A Computerized Intervention to Improve Timing of Outpatient Follow-up:

A Multicenter Randomized Trial in Patients Treated with Warfarin

*STEPHAN D. FIHN, MD, MPH, MARY B. McDONELL, MS, DOMOKOS VERMES, PhD, JORJA G. HENIKOFF, MS, DONALD C. MARTIN, PhD, CATHERINE M. CALLAHAN, MSHyg, DANIEL L. KENT, MD, RICHARD H. WHITE, MD, for the NATIONAL CONSORTIUM OF ANTICOAGULATION CLINICS**

Objective: To evaluate a computerized scheduling model that employs nonlinear optimization to recommend optimal follow-up intervals for patients taking warfarin.

Design: Randomized trial.

Setting: 5 anticoagulation clinics.

Patients/participants: 620 patients expected to receive warfarin for ≥ 6 weeks.

Interventions: Computer-generated recommendations for scheduling the next visit were presented to or withheld from practitioners.

Measurements and main results: The main outcome measures were the follow-up interval scheduled by the provider, the interval at which the patient actually returned to clinic, and the quality of anticoagulation control (computed as the absolute difference between the measured and target prothrombin times [PTRs] or international normalized ratios [INRs]). Follow-up intervals scheduled for the patients whose practitioners received computer-generated recommendations were significantly longer than those for control patients (mean, 4.4 vs 3.5 weeks, $p < 0.001$), despite the fact that the practitioners modified the suggested return interval by > 1 week on 40% of the visits. The interval at which the intervention group actually returned to clinic was also longer (mean, 4.4 vs 4.1 weeks, $p < 0.05$), even though the control patients tended to return at longer intervals than were scheduled by their practitioners. Control of anticoagulation was nearly the same among experimental and control patients. Life-threatening complications occurred in the care of three experimental patients and one control patient, while other serious complications occurred in the care of 16 experimental patients and 17 control patients.

Conclusions: Recommendations based on nonlinear optimization prompted clinicians to schedule less frequent follow-up for patients taking warfarin, with no deterioration in anticoagulation control. This approach to scheduling can potentially reduce utilization while maintaining quality of care for patients who require long-term monitoring.

Key words: patient; warfarin; follow-up; monitoring; scheduling; computers.

J GEN INTERN MED 1994;9:131–139.

A COMMON PROBLEM facing every outpatient practitioner is how soon to schedule routine follow-up for a chronic but stable medical problem. Although many factors enter into this decision, including the patient's preferences and social situation, the patient's income and insurance coverage, and future availability of appointments, the key determinant is the nature and severity of the patient's medical problem. The conscientious provider strives to select a return interval that balances the cost and inconvenience of an office visit against the risk of allowing development of a complication that might have been averted had the patient been seen back sooner. Because this is a highly subjective judgment, return intervals that physicians recommend for similar patients vary widely.[1–3] Little research, however, has been devoted to developing methods to match follow-up intervals to clinical circumstances. We hypothesized that concepts from control theory, often employed in industrial and engineering applications, could be applied to improve decisions about follow-up for medical problems.

We chose to explore whether these methods might aid in the management of patients taking warfarin. More than 2 million patients in the United States receive warfarin.[4] Studies showing the efficacy of warfarin in preventing stroke in atrial fibrillation and death following acute myocardial infarction have led to wide use of the drug.[5–10] Patients taking warfarin must be monitored using the prothrombin time ratio (PTR) or the international normalized ratio (INR).[11] A substantial proportion of PTR/INR determinations, ranging from 20% to 50% in various clinical reports, are outside the therapeutic range.[12–16] This is of concern because a PTR/INR prolonged beyond the therapeutic range predisposes a patient to serious bleeding, while subtherapeutic values

*Members of the National Consortium of Anticoagulation Clinics participating in the study († indicates no longer involved in the study): *Seattle VA Medical Center, Coordinating Center, Seattle, Washington*: Catherine M. Callahan, MSHyg, Stephan D. Fihn, MD, MPH, Jorja G. Henikoff, MS, Daniel L. Kent, MD, Esther Kohler, CRT†, Nancy Roben, CRNP, Mary B. McDonell, MS, Donald C. Martin, PhD, and James Reiss†; *Buffalo VA Medical Center, Buffalo, New York*: John Banas, MD, Mary Pasko, PharmD, and Marc Stern, MD†; *University of California at Davis Medical Center, Sacramento, California*: Rose Hutchinson, RN, and Richard H. White, MD; *Palo Alto VA Medical Center, Palo Alto, California*: Robert W. Coleman, RPh, and Frederick Yee, RPh; *University of Virginia Medical Center, Charlottesville, Virginia*: Daniel M. Becker, MD, Pam Buncher, RNP†, Jane Spencer-Bopp, RNP†, Linda Krongaard-DeMong, RN, and Frederick Walker III, MD†; and *Worchester Polytechnic Institute, Worchester, Massachusetts*: Domokos Vermes, PhD.

Presented at the annual meeting of the Society of General Internal Medicine, Washington, DC, April 29–May 1, 1992.

Supported by the Health Services Research and Development Program of the Department of Veterans Affairs (Grants IIR 87-063 and IIR 90-036), by the Center for Outcomes Research in Elderly Persons—A VA HSR&D Field Program, and by grants from Du Pont Pharma and Boehringer-Mannheim.

Address correspondence and reprint requests to Dr. Fihn: Section of General Internal Medicine, Seattle VA Medical Center (111M), 1660 South Columbian Way, Seattle, WA 98108.

entail an increased risk of recurrent thromboembolism.[11, 17–19] The incidence of major bleeding ranges from 0.8 to 4.1 episodes per 100 patient-years and the incidence of fatal hemorrhage ranges from 0.2 to 2.3 deaths per 100 patient-years.[20, 21]

In addition to lower therapeutic ranges for warfarin, strategies to reduce complications have involved formation of special anticoagulation clinics,[22–26] programs to improve patients' compliance with medications and medical advice, and computerized pharmacokinetic models to predict proper dosages.[27–31] Other investigators have evaluated expert consultation to physicians starting patients on anticoagulation therapy[32] or the use of home monitoring devices.[33] Of these interventions, only frequent home monitoring produced improved control of anticoagulation among outpatients and none reduced complications.

How often the PTR/INR must be measured to ensure stable control is uncertain. While many authorities state that patients with stable PTR/INR values can be followed at intervals of up to eight weeks, most providers follow patients at regular but rather arbitrary intervals of two to five weeks.[18, 21, 24, 34] Several studies suggest that frequent monitoring is not always necessary. Errichetti and colleagues reported that 89% of their patients required no change in warfarin dosage on more than 50% of visits.[24] Davis and associates found the same was true for 25% of their patients.[22] In a study by White and colleagues of patients recently started on warfarin therapy, only 46% of all PTR/INR determinations were associated with a change in warfarin dosage.[33]

We have developed a dynamically controlled stochastic model based on nonlinear optimization theory.[35] This model operates to achieve an optimal compromise for an individual patient between the costs of outpatient follow-up and the expected cost of possible complications resulting from inadequate control of anticoagulation. We have incorporated this model into an anticoagulation management system programmed onto notebook computers. We report here the results of a multicenter randomized trial to determine whether the model influenced how often patients were scheduled for follow-up or how well their anticoagulation was controlled.

METHODS

Setting/Patients

Five well-established anticoagulation clinics representing a mix of practice settings, geographic locations, and patient populations participated in the trial (Table 1). Patients who were actively enrolled in a clinic or who were newly referred to the clinic were eligible for the study if the planned duration of anticoagulation was six weeks or longer. At two clinics in university medical centers, the trial was exempted from requirements for verbal or written informed consent by the local institutional review board, and all eligible patients were enrolled. At three Veterans Affairs (VA) clinics where informed consent was required by local review boards, we invited all eligible patients to participate in person or by mail.

Intervention

At each clinic we randomly assigned patients to an intervention group, in which a computer-generated recommendation for scheduling the next visit was presented to the practitioner, or to a control group, in which a recommendation was generated but was not presented to the practitioner. Because we lacked conclusive prior evidence that the intervals recommended by the model were safe, we did not mandate that the provider adhere to them. The providers were permitted to modify the scheduling recommendations for intervention group patients in accordance with their clinical judgment. They were aware of the purpose of the scheduling model but were not informed about the results of the trial until data collection was complete.

Prior to the start of the trial, we abstracted the medical records of all active, eligible patients in each clinic. Data collected included the indication(s) for anticoagulation, all PTR/INR values, comorbid conditions, notations about alcohol consumption, and all medications prescribed, including the dates they were started and stopped. We identified all hemorrhagic and thromboembolic complications and reviewed all available hospital records.

At each visit to the anticoagulation clinic during the trial, all data, including PTR/INR results, warfarin dosages, complications, and follow-up plans, were entered directly on a laptop computer. The providers were required to specify a therapeutic target PTR/INR for each patient. Usually the target was the midpoint of the formerly recommended (INR = 2.0–3.0) or high (INR = 3.0–4.5) ranges established by the American College of Chest Physicians.[36] We maintained a record of the international sensitivity index (ISI) values for the laboratories used by each of the participating clinics, even those that reported only the PTRs, permitting conversion to INRs.

Scheduling Model

The stochastic model used to compute optimal return intervals has been described in detail elsewhere but is briefly summarized here.[35] At each clinic visit, the scheduling model computed an optimal return interval based on the PTR on the day of the patient's visit, the target PTR established by the practitioner, the number of prior visits to the clinic, the variability in the patient's PTRs over time, the cost of complications as a function of the PTR, and the cost of an anticoagulation visit. Deviation of the PTR from its ideal target value is modeled as a zero mean Brownian diffusion process with a patient-

specific diffusion coefficient based on the assumption that the PTR changes continuously over time as a result of many independent, relatively small perturbations, all of comparable magnitudes.[37] These input data are used by a nonlinear optimization algorithm to select the return interval that minimizes the total expected costs of management for each individual patient.

If the current PTR is comfortably within the range in which expected complication costs are low and the patient has exhibited low PTR variability, the recommended return interval is usually prolonged. If the PTR is approaching values at which there is a heightened risk of complications and/or the patient is unstable, early follow-up is suggested. No return interval is recommended until the patient has had at least three PTR measurements spaced a week or more apart in order to generate a minimal estimate of the PTR variability. The first few intervals thereafter tend to be short as the patient's stability remains uncertain. For patients who exhibit sustained stability of their PTRs, the model recommends progressively longer return intervals up to a maximum of approximately 12 weeks. The optimization model is highly dynamic and yields a unique recommendation for each patient at every visit at which a recommendation is possible.

At the start of the trial we made several modifications to our system to make it more clinically responsive. The major enhancement was that the practitioner was permitted to "reweight" or discount a patient's past history of PTR values if it was suspected the patient might become either more unpredictable or more stable. For example, if the provider was concerned because a previously abstinent alcoholic patient resumed drinking, recent PTR values could be weighted more heavily. In an extreme case, for example, immediately following a complicated hospital admission, the practitioner could "reset" the model, causing it to ignore all prior PTR values. The model was not reweighted or reset after changes in the dose of warfarin because we assumed the practitioner made informed dosage adjustments to keep the PTR at the target value.

Outcome Measures

The main outcome measures of this trial were the follow-up interval scheduled (i.e., intended by the practitioner), the interval at which the patient actually returned to the clinic, and the quality of anticoagulation control. We considered that a practitioner had accepted a model recommendation without modification if the patient was scheduled to return within three days of the recommended date. We computed quality of anticoagulation control as the absolute deviation of each INR determination from the practitioner's stated target INR. The mean of these deviations for the duration of the trial described a patient's overall control. We used this approach rather than the more common method of ascertaining the proportion of patients in-or out-of-control based on the therapeutic range, because practitioners were required to establish a target INR rather than a range for each patient. Providers at different clinics often set narrower or wider "therapeutic" ranges for different patients with identical target INRs. Thus, the same INR value for two patients with exactly the same target INRs but different therapeutic ranges could lead to the classification of one as "in-control" and the other as "out-of-control."

TABLE 1

Characteristics of the Patients in the Intervention and Control Groups

	Intervention (n = 301)	Control (n = 319)	Total (n = 620)
Gender—% men/% women	71/29	72/28	71/29
Age—mean (years)	61	60	61
Indication (%)			
Deep vein thrombosis	46 (15)	57 (18)	103 (17)
Pulmonary embolism	27 (9)	27 (8)	54 (9)
Atrial fibrillation	52 (17)	52 (16)	104 (17)
Systemic embolism	17 (6)	18 (6)	35 (6)
Stroke/TIA*	40 (13)	24 (8)	64 (10)
Valvular disease	69 (23)	82 (26)	151 (24)
Other	50 (17)	59 (18)	109 (18)

*TIA = transient ischemic attack.

Because their infrequent occurrence limits statistical power, hemorrhagic and thromboembolic complications were not designated primary outcome events. We did, however, carefully track these complications, which we classified as minor, serious, life-threatening, or fatal. Minor complications necessitated no additional testing, referral, or outpatient visit, but were remarkable enough to report to the provider. Examples of minor bleeding included mild nosebleeds, bruising, mild hemorrhoidal bleeding, and microscopic hematuria. Serious bleeding necessitated testing or treatment. Examples of serious bleeding included overt gastrointestinal bleeding, occult gastrointestinal bleeding for which endoscopic or radiographic studies were performed, gross hematuria that prompted cystoscopy or intravenous urography or lasted more than two days, and hemoptysis. If blood was transfused, two units or less was given. We defined life-threatening bleeding as that leading to cardiopulmonary arrest, to surgical or angiographic intervention, or to irreversible sequelae such as myocardial infarction, neurologic deficit consequent to intracerebral hemorrhage, or massive hemothorax. Bleeding was also considered to be life-threatening when it led to two of the three following consequences: 1) loss of $\geq$ 3 units of blood; 2) systolic hypotension ($<$ 90 mmHg); or 3) critical anemia (hematocrit $\leq$ 20). Fatal bleeding led directly to the patient's death.

Thromboembolic complications were also classified as minor, serious, life-threatening, or fatal. Mild superficial thrombophlebitis was, for example, minor. Serious thromboembolic events included transient ischemic

attacks, suspected stroke, deep venous thrombosis, and pulmonary embolism without respiratory or hemodynamic compromise. Life-threatening events included massive pulmonary embolism, stroke with residual neurologic deficit, and systemic embolism.

All serious, life-threatening, and fatal complications were independently reviewed by a physician investigator at the local site and three investigators at the coordinating center. Using standardized criteria, we determined whether deaths were related to bleeding caused by warfarin or to a thromboembolic complication. Disagreements were reconciled by discussion.

Statistical Analysis

We analyzed all patient visits for which the scheduling model had generated a follow-up recommendation and after which a return visit had occurred. Visits at which no recommendation had been generated (most often because the patient had not made the three or more visits necessary for the model to calculate a return interval) were not analyzed. In accordance with our method of randomization and usual assumptions about independence, we treated the individual patient as the unit of analysis for all hypothesis tests. The error structure of our data prevented us from adopting a random effects model analysis of variance approach, because the number of observations varied within subjects (ranging from two to 28 with a mean of ten) and the observations for a single subject represented a time series with unequal point spacing and unknown covariance structure. There were also site effects to be considered in the analysis.

We based test statistics for within-site comparisons on the subject means for the return interval in weeks and the absolute deviation of the PTR and INR from the target level. These means were independent, but equality of the variances could not be expected because both the number and spacing of within-subject observations varied across subjects. Even though a conventional two-group t-test was likely adequate for within-site tests of treatment effects for our relatively large sample size, we performed the computationally intensive permutation t-test (proposed by Fisher and Pitman in 1938), which does not assume equal variances.[38] A sample of 5,000 random permutations was used for each test. We repeated the weighted permutation t-tests varying the weights according to the number of observations for each subject and an assumed interclass correlation. In all cases, including the ordinary t-test, these variations had no effect on the interpretation of our results. Based on these results, we elected to apply the unweighted permutation t-test to our data.

We also used a permutation pooled t-test to test the effects of the intervention across all sites. This test is nonparametric and does not require a homogeneity of variance assumption.[38, 39] The test statistic was the sum of the differences between the intervention and the control groups within each of the five sites, eliminating any fixed-site effects. The null distribution was found by randomly permuting the intervention and control groups within each site 5,000 times and computing the statistic. The p values reported are based on the number of random permutations that were as great as or greater than the observed value of the statistic. This test is a nonparametric analog of a five-site-by-two-treatment design where the summary test is a contrast that estimates the treatment effect within each site and pools these estimates across sites. Alternately, the test is similar to pooling the five within-site t-tests into a single t-test. We assessed the effects of different weighting schemes, including weights based on both the sample sizes at each site and the number of observations for each patient, and combinations thereof. Once again, the results were essentially the same for all four methods and the unweighted test results are reported.

We performed all analyses with and without the 14 eligible patients who had withdrawn from the study and consistently found nearly identical results. We performed one-sided tests of significance to compare differences in scheduled and actual return intervals and two-sided tests to compare qualities of anticoagulation and to determine differences in the durations of computer-generated scheduling intervals. A p value of 0.05 was considered significant.

RESULTS

Characteristics of Patients

There were 849 patients randomized to the study, but 19 (2.2%) later withdrew. Withdrawals occurred at four of the five clinics and ranged from 2% to 4% of the eligible patients. All withdrawals were among the patients assigned to receive computer-generated follow-up recommendations, because control patients would notice no change in their medical care. Of the 830 remaining patients, 620 made at least one visit at which a model recommendation was generated and a subsequent follow-up visit was completed. The remaining 210 patients did not have at least one scheduling recommendation made and were excluded from all analyses. Of the 19 refusals, only 14 had made enough visits to permit a scheduling recommendation. Forty-three percent of the excluded patients were assigned to the intervention group and 57% to the control group. Far and away the most common reason for not having at least one scheduling recommendation was that a patient had an insufficient number of PTR/INR determinations to obtain a stable estimate of variability. Most of these patients were located at three sites that started a large number of new patients on warfarin therapy while the trial was in progress. Of the 210 patients excluded from the analysis, 80% enrolled six months or more after the study was started and 50% entered during the last four months.

No computerized scheduling recommendation was generated for 78% of these excluded patients; and 25% made no return visit to the clinic within the study period. This contrasts with the patients included in the analysis, 73% of whom entered the study during the first six months, while another 12% enrolled within the next three months. The patients excluded from the analysis were similar to those included in the analysis except they were less likely be receiving anticoagulation for valvular heart disease (10% vs 25%).

Seventy-one percent of the patients in the analysis were men; the average age was 61 years (Table 1). The most common indications for anticoagulation were pulmonary embolism or deep venous thrombosis (26% of the patients), valvular heart disease (24%), and atrial fibrillation (17%). The study participants had an average of 3.6 chronic medical conditions in addition to the indication(s) for anticoagulation. Most common among these were a current or prior history of malignancy (24% of patients), chronic obstructive pulmonary disease (16%), and coronary artery disease (12%).

The mean duration of warfarin therapy for patients who were taking warfarin at the start of the trial was 34 months (range one to 369 months). The mean duration of follow-up for all patients was eight months. The total durations of follow-up were similar for the intervention and control groups (208 vs 203 patient-years). Total follow-up contributed by participating clinics varied from 61 to 98 patient-years.

Effects of the Scheduling Program on Return Interval

Recommendations by the scheduling model prompted practitioners to plan significantly longer intervals between visits for the intervention group (4.4 weeks) compared with the control group (3.5 weeks, $p < 0.001$, Table 2). Of 2,472 recommendations made by the computerized model for patients in the intervention group, 60% were accepted by practitioners without modification. This is shown in Figure 1 by the large number of points lying on the diagonals of identity in the graphs of recommended versus scheduled return intervals for the intervention group. Moreover, the correlations between recommended and scheduled follow-up intervals are consistently higher among the intervention group.

The practitioners scheduled a follow-up visit that was a week or more different from the recommendation at 40% of the visits, usually reducing it (Table 3). The frequency of these modifications varied greatly among clinics, ranging from 14% to 60%. The practitioners seemed to view the scheduling recommendation as an upper boundary and infrequently exceeded it, as shown in Figure 1 by the relatively few points located above the diagonals of identity for the intervention group. The main reasons for modifying a computer-generated rec-

TABLE 2

Effects of Optimized Scheduling Intervention on Return Interval and Control of Prothrombin Time Ratio (PTR)

	Clinic 1		Clinic 2		Clinic 3		Clinic 4		Clinic 5		Total	
	Optimized Computer Follow-up (n = 38)	Control (n = 30)	Optimized Computer Follow-up (n = 75)	Control (n = 98)	Optimized Computer Follow-up (n = 75)	Control (n = 72)	Optimized Computer Follow-up (n = 27)	Control (n = 37)	Optimized Computer Follow-up (n = 86)	Control (n = 82)	Optimized Computer Follow-up (n = 301)	Control (n = 319)
Return interval (weeks, mean ± SD)												
Recommended	6.2	5.8	4.7	4.5	6.2	5.9	5.8	5.2	5.2	5.1	5.5 ± 2.1	5.2 ± 2.2
Scheduled	5.6	4.7*	4.3	3.2†	5.2	4.2†	3.8	3.1†	3.6	3.1‡	4.4 ± 1.8	3.5 ± 1.4†
Actual	5.6	4.9‡	4.4	4.3	5.0	4.6‡	3.5	3.3	3.5	3.5	4.4 ± 1.8	4.1 ± 1.8‡
Mean deviation of PTR and INR§ from target value (|target−actual|)												
ISI value¶	2.22		2.04		2.43		2.10		2.04		—	
PTR	0.19	0.19	0.22	0.22	0.19	0.17	0.16	0.19	0.15	0.16	0.19 ± 0.16	0.18 ± 0.09
INR	0.76	0.73	0.76	0.75	0.96	0.72	0.55	0.67	0.48	0.49	0.71 ± 1.21	0.66 ± 0.40

*$p \leq 0.01$.
†$p \leq 0.001$.
‡$p \leq 0.05$.
§INR = international normalized ratio.
¶Most recent international sensitivity index (ISI) value listed. The ISI values changed during the study but the INRs are based on correct ISI values.

TABLE 3

Frequency of and Reasons for Modification by Practitioners of Scheduling Recommendations for Patients in the Intervention Group

	Clinic 1	Clinic 2	Clinic 3	Clinic 4	Clinic 5	Total
Number of visits with recommendation	339	468	527	398	740	2,472
Number of modifications (%)						
Total	151 (45)	67 (14)	183 (35)	238 (60)	353 (48)	992 (40)
Longer than recommended	53 (35)	11 (16)	19 (10)	8 (3)	8 (2)	99 (10)
Shorter than recommended	98 (65)	56 (84)	164 (90)	230 (97)	345 (98)	893 (90)
Mean length of modification (weeks)						
Longer than recommended	2.5	2.2	2.1	1.5	1.3	2.2
Shorter than recommended	3.3	3.6	3.6	3.7	3.3	3.5
Reason for modification (%)						
Scheduling convenience	25 (17)	19 (28)	68 (37)	5 (2)	14 (4)	131 (13)
Interval not acceptable	123 (81)	46 (69)	106 (58)	207 (87)	325 (92)	807 (81)
Other	3 (2)	2 (3)	9 (5)	26 (11)	14 (4)	54 (5)

ommendation were concern over the length of the interval (81%) and change to a more convenient time (13%). The computer-generated recommendations for follow-up that were presented to the practitioners when seeing the patients in the intervention group averaged 5.5 weeks (range 3 days to 10.5 weeks), and the recommendations for the control subjects that were computed but not presented averaged 5.2 weeks (range 2 days to 10 weeks, Table 2).

The interval at which patients *actually* returned to the clinic was often different from that intended by the practitioner, particularly among the control patients, who frequently returned later than planned. Nonetheless, the actual intervals at which the patients made follow-up visits were still significantly longer for the intervention group than for the control group (4.4 vs 4.1 weeks, Table 2).

Effects of the Scheduling Program on Anticoagulation Control and Complications

Values for control of the PTR/INR (absolute difference between actual PTR/INR and target PTR/INR) were not significantly different between the intervention and control patients (Table 2, $p = 0.50$). The mean absolute deviations of the PTR and the INR from the target values were 0.19 PTR units and 0.71 INR units, respectively, for patients in the intervention group and 0.18 PTR units and 0.66 INR units for those in the control group. There was no important difference in the frequency of dosage changes between the intervention and control groups (11.2 vs 11.8 dose changes per year, respectively).

Apparent from Table 2 is that frequency of follow-up bore little relationship to the quality of anticoagulation control. For example, the patients in the intervention group at clinic 1 were seen an average of every 5.6 weeks compared with every 3.5 weeks at clinic 5, a difference of 40%—yet the two clinics were not significantly different in anticoagulation control. Control of anticoagulation was significantly worse, however, for patients whose actual return interval exceeded the model recommendation (mean absolute PTR deviation = 0.23 units) compared with those returning on or before the recommended date (mean absolute PTR deviation = 0.18, $p = 0.02$, adjusting for site).

There was no fatality. Clinically important bleeding complications occurred in the care of 13 intervention patients (11 serious and two life-threatening) and 15 control patients (14 serious and one life-threatening), a nonsignificant difference ($p = 0.74$). Nearly half of hemorrhagic complications in the care of control patients occurred while the PTR ratio exceeded 2.0 as compared with only 15% of those occurring in the intervention group. There were six thromboembolic complications (five serious, one life-threatening) in the intervention group and three (all serious) in the control group ($p = 0.28$).

The incidence rates of life-threatening and serious bleeding complications were 1.0 and 5.4 per 100 patient-years, respectively, in the intervention group and 0.5 and 6.7 events per 100 patient-years in the control group. The incidence rates of life-threatening and serious thromboembolic complications were 0.5 and 2.4 events per 100 patient-years in the intervention group and 0 and 1.4 events per 100 patient-years in the control group. After adjusting for intensity of anticoagulation, the risks of bleeding and thromboembolic complications in the intervention group were not significantly different from those in the control group (relative risk = 1.1 [95% CI = 0.5, 2.3] and 2.1 [95% CI = 0.5, 8.4], respectively). Three intervention patients and three control patients experienced a second complication during the study.

DISCUSSION

We used a sophisticated computerized model to assist practitioners in anticoagulation clinics with scheduling follow-up for patients taking warfarin. The model prompted the practitioners to extend the average in-

terval at which they scheduled the patients to return to clinic by 25%, as compared with control patients. Because the control patients typically did not return as often as scheduled, the effect of the model on the actual return interval for the study sample was less impressive but still statistically significant. The patients in the intervention group made 7% fewer visits than did the control patients without affecting the accuracy of their anticoagulation control. As an average of nearly 2,500 return intervals, this difference in visit frequencies was statistically significant, but of arguable significance when viewed from a clinical or managerial perspective. We believe that patients with atrial fibrillation, recurrent thromboembolism, or mechanical cardiac valves who require lifelong anticoagulation and those paying for this care would find a 7% reduction in monitoring costs to be consequential.

There are several important strengths of this study. The patient sample was large and diverse, reflecting a broad spectrum of age, gender, and clinical characteristics. The intervention was assigned randomly, a design feature often missing from other studies of new approaches to health care delivery. Furthermore, complete follow-up was obtained for all participants. The overall incidence of 0.73 life-threatening bleeding events per 100 years of patient follow-up is lower than the rates of 1.7 to 3.7 reported from recent studies of anticoagulation in patients with atrial fibrillation.[5–9] This fact, along with the absence of any fatal bleeding complications, attests to the high quality of care delivered in our participating anticoagulation clinics. Finally, our results consistently remained significant using a variety of analytic techniques.

The methodologic liabilities of this study also deserve comment. First, there was differential withdrawal after randomization: 14 eligible patients in the intervention group but no control patients. This represented less than 5% of the intervention group, and any effect on our results was negligible. Because the patients who withdrew from the scheduling trial were still followed clinically using our computerized database, we had access to data regarding their subsequent follow-up intervals, PTR/INR values, and complications. We found no difference in our results whether or not these patients were included in the analyses. We suspect that patients who withdrew may have misunderstood the intervention or may have objected to alteration in their standing follow-up arrangements. The majority of patients did not appear to share this attitude, and most patients we approached about the study were attracted to the possibility that they would need to attend the clinic less often. Control patients, on the other hand, had no cause to withdraw because their care was not altered in any way.

A second limitation was that simply making a recommendation to a busy provider was a rather weak intervention, not only because the provider could choose to ignore the recommendation, but also because the medical delivery system or the patient could act to defeat the intervention. This problem was illustrated by the lack of concordance between the providers' scheduled follow-up times and when the patients actually returned. This study also demonstrates that simply influencing the intentions and actions of providers does not necessarily improve patient outcomes.

It is also worth remarking about some of the difficulties we encountered in implementing this system in participating clinics. First, many of the practitioners initially found it difficult to accede to recommendations for follow-up intervals as long as eight to ten weeks. The practitioners were also hesitant at first to state a specific

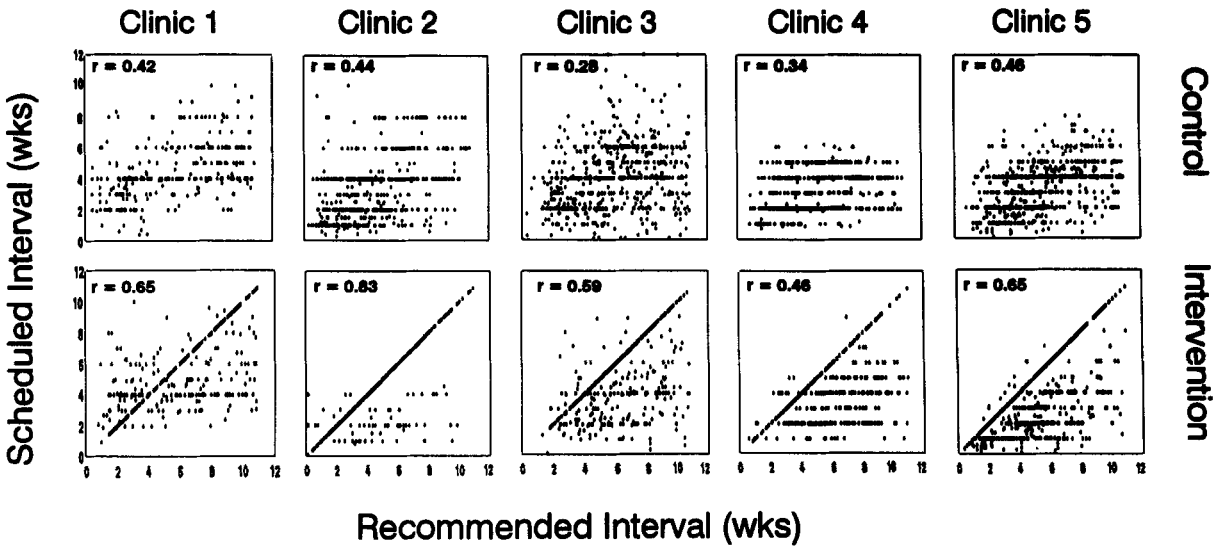

FIGURE 1. Correlation between the follow-up intervals recommended by the computerized scheduling model and the intervals scheduled by the practitioners. Recommendations were presented only to the practitioners of patients in the intervention group (bottom panels). The intervals of control patients are shown above.

target PTR/INR for each patient. As they gained confidence in the system, the practitioners became more comfortable in accepting longer recommendations, and when they did modify a recommended return interval, the reduction often was small. Second, we did not anticipate the large influx of new patients that occurred at all participating clinics during the study, approaching 100% at several sites. Because the model requires three visits before a scheduling recommendation is presented, a substantial number of visits made by newly referred patients could not be analyzed. This did not affect the overall results of the study, because when we included the 2,269 visits at which no recommendation was made, the differences in scheduled and actual return intervals were still significantly different though smaller. Moreover, the model does not recommend intervals longer than three to four weeks until data from several visits (six to 12) have provided the basis for an accurate estimate of a patient's PTR variability. The shorter return intervals recommended by the model during the first few months of therapy are consistent with the high risk of complications during this period.[21]

Our study addresses one aspect of a larger question: how frequently should stable medical outpatients receive routine follow-up? This issue has been largely ignored by clinical researchers despite important implications for cost and quality of care. Dittus and Tierney found that return intervals selected by physicians in the same institution varied up to threefold after adjusting for comorbidity and disease severity.[1] Lichtenstein et al. found that the average follow-up interval for hypertensive patients varied between practices by up to 200%.[2] A more recent study showed similar variation among physicians for a wide variety of common medical conditions.[3] These differences are similar in magnitude to the geographic variations in surgical procedures noted by several investigators.[40,41] As the primary locus of medical care shifts from the hospital to the outpatient clinic, variations in the intensity of outpatient follow-up will become an increasingly important issue. Timing of follow-up is important since expenditures for procedures and services are rarely made unless a visit occurs.

Stable patients are often routinely scheduled for ambulatory visits simply to monitor a given condition. Up to 45% of all visits to internists are for routine monitoring of chronic conditions.[4] This usually entails tracking a key clinical parameter (e.g., PTR, blood pressure, blood glucose, or serum cholesterol) to see that it remains within a desirable range. Strategies to tailor follow-up to the individual patient can potentially introduce significant efficiencies. In one example, Kent and associates applied nonlinear optimization and queuing theories to scheduling follow-up cystoscopies for recurrent bladder cancer. Compared with standard recommendations, their method shifted visits from low-risk to high-risk patients, reducing the predicted delay in detection of new tumors by 50%.[42,43]

Our work extends these earlier efforts by elaborating the theoretical basis and by developing an actual clinical application that we then tested. Ours is the first prospective study demonstrating that computerized scheduling models can potentially improve the efficiency of care while having no deleterious effect on the quality of care. This study demonstrates the feasibility of incorporating computerized support for scheduling follow-up visits and suggests that this new technology has the potential to reduce the costs and inconvenience of frequent follow-up while maintaining the quality of care. As experience with this and similar approaches accumulates, extension to other chronic conditions that involve monitoring a key clinical parameter should be pursued.

REFERENCES

1. Dittus RS, Tierney WM. Scheduling follow-up office visits: physician variability [abstract]. Clin Res. 1987;35:738A.
2. Lichtenstein MJ, Steele MA, Hoehn TP, Bulpitt CJ Coles EC. Visit frequency for essential hypertension: observed associations. J Fam Pract. 1989;28:667–72.
3. Tobacman JK, Zeitler R, Cilursu AM, Mori M. Variation in physician opinion about scheduling of return visits for common ambulatory care conditions. J Gen Intern Med. 1992;7:312–6.
4. National Center for Health Statistics, Cypress BK. Patterns of ambulatory care in internal medicine. The National Ambulatory Medical Care Survey. United States, January 1980–December 1981. Vital and Health Statistics. Series 13, No. 80. DHHS Pub. no. 84-1741. Public Health Service, Washington, DC: U.S. Government Printing Office, Sep 1984.
5. Petersen P, Boysen G, Godtfredsen J, Andersen ED, Andersen B. Placebo-controlled, randomised trial of warfarin and aspirin for prevention of thromboembolic complications in chronic atrial fibrillation: the Copenhagen AFASAK Study. Lancet. 1989;i:175–9.
6. Stroke Prevention in Atrial Fibrillation Study Group Investigators. Preliminary report of the Stroke Prevention in Atrial Fibrillation Study. N Engl J Med. 1990;322:863–8.
7. The Boston Area Anticoagulation Trial for Atrial Fibrillation Investigators. The effect of low-dose warfarin on the risk of stroke in patients with nonrheumatic atrial fibrillation. N Engl J Med. 1990;323:1505–11.
8. Connolly SJ, Laupacis A, Gent M, et al. Canadian atrial fibrillation anticoagulation study. J Am Coll Cardiol. 1991;18:349–55.
9. Ezekowitz MD, Bridgers SL, James KE, et al. Warfarin in the prevention of stroke associated with nonrheumatic atrial fibrillation. N Engl J Med. 1992;327:1406–12.
10. Smith P. The effect of warfarin on mortality and reinfarction after myocardial infarction. N Engl J Med. 1990;323:147–52.
11. Hirsh J, Dalen JE, Deykin D, Poller L. Oral anticoagulants. Mechanism of action, clinical effectiveness, and optimal therapeutic range. Chest. 1992;102(suppl):312S–326S.
12. Duxbury BM. Therapeutic control of anticoagulant treatment. BMJ. 1982;284:702–4.
13. Harries AD, Birtwell AJ, Jones DB. Anticoagulant control [letter]. Lancet. 1981;i:1320.
14. McInnes GT. Efficacy of anticoagulation in the UK [letter]. Lancet. 1981;ii:88.
15. Pegg M, Bourne J, Mackay AD, Lawton WA, Cole RB. The role of the pharmacist in the anticoagulant clinic. J R Coll Physicians Lond. 1985;19:39–44.
16. Stamp EJ, Jones SJ, Ryrie DR, Hedley AJ. Oral anticoagulants: a cost–effectiveness approach. J R Coll Physicians Lond. 1985;19:105–8.
17. Forfar JC. Prediction of hemorrhage during long-term oral coumarin anticoagulation by excessive prothrombin ratio. Am Heart J. 1982;103:445–6.
18. Landefeld CS, Flatley M, Weisberg M, Cook EF, Goldman L. Iden-

tification and preliminary validation of predictors of major bleeding in hospitalized patients starting anticoagulant therapy. Am J Med. 1987;82:703–13.
19. Landefeld CS, Rosenblatt MW, Goldman L. Bleeding in outpatients treated with warfarin: relation to the prothrombin time and important remediable lesions. Am J Med. 1989;87:153–9.
20. Levine MN, Hirsh J, Landefeld S, Raskob G. Hemorrhagic complications of anticoagulant treatment. Chest. 1992;102(suppl):352S–363S.
21. Fihn SD, McDonell M, Martin D, et al., for the Warfarin Optimized Outpatient Follow-up Study Group. Risk factors for complications of chronic anticoagulation. A multicenter study. Ann Intern Med. 1993;118:511–20.
22. Davis FB, Estruch MT, Samson-Corvera EB, Voight GC, Tobin JD. Management of anticoagulation in outpatients. Experience with an anticoagulation service in a municipal hospital setting. Arch Intern Med. 1977;137:197–202.
23. Teinders TP, Steinke WE. Pharmacist management of anticoagulant therapy in ambulant patients. Am J Hosp Pharm. 1979;36:645–58.
24. Errichetti AM, Holden A, Ansell J. Management of oral anticoagulant therapy. Experience with an anticoagulation clinic. Arch Intern Med. 1984;144:1966–8.
25. Garabedian-Ruffalo SM, Gray DR, Sax MJ, Ruffalo RL. Retrospective evaluation of a pharmacist managed warfarin anticoagulation clinic. Am J Hosp Pharm. 1985;42:304–8.
26. Cohen IA, Hutchinson TA, Kirking DM, Shue ME. Evaluation of a pharmacist-managed anticoagulation clinic. J Clin Hosp Pharm. 1985;10:167–75.
27. Hoffer EP, Marble KD, Yurchak PM, Barnett GO. A computer-based information system for managing patients on long-term oral anticoagulants. Comput Biomed Res. 1975;8:573–9.
28. Wiegman H, Vossepoel AM. A computer program for long-term anticoagulation control. Comput Methods Programs Biomed. 1977; 7:71–84.
29. Sawyer WT, Finn A. Digital computer-assisted warfarin therapy: comparison of two models. Comput Biomed Res. 1979;12:221–31.
30. O'Leary TJ, Abbrecht OH. Predicting oral anticoagulant response using a pharmacodynamic model. Ann Biomed Eng. 1981;9:199–216.
31. Wilson R, James AH. Computer-assisted management of warfarin treatment. BMJ. 1984;289:422–4.
32. Landefeld CS, Anderson PA. Guideline-based consultation to prevent anticoagulant-related bleeding. A randomized controlled trial in a teaching hospital. Ann Intern Med. 1992;116:829–37.
33. White RH, McCurdy SA, von Marensdorff H, Woodruff DE Jr, Leftgoff L. Home prothrombin time monitoring after the initiation of warfarin therapy. A randomized, prospective study. Ann Intern Med. 1989;111:730–7.
34. Gottlieb LK, Salem-Schatz S. Anticoagulation in atrial fibrillation: does efficacy in clinical trials translate into effectiveness in practice? [abstract]. Clin Res. 1992;40:582A.
35. Kent DL, Vermes D, McDonell M, Henikoff J, Fihn SD, and the Warfarin Optimal Outpatient Follow-up Study Group. A model for planning optimal follow-up for outpatients on warfarin anticoagulation. Med Decis Making. 1992;12:132–41.
36. Hirsh J, Poller L, Deykin D, Levine M, Dalen J. Optimal therapeutic range for oral anticoagulants. Second ACCP-NHLBI Conference on Antithrombotic Therapy. Chest. 1989;95(suppl):5S–11S.
37. Karatzas I, Shreve LE. Brownian Motion and Stochastic Calculus. New York: Springer-Verlag, 1988.
38. Edgington ES. Randomization Tests, 2nd Ed. New York: Marcel Dekker, 1987.
39. Noreen EW. Computer Methods of Hypothesis Testing: An Introduction. New York: John Wiley, 1989.
40. Wennberg J, Gittlesohn A. Variations in medical care among small areas. Sci Am. 1982;246:120–35.
41. Chasin MR, Kosecoff J, Park RE, et al. Does inappropriate use explain geographic variations in the use of health care services? A study of three procedures. JAMA. 1987;258:2533–7.
42. Kent DL, Schachter RD, Sox HC, et al. Efficient scheduling of cystoscopies in monitoring for recurrent bladder cancer. Med Decis Making. 1989;9:26–37.
43. Kent DL, Nease RA, Sox HC, Shortliffe LD, Schachter RD, and the Bladder Cancer Follow-up Group. Evaluation of nonlinear optimization for scheduling of follow-up cystoscopy to detect recurrent bladder cancer. Med Decis Making. 1991;11:240–8.

REFLECTIONS

The whole imposing edifice of modern medicine is like the celebrated tower of Pisa—slightly off balance.—CHARLES, PRINCE OF WALES (1948–), eldest son of Elizabeth II.