group as compared with the clinic-testing group.

The results of THINRS do not establish the superiority of self-testing over high-quality clinic testing in preventing major clinical outcomes but do provide evidence of modest improvements in time within the therapeutic INR range, patient satisfaction with anticoagulation therapy, and quality of life. In light of the poor record of usual care and the value of anticoagulation in preventing major events, we recommend that self-testing be considered for patients whose access to high-quality anticoagulation care is limited by disability, geographic distance, or other factors, if the alternative would be to withhold a highly effective treatment.

Supported by the Department of Veterans Affairs Cooperative Studies Program.

Dr. Matchar reports receiving consulting fees from Boehringer Ingelheim and receiving honoraria from Roche. Dr. Jacobson reports being an advisory board member for the Anticoagulation Forum and the Loma Linda Veterans Association for Research, receiving consulting fees from Biosite, Boehringer Ingelheim, Daiichi Sankyo, Farallon Medical, HemoSense, Inverness Medical, Pfizer Medical, Quality Assured Services, Roche Diagnostics, Sanofi-Aventis, and Tapestry Medical, receiving grants or grants pending from Biosite, Boehringer Ingelheim, Farallon Medical, HemoSense, Inverness Medical, and Sanofi-Aventis, receiving honoraria from GlaxoSmithKline and Tapestry Medical, and receiving fees for the development of educational presentations and for serving on speakers' bureaus from Boehringer Ingelheim and GlaxoSmithKline, and receiving travel support from Biosite, Boehringer Ingelheim, Daiichi Sankyo, GlaxoSmithKline, and Sanofi-Aventis. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

REFERENCES

1. Ansell J, Hirsh J, Hylek E, Jacobson A, Crowther M, Palareti G. Pharmacology and management of the vitamin K antagonists: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines (8th edition). Chest 2008;133:Suppl:160S-198S.
2. Matchar DB, Samsa GP, Cohen SJ, Oddone EZ, Jurgelski AE. Improving the quality of anticoagulation of patients with atrial fibrillation in managed care organizations: results of the Managing Anticoagulation Services Trial. Am J Med 2002;113:42-51.
3. Heneghan C, Alonso-Coello P, Garcia-Alamino J, Perera R, Meats E, Glasziou P. Self-monitoring of oral anticoagulation: a systematic review and meta-analysis. Lancet 2006;367:404-11.
4. Matchar DB, Jacobson AK, Edson RG, et al. The impact of patient self-testing of prothrombin time for managing anticoagulation: rationale and design of VA Cooperative Study #481 — the Home INR Study (THINRS). J Thromb Thrombolysis 2005;19:163-72.
5. Matchar DB, Samsa GP, Cohen SJ, Oddone EZ. Community impact of anticoagulation services: rationale and design of the Managing Anticoagulation Services Trial (MAST). J Thromb Thrombolysis 2000;9:Suppl 1:S7-S11.
6. Dolor RJ, Ruybalid RL, Uyeda L, et al. An evaluation of patient self-testing competency of prothrombin time for managing anticoagulation: pre-randomization results of VA Cooperative Study #481 — the Home INR Study (THINRS). J Thromb Thrombolysis 2010 July 14 (Epub ahead of print).
7. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 1975;31:103-15.
8. Schulman S. Care of patients receiving long-term anticoagulant therapy. N Engl J Med 2003;349:675-83.
9. Peto R, Pike MC, Armitage P, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. Br J Cancer 1976;34:585-612.
10. Haybittle JL. Repeated assessment of results in clinical trials of cancer treatment. Br J Radiol 1971;44:793-7.
11. Rosendaal FR, Cannegieter SC, van der Meer F, Briët E. A method to determine the optimal intensity of oral anticoagulant therapy. Thromb Haemost 1993;69:236-9.
12. Samsa G, Matchar DB, Dolor RJ, et al. A new instrument for measuring anticoagulation-related quality of life: development and preliminary validation. Health Qual Life Outcomes 2004;2:22.
13. Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. Health Qual Life Outcomes 2003;1:54.
14. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457-81.
15. Cox DR. Regression models and lifetables. J R Stat Soc B 1972;34:187-202.
16. Harrell FE Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer-Verlag, 2001.
17. Fine J, Gray R. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999;94:496-509.
18. Gage BF, Waterman AD, Shannon W, Boechler M, Rich MW, Radford MJ. Validation of clinical classification schemes for predicting stroke: results from the National Registry of Atrial Fibrillation. JAMA 2001;285:2864-70.
19. Fairclough DL. Design and analysis of quality of life studies in clinical trials. Boca Raton, FL: Chapman & Hall/CRC, 2002.
20. Ansell J, Caro JJ, Salas M, et al. Quality of clinical documentation and anticoagulation control in patients with chronic nonvalvular atrial fibrillation in routine medical care. Am J Med Qual 2007;22:327-33.
21. Ansell J, Jacobson A, Levy J, Völler H, Hasenkam J. Guidelines for implementation of patient self-testing and patient self-management of oral anticoagulation: international consensus guidelines prepared by International Self-Monitoring Association for Oral Anticoagulation. Int J Cardiol 2005;99:37-45.

CLINICAL TRIAL REGISTRATION

The *Journal* requires investigators to register their clinical trials in a public trials registry. The members of the International Committee of Medical Journal Editors (ICMJE) will consider most clinical trials for publication only if they have been registered (see N Engl J Med 2004;351:1250-1). Current information on requirements and appropriate registries is available at www.icmje.org/faq.pdf.