Comparison of Three Methods to Assess Therapeutic Quality Control of Treatment with Vitamin K Antagonists

Barbara A. Hutten1, Martin H. Prins1, W. Ken Redekop1, Jan G.P. Tijssen1, Siem H. Heisterkamp1, Harry R. Büller2

From the 1Department of Clinical Epidemiology and Biostatistics, 2Department of Vascular Medicine, Academic Medical Center, University of Amsterdam, The Netherlands

Summary

During treatment with vitamin K antagonists, International Normalized Ratios (INR) are determined periodically to maintain a therapeutic level of anticoagulation. We evaluated two existing methods for therapeutic quality control (linear interpolation and equidivision), with regard to their validity and reproducibility. In addition, we proposed and evaluated a (hybrid) method that takes into account potential effects of dosage modifications when INRs are far out of the target range. Validity was assessed by deleting intermediary INR results and estimating this INR based on the two surrounding INRs with each of the three methods. The estimated INRs were then compared with the observed INR.

Reproducibility of time spent in an INR range was evaluated for each of the three methods by deleting at random increasing proportions of INRs and comparing these estimates with the situation without deletions. We found that estimates of time spent in INR categories obtained with equidivision were most reproducible, but least valid. The hybrid method showed slightly higher validity and reproducibility in comparison with linear interpolation. Since these differences were small, linear interpolation is preferable to the hybrid method, since the calculations involved are easier.

Introduction

The clinical effectiveness of vitamin K antagonists has been firmly established in the primary and secondary prevention of venous thromboembolism and in the prevention of systemic embolism in patients with atrial fibrillation and prosthetic heart valves (1). The risk of thromboembolic, as well as bleeding complications, is dependent on the achieved intensity of oral anticoagulation. Under-anticoagulation gives inadequate protection against thromboembolic events, whereas over-anticoagulation increases the bleeding risk (2, 3).

Even when the target intensity of treatment with vitamin K antagonists has been defined, repeated measurements of the International Normalized Ratios (INR) are needed because this target level is often not maintained due to factors such as individual variation in pharmacodynamic response, influence of diet, other medication and variation in compliance. For this purpose INRs are determined periodically, with intervals varying from days to weeks.

Patients and Methods

Patients

Patients who took part in two previous trials, which have been described in detail elsewhere, comprised the study population (9, 10). Both were open, multicenter, randomized, clinical trials, which evaluated the efficacy and safety of two treatment strategies, i.e. intravenous unfractionated heparin versus subcutaneous low-molecular-weight heparin, in the initial treatment of patients with symptomatic venous thromboembolism. In both studies, patients were randomly assigned to receive study drug for at least 5 days. Treatment with vitamin K antagonists was started either directly or the next day and continued for three months with a target INR range between 2.0 and 3.0. An INR measurement was obtained every two or three weeks or more often, if indicated. For the present study, we used the data from the period between cessation of initial heparin treatment and the last INR measurement before discontinuation.

Description of Methods to Estimate Time Spent in Different INR Ranges

To assess therapeutic quality control, several methods are used, such as the number of INR measurements within the target range expressed as a percentage of the total number obtained, or the cross-section-of-the-files technique (4). A disadvantage of these methods is that they do not incorporate time and therefore cannot be used to calculate incidence rates of recurrences and bleeding complications at different INR levels. Time is incorporated in the method proposed by Duxbury, which assumes that the change between two consecutive INR measurements occurred halfway the interval (equidivision) (5). Also Rosendaal et al. proposed a method to estimate the time spent in INR ranges which assumes that the INR between two measurements varies linearly from the first INR, to the second INR (linear interpolation) (6). The above mentioned methods to assess therapeutic control were compared in two studies (7, 8). However, a formal method to decide on the superiority of any of these methods was not used and the conclusions were based on plausibility of the methods.

In the present study we evaluated the equidivision and linear interpolation methods with regard to their validity and reproducibility. In addition, we proposed and evaluated a new method that takes into account effects of dosage modifications in the case of INRs that are far out of the target range, in order to reflect more adequately the interaction between INR results and dosing policy.
For linear interpolation, it is assumed that the INR value between measurements varies linearly from the value of the first to the second measurement (Fig. 1b). For each moment between the measurements, the INR can be estimated.

For the hybrid method (Fig. 1c) we distinguished three scenarios (Fig. 1d).

Scenario 1: if the first INR measurement was more than 0.5 units of INR above the targeted range and the second INR was lower than the first. Scenario 2: if the first INR measurement was more than 0.5 units of INR beneath the targeted range and the second INR was higher than the first. Scenario 3: all other situations. Based on pharmacokinetics data (11) we made the following assumption for scenarios 1 and 2: the value of the second INR was attained by three days and maintained for the period thereafter. Linear interpolation was used for these first three days, hereafter time is allocated to the second INR value. In all other situations (scenario 3) linear interpolation was used.

For all methods evaluated the INR was considered not predictable for the mid-part of the interval, between two consecutive INR measurements, if it exceeded 28 days.

**Prediction of Intermediary INR Results**

To assess validity of INR interpolation for the three methods we calculated intermediary INR results and compared these calculated INR values with the observed results. In this analysis, all sets of three consecutive INR measurements (triplets) were included when the number of days between the first and the last INR of a triplet did not exceed 28 days.

For each triplet, the middle INR value was predicted with each of the three methods, based on the first and last INR. After this, the absolute difference between the predicted INR value obtained by each of the three methods and the observed INR was calculated. For each patient these absolute differences were added and subsequently divided by the number of triplets belonging to each patient. This resulted in three mean absolute differences (for each method one) per patient. The absolute differences between the observed INR value and the predicted value were compared between methods using the Student’s paired t-test.

**Reproducibility of Time Spent in INR Category**

Estimates of time spent in different INR categories involve interpolation of INR results for days when no INR results were determined. Therefore, we evaluated the stability of these estimates for each of the three interpolation methods, when different proportions of INR determinations were deleted, and compared these estimates with their own standard estimates (observed values without deletions). For comparison between the methods, eight different INR ranges were created, with steps of 0.5 units of INR. The lowest category used was ‘INR lower than or equal to 1.5’, and the highest category ‘INR greater than 4.5’.

Firstly, the time spent in different INR ranges for all patients was calculated with each of the three methods using all of the available INR measurements (standard estimate for each method). Hereafter, these calculations were repeated after randomly deleting a proportion (10, 20 or 30 percent) of the INR values. This last step was reiterated 500 times for each of the deleted proportions. Based on the 500 estimates, a mean and its non-parametric 95 percent confidence interval was calculated for each INR range.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Mean of the difference and absolute difference between the observed INR value and the predicted INR value for the three methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>no. of triplets (no. of patients)</td>
<td>mean absolute difference</td>
</tr>
<tr>
<td>All cases</td>
<td></td>
</tr>
<tr>
<td>equidivision</td>
<td>12131 (1305)</td>
</tr>
<tr>
<td>linear interpolation</td>
<td>12131 (1305)</td>
</tr>
<tr>
<td>hybrid</td>
<td>12131 (1305)</td>
</tr>
<tr>
<td>For cases where treatment policy can be expected</td>
<td></td>
</tr>
<tr>
<td>equidivision</td>
<td>1450 (725)</td>
</tr>
<tr>
<td>linear interpolation</td>
<td>1450 (725)</td>
</tr>
<tr>
<td>hybrid</td>
<td>1450 (725)</td>
</tr>
<tr>
<td>For cases where no treatment policy can be expected</td>
<td></td>
</tr>
<tr>
<td>equidivision</td>
<td>10681 (1298)</td>
</tr>
<tr>
<td>linear interpolation</td>
<td>10681 (1298)</td>
</tr>
<tr>
<td>hybrid</td>
<td>10681 (1298)</td>
</tr>
</tbody>
</table>

1 A triplet consisted of three consecutive INR measurements, the number of days between the first and the last INR does not exceed 28 days
2 see figure 1d, scenario 1 and 2
3 see figure 1d, scenario 3

For personal or educational use only. No other uses without permission. All rights reserved.
dence interval (defined by the 2.5th and 97.5th percentiles) was calculated for each of the three methods and deleted proportions. The difference between the mean and the standard estimate was calculated and it was assessed whether the 95% confidence interval of the estimated mean included its standard estimate.

**Results**

**Prediction of Intermediary INR Results**

A total of 14,708 INR values in the 1421 patients was available for analysis. Of these, 12,131 INR values (in 1305 patients) were surrounded by a previous as well as a subsequent INR determination (triplets).

The mean of the absolute differences between the observed INR value and predicted INR value for each of the three methods are shown in Table 1. The results for linear interpolation and hybrid method were almost identical, 0.59 and 0.58 units of INR respectively. The mean absolute difference obtained by equidivision was 0.74 units of INR and this differed statistically significantly (p < 0.001) from both the means obtained by linear interpolation and the hybrid method. In 1450 (12%) of the 12,131 triplets an effect of dosing regimens could be expected (Fig. 1d, scenarios 1 and 2). The mean of the absolute differences between the observed INR value and predicted INR value, in these 1450 instances, for each of the three methods are shown in Table 1. Again, the results for linear interpolation and the hybrid method were of the same order of magnitude (0.96 and 0.84 units of INR, respectively), although the mean absolute difference was slightly lower with the hybrid method (p = 0.039). The mean deviation from the observed value for equidivision (1.50 units of INR) was more than 0.5 units of INR larger in comparison with linear interpolation and the hybrid method. This difference was statistically significant compared to both the results of linear interpolation and the hybrid method (p < 0.001).

When only triplets were considered where no treatment policy could be expected (a total of 10,681 instances), the mean differences for each of the methods were positive (equidivision, 0.157; linear interpolation and hybrid, 0.081 units of INR) which indicates overestimation of the predicted value. In contrast, for triplets where actually a treatment policy could be expected, the equidivision and linear interpolation methods gave a relatively large underestimation (~0.93 and ~0.47 units of INR, respectively) while the hybrid method gave a slight overestimation (0.028 units of INR). Overall, this resulted in an underestimation when using the equidivision and linear interpolation methods, and an overestimation using the hybrid method.

**Reproductibility of Time Spent in Different INR Categories**

A total of 1329 patients was available for analysis. The reasons for non-availability were lack of or incomplete INR results (71 patients), an unknown oral anticoagulant (20 patients) or heparin cessation date (1 patient).

For all patients the percentage of time spent in the different INR categories was calculated with the three methods. The means of these percentages for each of the three methods are shown in Fig. 2. As can be seen, the distribution of time spent in the INR categories appears similar between the methods. However, the percentage of time spent in the therapeutic range was highest (57%) when calculated with the hybrid method and lowest (49%) when calculated using equidivision. For linear interpolation this figure was 55%. On the other hand, the percentage time spent in the lowest INR category for equidivision was almost twice the percentage time calculated using the linear and hybrid method.

The effect on the estimation of time spent in each of the INR categories when deleting various proportions of INR values is shown in Fig. 3. As can been seen, equidivision is not sensitive to missing values since even after 30% of INR values were deleted, the confidence intervals of the estimates included the original value. For linear interpolation, the estimates included their original value up to deletion of 10% of the INR results. When 20% was deleted the percentage time was underestimated for the INR category 1.5 to 2.0. This became more evident if a higher proportion was deleted. The estimates obtained with
the hybrid method included their original value up to deletion of 20% of the INR results. Only if 30% of the observations was deleted was the original value not included in the 95% confidence interval for the INR category 1.5 to 2.0.

**Discussion**

The calculation of the percentage of time that patients actually spent within the targeted INR range has been suggested as the preferred method to assess the quality of oral anticoagulant control, since it takes the variation in time between INR measurements into account and allows the calculation of incidence rates at different INR levels (8). Because INR measurements are not performed daily, INRs have to be estimated for the days between the available measurements. For this purpose linear interpolation and equidivision methods are currently used (2, 3, 9, 10, 12). However, it is unknown how well the estimated INR values based on these methods represent the true INR values and how sensitive these methods are for variability in the length of period with missing INR values (i.e. the period between two available measurements). Consequently, it is uncertain how much confidence we can have in results obtained using these methods. Furthermore, both methods do not take into account the response of the physician if extremely high or low INR values are encountered. Therefore, a new method was introduced (the hybrid method), which is a refinement of linear interpolation and takes into account the possible interaction between INR results and treatment policy. This new method as well as the two existing methods were evaluated with respect to validity and sensitivity to variability of the time between INR measurements.

Our results showed that the estimates of time spent in INR categories obtained with equidivision were insensitive for variation in length of time between INR measurements, since these estimates were stable, even if 30 percent of the observations was deleted. However the calculated intermediary INR results using this method were inferior to those obtained with the other two methods.

Linear interpolation and the hybrid method performed equally well in the prediction of intermediary INR results, although the hybrid method predicted slightly better intermediary INR results, especially when treatment responses could be expected. Although the hybrid method was somewhat less sensitive to variation in length of time between INR measurements than linear interpolation the differences were not large.

It should be realized that in the first months of treatment with vitamin K antagonists, INR monitoring is likely to be relatively frequent. Since potential differences between linear interpolation and the hybrid method attenuate when INR measurements are performed over short periods, differences between these could have been larger for chronic treatment. In addition, for the hybrid method we assumed that the INR on day 3 and further would correspond to the value of the second INR for all types of vitamin K antagonists. Since different types of vitamin K antagonists with corresponding pharmacokinetic and pharmacodynamic profiles were used, this could have influenced the results for this method.

For the evaluation of these methods we used a data set which was derived from and reflects daily practice. To gain a better understanding of the fluctuations of INR values in the period between two consecutive measurements (with a possible adaptation of the dosage), ideally a daily INR measurement for each patient should be available without impact on dosage. This is difficult to realize in outpatients, since a daily visit to a clinical laboratory would be required. Recently, evaluations of portable instant prothrombin time measurement devices have been reported (13). If these devices are sufficiently validated, daily measurements will potentially become feasible. This would allow further refinement of the existing methods to estimate intermediate INRs.

In summary, we evaluated two existing methods and proposed a new method to assess therapeutic quality control of treatment with vitamin K antagonists. We found that the results obtained with equidivision were most reproducible, but the least valid. Although the hybrid method showed higher validity and reproducibility in comparison with linear interpolation, the differences were not large. Therefore, linear interpolation is preferable to the hybrid method, since the calculations involved are easier.

**Acknowledgements**

The authors sincerely wish to thank all the Columbus and Tasman investigators for the opportunity to evaluate the data collected on the patients included in these randomized clinical trials.

**References**

13. Besselaar van den AM, Breddin K, Lutze G, Parker-Williams J, Taborski U, Vogel G, Tritschler W, Zerback R, Leinberger R. Multicenter evaluation of able instant prothrombin time measurement devices have been reported (13). If these devices are sufficiently validated, daily measurements will potentially become feasible. This would allow further refinement of the existing methods to estimate intermediate INRs.

In summary, we evaluated two existing methods and proposed a new method to assess therapeutic quality control of treatment with vitamin K antagonists. We found that the results obtained with equidivision were most reproducible, but the least valid. Although the hybrid method showed higher validity and reproducibility in comparison with linear interpolation, the differences were not large. Therefore, linear interpolation is preferable to the hybrid method, since the calculations involved are easier.

**Acknowledgements**

The authors sincerely wish to thank all the Columbus and Tasman investigators for the opportunity to evaluate the data collected on the patients included in these randomized clinical trials.

**References**


Received March 23, 1999 Accepted after revision June 17, 1999