

PAPERS

Can the long term outcome of individual patients with transient ischaemic attacks be predicted accurately?

Graeme J Hankey, James M Slattery, Charles P Warlow

Abstract

The prognosis of individual patients with transient ischaemic attacks (TIAs) is extremely variable; some patients are at high risk and others at low risk of a serious vascular event. Prediction equations of outcome were developed, based on eight clinical prognostic factors, from a cohort of 469 hospital-referred TIA patients ("training" data set), that enable high (and low) risk patients to be identified and for whom costly and risky treatments may (or may not) be targeted. The study aimed to determine whether these equations are externally valid and can predict outcome, with reliability and discrimination, in two independent cohorts of TIA patients ("test" data sets): 1653 TIA patients in the UK-TIA aspirin trial and 107 TIA patients in the Oxfordshire Community Stroke Project. Predicted outcomes agreed closely with the observed outcomes in the "test" data sets (reliability) for all outcome events at low five year risk (<30%) but the estimates of risk were less precise in groups predicted to have a high five year risk (>40%). The prediction equations were fairly accurate in discriminating between patients who subsequently suffered the outcome event of interest and those who survived free of the event at five years after the TIA, particularly at lower cut-off levels distinguishing high and low risk (for example, <30% vs >30% at five years). It is very difficult to achieve perfect discrimination because there is no single important prognostic factor for TIA patients that indicates whether a patient is going to suffer an event or not. These equations can be used to provide a reliable estimate of the absolute five year risk of a serious vascular event in hospital-referred TIA patients but they cannot, as yet, be used with confidence to distinguish patients at high risk from patients at low risk.

(*J Neurol Neurosurg Psychiatry* 1993;56:752-759)

As a *group*, patients with transient ischaemic attacks (TIAs) have an increased risk of stroke and other serious vascular events.^{1,2} As

individuals, however, the prognosis is extremely variable. Now that effective treatments are available for TIA patients,³⁻⁸ some of which are expensive and risky,^{7,8} it is important that patients at high and low risk can be distinguished so that effective but costly and risky treatments can be targeted to the high risk patients in addition to less expensive and safer treatments which can be offered to a wider range of patients.

We have recently developed prediction equations (that is, mathematical models) for TIA patients for three important vascular outcome events (i. *stroke*, ii. *coronary event*, and iii. the composite outcome event *stroke, myocardial infarction or vascular death*). These were based on eight simple clinical variables collected at presentation in a cohort of 469 hospital-referred patients with TIA due to presumed large vessel atherothromboembolism, small vessel lipohyalinosis or cardiogenic embolism.⁹ However, we cautioned against their widespread acceptance because they had not been validated externally in another independent cohort of TIA patients^{10,11}; only one other attempt has been made to validate a prediction equation for TIA patients but this was unsuccessful, perhaps because the original equation was based on a retrospective study of a rather small number of outcome events.¹² In this paper we examine the external validity of our predictive equations in two independent groups of TIA patients from the UK-TIA aspirin trial¹³ and the Oxfordshire Community Stroke Project.^{1,14,15}

Patients and methods

The original prediction equations for i) first-ever *stroke*, ii) *coronary event*, and iii) the composite event *stroke, myocardial infarction (MI) or vascular death*, were derived from a Cox proportional hazards multiple regression analysis of the clinical and survival data of a cohort of 469 hospital-referred TIA patients (the "training" data set).⁹ A prediction equation for *stroke, MI or death* was also developed (unpublished) to examine the effect of censorship by death before five years on the validation of the prediction equations, and to demonstrate that there were no major qualitative differences in the validity of the equations for all four outcome events that might have been caused by the limited violation of the

Department of Clinical Neurosciences, Western General Hospital, Edinburgh, UK

G J Hankey
J M Slattery
C P Warlow

Correspondence to:
Dr Hankey, Department of Neurology, Royal Perth Hospital, Wellington Street, Perth, Western Australia, Australia 6001

Received 25 June 1992
and in revised form
23 November 1992.
Accepted 27 November 1992

assumption that any patient who is censored has the same chance of suffering an outcome event as an uncensored individual with similar prognostic variables.

The *internal* validity (or predictive accuracy) of the equations was examined by applying them to the "training" data set from which they were derived. The *external* validity of the equations was examined by applying them to two independent "test" data sets: a cohort of 1653 TIA patients in the UK-TIA aspirin trial¹³ and 107 TIA patients in the Oxfordshire Community Stroke Project (OCSP).^{14,15} For the purposes of this study, 601 patients in the UK-TIA aspirin trial who presented with minor stroke and 13 patients who presented with a TIA that was probably not due to large artery atherothromboembolism, lipohyalinosis or cardiogenic embolism were excluded from the original cohort of 2435 patients to preserve diagnostic uniformity. In addition, 168 TIA patients in the original UK-TIA aspirin trial cohort and 77 TIA patients in the original OCSP cohort who had been included in the hospital-referred "training" data set were excluded to avoid any overlap between cohorts (this overlap of patients was because all three studies were in progress in Oxford at the same time). The patients in the UK-TIA aspirin trial were all followed up until October 1990.

Two aspects of validity were examined: reliability and discrimination.¹⁶ *Reliability* is a measure of how well the predictions of the equations correspond with observed outcomes. *Discrimination* is a measure of how well the equations separate individual patients with good and poor outcomes (that is, at low and high risk).¹⁶

Reliability was assessed by dividing the independent "test" sample of TIA patients into (prognostic) groups according to each patient's predicted probability of an outcome event at five years. We then compared the *predicted* mean probability of the outcome event of interest in each prognostic group with the "*observed*" proportion in each group, which was derived from the Kaplan-Meier estimate.^{16,17} (The "observed" proportions [and numbers] of patients are actuarial estimates because the Kaplan-Meier technique censors individuals from the analysis if they die of an event not being considered in that particular analysis, or are not followed up for a full five years.) The mean predicted and "observed" outcome probabilities in each prognostic group were plotted against each other; perfect reliability being represented by a diagonal line of identity. Reliability was not assessed on an individual patient basis because an individual's survival at five years can only be 0% or 100%. The reliability of the prediction equations was only assessed in the UK-TIA aspirin trial cohort because the number of outcome events in each arbitrarily defined prognostic group in the OCSP cohort was too small for meaningful analysis.

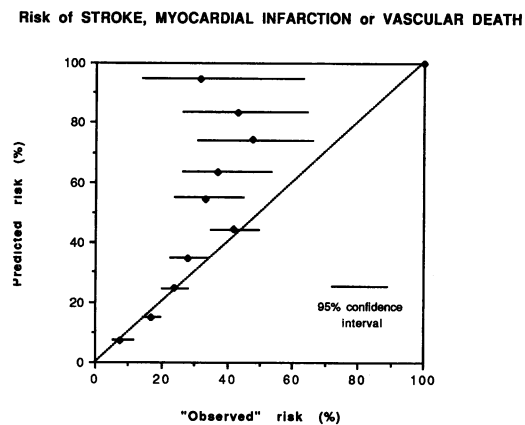
The discrimination of the prediction equations was assessed on an individual patient by patient basis. The "test" data set was divided

into two groups of patients according to a threshold probability p : any patient receiving a score of at least p was predicted to be at "high" risk of suffering the outcome event of interest at an arbitrarily appointed time (five years, in this case) whereas any patient receiving a score of less than p was classified as "low risk". For each of several arbitrarily defined "thresholds" of probability (e.g. $p < \text{or } \geq 10\%$, $< \text{or } \geq 20\%$, $< \text{or } \geq 30\%$, etc.) different groups of patients predicted to be at high and low risk were obtained. Within these groups each patient suffered the outcome event, did not suffer the outcome event but was observed for a full five years, or was censored by either the end of the study or by death from some other cause.

The "observed" number of patients in each predicted (high and low) risk group who had, and had not, suffered the outcome event of interest at five years was estimated by means of the Kaplan-Meier technique,¹⁷ taking into account those individuals who had been censored before five years.

For each threshold of probability, a two-by-two table was constructed using a) the number of patients predicted to be at low risk who were "observed" to have *not* suffered the outcome event; b) the number of patients predicted to be at low risk who were "observed" to have suffered the outcome event; c) the number of patients predicted to be high risk who were "observed" to have *not* suffered the outcome event, and d) the number of patients predicted to be high risk who were "observed" to have suffered the outcome event. The sensitivity and specificity¹¹ of the predictions were calculated; the sensitivity being the proportion of patients "observed" to have suffered the outcome event who were predicted to be at high risk of suffering the event, and the specificity being the proportion of patients "observed" to have not suffered the event who were predicted to be at low risk. An odds ratio was also calculated, describing the odds of suffering an event if predicted to be at high risk to the odds of an event if predicted to be at low risk. We then plotted the "sensitivity" (or true positive rate) against "1-specificity" (or false positive rate) of the prediction equation for each of the arbitrarily designated thresholds of probability (or cut-off points) separating high and low risk groups and thereby constructed a receiver-operator characteristic (ROC)^{18,19} graph. The overall accuracy of a prediction equation can be described by the area under the ROC curve, which ranges from 0 to 1. An area of 1.0 indicates perfect discrimination, an area of 0.5 represents no discrimination (that is, the equivalent of a random performance), and values less than 0.5 are associated with a worse-than-random performance.^{18,19} The area under the ROC curve can be interpreted as the probability that randomly chosen high and low risk patients will be ranked correctly. Another way of estimating the same quantity (probability of a correct ranking) is by calculating the non-parametric Wilcoxon or Mann-Whitney statistic (W),

Figure 1 Reliability of the prediction equation for stroke, MI or vascular death tested in the UK-TIA aspirin trial cohort of 1653 TIA patients. The dotted diagonal line represents the line of identity, indicating perfect reliability.



using standard methods.¹⁹ The result is a single numerical value which identifies the overall location of an ROC curve relative to the (non-informative) diagonal. It may be regarded as a summary measure of the discriminating power of the equation over the whole range of criteria (or cut-off points) for distinguishing high from low risk groups. The 95% confidence interval (CI) of W quantifies how variable W (or the area under the ROC curve) will be in different samples of similar size and if the lower CI is greater than 0.5, the equation's performance is significantly better than chance. The ROC curve not only illustrates overall discriminative validity but also discriminative validity over a range of different cut-off probability thresholds (or risks). It shows how severe the trade-off is between sensitivity and specificity and indicates where the best cut-off point may be.

Table 1 Baseline characteristics and treatments of the TIA patients in each of the three cohorts

| Characteristics (for definitions, see reference ²) | Hospital-Referred ("training" data set) | UK-TIA aspirin trial ("test" data set) | Community (OCSP) ("test" data set) |
|---|--|---|---------------------------------------|
| No. patients presenting with TIA (brain or eye) | 469 | 1653 | 107 |
| Presenting condition: | | | |
| TIA brain | 311 | 1278 | 95 |
| TIA eye (amaurosis fugax) | 158 | 252 | 7 |
| TIA brain and TIA eye | | 123 | 5 |
| Presumed arterial territory | | | |
| No. (%) carotid | 346 (74) | 1231 (74) | 77 (72) |
| No. (%) carotid and vertebrobasilar | 36 (8) | 90 (5) | 7 (7) |
| No. (%) vertebrobasilar or uncertain | 87 (19) | 332 (20) | 23 (21) |
| Mean (SD) age in years | 62.1 (12) | 60.0 (9) | 72.1 (12) |
| No. (%) with age >65 years | 218 (46) | 494 (30) | 78 (73) |
| No. (%) of men | 317 (68) | 1118 (68) | 49 (46) |
| No. (%) of current smokers | 221 (47) | 882 (53) | 27 (25) |
| No. (%) with hypertension* | 198 (42) | 598 (36) | 66 (62) |
| No. (%) with ischaemic heart disease* | 99 (21) | 341 (21) | 28 (26) |
| No. (%) with atrial fibrillation | 18 (4) | 30 (2) | 23 (21) |
| No. (%) with ECG evidence of LV strain or LVH | 46 (10) | 107 (6) | 21 (20) |
| No. (%) with peripheral vascular disease* | 79 (17) | 195 (12) | 11 (10) |
| No. (%) with residual neurological signs | 26 (6) | 103 (6) | 13 (12) |
| No. (%) with diabetes* | 25 (5) | 60 (4) | 12 (11) |
| No. (%) with hypercholesterolaemia | 201 (42) | 564 (34) | 43 (40) |
| No. (%) with haematocrit > 0.50 | 46 (10) | 53 (3) | 10 (9) |
| No. (%) with presumed infarction on CT scan/no. of scans | 52/302 (17) | 107/718 (15) | 24/73 (33) |
| No. (%) having cerebral angiography | 229 (49) | 441 (27) | 13 (12) |
| Treatment before/at presentation | | | |
| Antihypertensive | 144 (31) | 417 (25) | — |
| Antiplatelet | 8 (2) | 277 (17) | 3 (3) |
| Anticoagulant | 2 (0) | 10 (<1) | 4 (4) |
| Treatment during follow-up | | | |
| Antihypertensive | 179 (38) | — | — |
| Antiplatelet | 239** (51) | 1187 (72) | 52 (49) |
| Anticoagulant | 11 (2) | 102 (6) | 11 (10) |
| Carotid endarterectomy | 55** (12) | 78 (5) | 2 (2) |

*Slightly different definitions of disease/risk factor between the OCSP and hospital studies; that is, same definitions for hospital series and UK-TIA aspirin trial

**an additional 54 (12%) patients were taking placebo in the UK-TIA aspirin trial

***an additional 27 (6%) patients were randomised to "no surgery" in the European Carotid Surgery Trial

Table 2 Reliability of the prediction equation for stroke, myocardial infarction or vascular death tested in the UK-TIA aspirin trial cohort of 1653 TIA patients

| PREDICTED risk of STROKE, MI or VASCULAR DEATH at five years (%) | | | "OBSERVED" risk of STROKE, MI or VASCULAR DEATH at five years (%) | | | |
|--|-----------|----------------------------|---|--|-----------------------------|--|
| Predicted risk at 5 years (%) | (mean, %) | Number of patients at risk | Number of patients predicted to have an event | Observed risk at 5 years Kaplan-Meier estimate (%) | 95% confidence interval (%) | No. of patients "observed" to have had an event (Kaplan-Meier) |
| 90-100 | 94.7 | 13 | 12 | 31.6 | 13 to 64 | 4 |
| 80-90 | 83.4 | 24 | 20 | 43.0 | 26 to 65 | 10 |
| 70-80 | 74.2 | 30 | 22 | 47.6 | 31 to 68 | 14 |
| 60-70 | 63.7 | 42 | 27 | 36.8 | 24 to 54 | 16 |
| 50-60 | 54.6 | 73 | 40 | 33.0 | 23 to 45 | 24 |
| 40-50 | 44.4 | 117 | 52 | 41.6 | 33 to 51 | 49 |
| 30-40 | 34.5 | 220 | 76 | 27.9 | 22 to 34 | 61 |
| 20-30 | 24.7 | 389 | 96 | 23.7 | 20 to 28 | 92 |
| 10-20 | 14.8 | 509 | 75 | 16.9 | 14 to 20 | 86 |
| 0-10 | 7.3 | 236 | 17 | 7.3 | 5 to 12 | 17 |
| Entire cohort | 26.5 | 1653 | 437 | 22.5 | 21 to 25 | 373 |

STROKE, MYOCARDIAL INFARCTION or VASCULAR DEATH

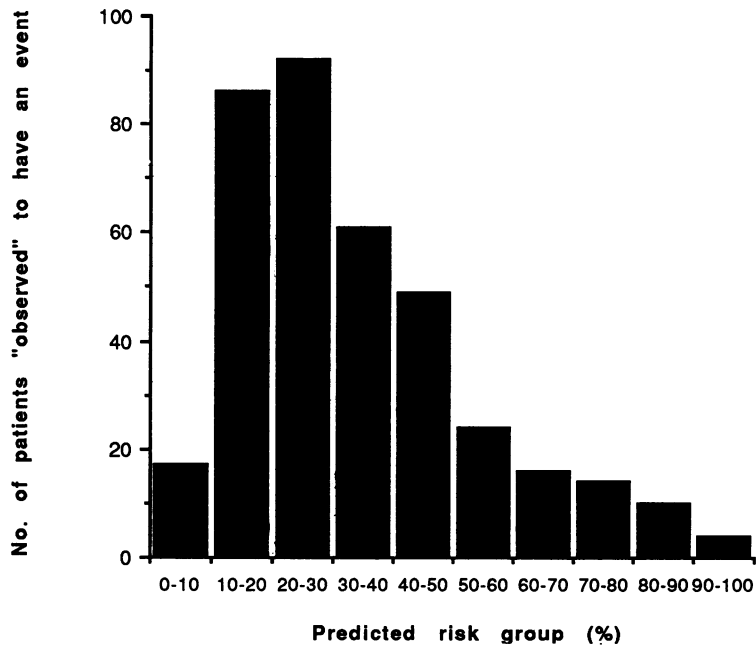


Figure 2 Histogram showing the number of patients in each predicted risk group who were "observed" to suffer a stroke, MI or vascular death during the first five years after TIA in the UK-TIA aspirin trial cohort of 1653 patients.

External validation of the prediction equation for
STROKE, MYOCARDIAL INFARCTION or VASCULAR DEATH
in the UK-TIA aspirin trial TIA patients (n=1653)

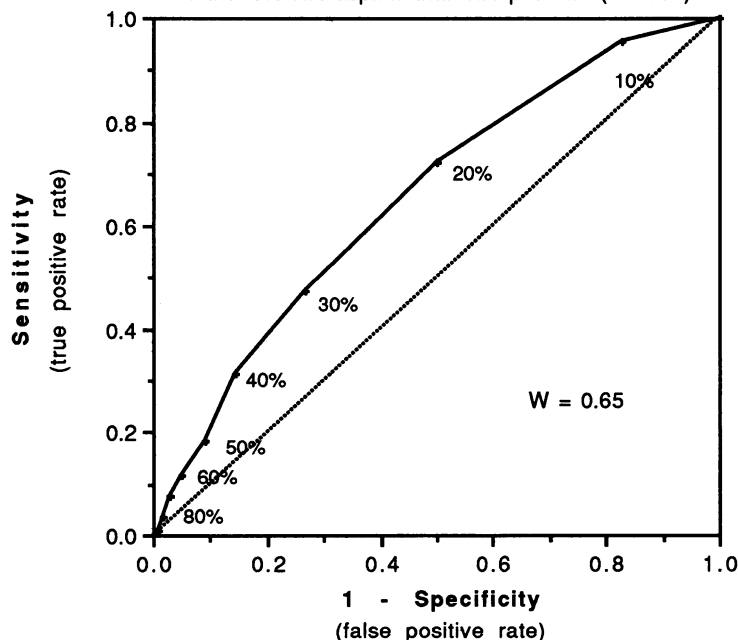


Figure 3 Receiver operator characteristic (ROC) curve illustrating the external discrimination of the prediction equation for stroke, MI or vascular death, tested in the UK-TIA aspirin trial cohort of 1653 TIA patients. The percentages at different points along the ROC curve correspond to different cut-off points, or threshold probabilities of suffering a stroke, MI or vascular death at five years, at which high and low risk are distinguished. The dotted diagonal line running from lower left to upper right describes an equation that contributes no discriminative information (area under the curve = 0.5). Prediction equations that discriminate well have ROC curves which approach the upper left hand corner of the graph.

Results

Base-line characteristics (table 1)

In comparison with the hospital-referred series, the patients in the UK-TIA aspirin trial were younger and had a higher prevalence of cigarette smokers and a lower prevalence of amaurosis fugax (AFx), hypertension, atrial fibrillation (AF), left ventricular hypertrophy (LVH), peripheral vascular disease, hypercholesterolaemia, and elevated haematocrit. Antiplatelet and anticoagulant therapy were more commonly prescribed in the UK-TIA aspirin trial whereas carotid angiography and endarterectomy were more frequently performed in the hospital-referred series. The patients in the OCSF were older than the hospital-referred series and had a greater prevalence of women, hypertension, AF, LVH, residual neurological signs, diabetes mellitus and cranial CT scan evidence of a low density area(s) consistent with infarction, and a lower prevalence of AFx, cigarette smokers and carotid bruit. Again, carotid angiography and endarterectomy was performed more frequently in the hospital-referred series while a greater proportion of OCSF patients were treated with anticoagulant therapy.

Validation of the prediction equations (models)

In the interests of keeping the size of this paper to a minimum, we will present only the results of the validation studies for *stroke, MI or vascular death*. The results are available for the other outcome events and the results are similar.

a) Reliability

Table 2 and fig 1 illustrate the reliability of the prediction equation for *stroke, MI or vascular death* as tested in the UK-TIA aspirin trial cohort. Fig 1 shows that the predicted and "observed" risks closely follow the diagonal line representing equal predicted and "observed" risks. The estimates of risk were less precise, with wider 95% confidence intervals, in groups predicted to have a high risk (that is, greater than about 40% at five years) because the number of outcome events was small in these groups. Figure 2, derived from table 2, shows that most of the outcome events occurred in the TIA patients at lowest predicted risk.

b) Discrimination

In each arbitrarily defined predicted risk group for *stroke, MI or vascular death* at five years, table 3 shows: the number of UK-TIA aspirin trial patients in each predicted risk group; the "observed" probability of surviving free of this outcome event at five years; the "observed" number of survivors free of the outcome; the "observed" number of patients who did not survive free of the outcome; the sensitivity (true positive rate) of the prediction of high risk; one minus the specificity (false positive rate); and the ratio of the odds of patients suffering the outcome event if predicted to be at high risk to the odds of patients suffering the outcome event

Table 3 External discrimination of patients at high and low risk in the UK-TIA aspirin trial. Kaplan-Meier estimates of outcome at five years, by prognostic group, in the UK-TIA aspirin trial

| Risk group (predicted) | Number of patients in UK-TIA trial with this level of predicted risk (n) | STROKE, MYOCARDIAL INFARCTION or VASCULAR DEATH | | | Sensitivity | 1-Specificity | Odds Ratio* |
|------------------------|--|--|---|--|-------------|---------------|-------------|
| | | Kaplan-Meier estimate of the "observed" survival-free of the event | "Observed" number free of the event (KM estimate) | "Observed" number with the event (n-KM estimate) | | | |
| <90% | 1640 | 0.775 | 1271 | 369 | 0.01 | 0.01 | 1.6 |
| ≥90% | 13 | 0.684 | 8.89 | 4.11 | | | |
| <80% | 1616 | 0.778 | 1257.2 | 358.8 | 0.04 | 0.02 | 2.3 |
| ≥80% | 37 | 0.609 | 22.5 | 14.5 | | | |
| <70% | 1586 | 0.783 | 1241.8 | 344.2 | 0.08 | 0.03 | 2.7 |
| ≥70% | 67 | 0.574 | 38.5 | 28.5 | | | |
| <60% | 1544 | 0.787 | 1215.1 | 328.9 | 0.12 | 0.05 | 2.5 |
| ≥60% | 109 | 0.597 | 65.1 | 43.9 | | | |
| <50% | 1471 | 0.793 | 1166.5 | 304.5 | 0.18 | 0.09 | 2.3 |
| ≥50% | 182 | 0.626 | 113.9 | 68.1 | | | |
| <40% | 1354 | 0.811 | 1098.1 | 255.9 | 0.31 | 0.14 | 2.8 |
| ≥40% | 299 | 0.609 | 182.1 | 116.9 | | | |
| <30% | 1134 | 0.828 | 939 | 195 | 0.48 | 0.27 | 2.5 |
| ≥30% | 519 | 0.657 | 341 | 178 | | | |
| <20% | 745 | 0.861 | 641.4 | 103.6 | 0.72 | 0.50 | 2.6 |
| ≥20% | 908 | 0.703 | 638.3 | 269.7 | | | |
| <10% | 236 | 0.927 | 218.7 | 17.2 | 0.95 | 0.83 | 4.3 |
| ≥10% | 1417 | 0.749 | 1061.3 | 355.7 | | | |
| OVERALL | 1653 | 0.774 | 1279.4 | 373.6 | | | |

W = 0.65 (95% confidence interval 0.62 to 0.68)

KM: Kaplan-Meier

*The odds ratio in this context is the ratio of the odds of a patient suffering an event (stroke, MI or vascular death) if predicted to be at high risk of the event, to the odds of a patient suffering an event if predicted to be at low risk.

if predicted to be at low risk. For example, if we consider 90% as the cut off between high and low risk, table 3 shows that the Kaplan-Meier estimate of survival free of *stroke, MI or vascular death* at five years in the "test" data set of UK-TIA aspirin trial patients was 0.775 in the group of patients predicted to be at low risk and 0.684 in the group predicted to be at high risk. There were 1640 patients in the group predicted to be at low risk and 13 patients in the high risk group. Among the 1640 low risk patients, 1271 were "observed" to survive free of *stroke, MI or vascular death* at five years, and 369 were "observed" to have suffered a *stroke, MI or vascular death* at five years. Among the 13 high risk patients, 9 were "observed" to survive free of *stroke, MI or vascular death* at five years (actually 8.89 [13 × 0.684], based on the Kaplan-Meier estimate), and 4 were "observed" to have suffered a *stroke, MI or vascular death* at five years (actually 4.11). The sensitivity of the prediction by the equation was 0.01 (4.11/[4.11 + 369]), and one minus the specificity was 0.01 (8.89/[8.89 + 1271]). The ratio of the odds of patients suffering the outcome event if predicted to be at high risk of the outcome event (4.11/8.89 = 0.46) to the odds of patients suffering the outcome event if predicted to be at low risk (369/1271 = 0.29) was 1.6 (0.46/0.29).

Figure 3 is an ROC curve showing the discriminative external validity of the prediction equation for *stroke, MI or vascular death* at five years in the UK-TIA aspirin trial cohort. The index of discrimination, based on the W statistic for *stroke, MI or vascular death*, is tabulated at the bottom of table 3.

Comparison of internal validity and external validity

Figure 4 compares the ROC curves of the hospital referred series (internal validation) with the UK-TIA aspirin trial and OCSF

cohorts (external validation) for *stroke, MI or vascular death* at five years. In the UK-TIA aspirin trial cohort, the prediction equations performed best when the cut-off for discriminating between high and low risk was at a low risk level (less than about 30% to 40% risk at five years), corresponding to the right side of the ROC curves. At these lower cut-off levels of risk, there was little difference between the internal and external validation curves, perhaps because the number of patients and outcome events were large and the estimates therefore fairly precise. At cut-off points above about 30 to 40% (refer to fig 3 for approximate location of cut-off points along the ROC curve), the number of patients and outcome events were smaller. Consequently, the results were prone to sampling error and less precise. Similarly, the greater irregularity of estimates in the OCSF reflected the smaller number of patients and outcome events and potential for sampling error.

Discussion

Clinical prediction equations are a form of health care technology which demand the same rigorous evaluation as new diagnostic techniques and interventions. Not only must the outcome events and predictive findings be clearly defined and relevant to medical practice, but the prediction equations must be shown to be valid. Moreover, the "acid test" of a prediction equation is not the quality of prediction that it achieves in the "training" data set from which it was derived (internal validity) but its predictive accuracy (external validity) in other, independent "test" data sets.^{10 11} Of course, it is highly unlikely that a prediction equation derived from patients in the "training" data set will fit another data set quite so well.^{10 16}

In this study, the original prediction equations were validated in two independent

STROKE, MYOCARDIAL INFARCTION or VASCULAR DEATH

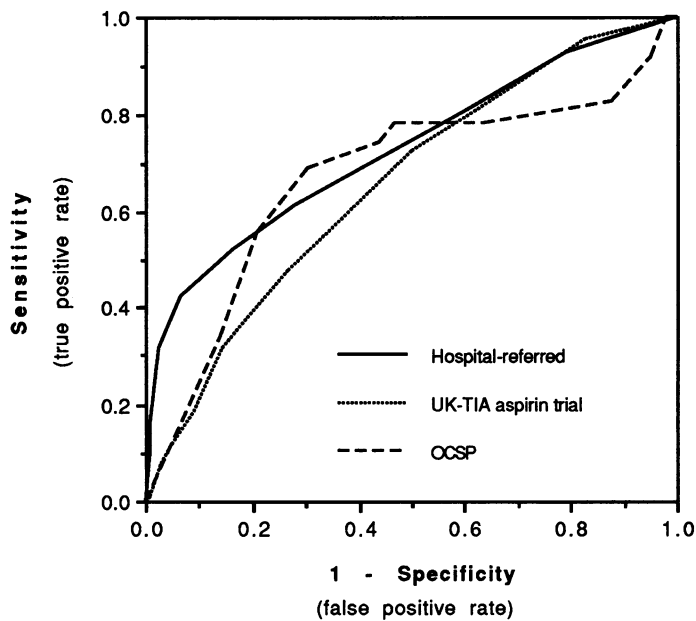


Figure 4 Receiver operator characteristic (ROC) curves comparing the discrimination of the prediction equation for stroke, MI or vascular death in the hospital-referred cohort of 469 TIA patients (internal discrimination), and the UK-TIA aspirin trial cohort of 1653 TIA patients and the Oxfordshire Community Stroke Project cohort of 107 TIA patients (external discrimination).

cohorts in terms of reliability (the capacity to accurately predict the chance of an outcome event in groups of patients with a certain predicted risk) and discrimination (the capacity to discriminate which individual patients are going to suffer, and not suffer, an outcome event).

The reliability of the equations was good for lower risk patients (0 to 30% chance of an outcome event in the first five years after a TIA) but for patients at higher risk (for example, above about 30 to 40%), each of the prediction equations tended to consistently overestimate risk. Although the estimates at high predicted risk were more imprecise, with wide 95% confidence intervals, this was not purely a chance effect due to small numbers because the deviation was consistent for all outcome events and all increments of risk. This may be explained by the different incidences of *stroke, MI or vascular death* in the "training" and "test" data sets and by the phenomenon of regression (of estimates) towards the mean. We have not attempted to calibrate the equations at high risk points because a risk of greater than one in three (33%) of a serious vascular event over the next five years is high risk by anyone's standards and well in excess of the risks of high risk treatments (such as carotid endarterectomy).

The clinical usefulness of a reliable prediction equation is that it allows the risks of the disease to be compared with the risks of a treatment. However, for clinicians confronted with a decision to make about a patient's management, the reliability (or precision) of a prediction equation is probably not as important as its ability to discriminate between indi-

vidual patients who will suffer the outcome event and those who will not. For example, knowing that a patient has a 15% risk of a stroke in the next five years (with reliability) is not as useful as knowing whether the patient is likely to suffer an event or not (discrimination). In addition, no adjustment or calibration can correct an equation if discrimination deteriorates whereas, with good discrimination, the predictor can be recalibrated to attain better reliability.¹⁶

In this study, the overall discrimination of the prediction equations was significantly better than a random chance performance (the lowest 95% confidence intervals of W did not overlap with 0.5) and there was little difference between the internal and external validation, particularly for low risk patients. However, the overall discrimination was not as high as we would have hoped; the areas under the ROC curves did not even approach one (the highest upper 95% confidence interval for W was 0.83 for *stroke, MI or death* in the OCSP). However, it must be emphasised that it is an inherently difficult task to discriminate which individual TIA patients are going to suffer, and not suffer, serious vascular events. This is because there is no powerful single clinical prognostic factor for all TIA patients.

Having examined the performance of a prediction equation over a wide range of cut-off points, or thresholds of probability, at which patients at "high risk" and "low risk" are discriminated, it is important to examine the sensitivity and specificity of the prediction equations at different "cut-off" points, and to determine where an acceptable trade-off lies between decreasing the number of false negatives (that is, improving sensitivity) and therefore increasing the number of false-positives (that is, deteriorating specificity), and vice-versa. The "best" cut-off may be defined in a number of different ways depending on the management decisions which are to be made on the basis of the patient's predicted risk. Ideally, if the costs and risks of any intervention for a patient with any given level of predicted risk were known, the best cut-off point would be the point at which the penalties of an intervention exceeds the penalties of no intervention. However, where little information is available on the medical and economic penalties of making inappropriate treatment decisions, many other, somewhat arbitrary criteria, exist for selection of cut-off points. These include a) maximising the ratio of odds of an adverse event within the group predicted to be high risk to the odds of an adverse event within the low risk group; b) simply minimising the total number of inappropriate treatment decisions (such as aggressive treatment of someone who would not have had an adverse event, or failure to treat someone who would have had an adverse event), and c) identifying the point at which the ROC curve maximally approaches the upper left hand corner of the box. The problem with these in the context of TIA patients is that the ROC curve is based on a mathe-

mathematical model that is trying to distinguish two distinct (high and low risk) groups. If the two groups were clearly distinct it would be sensible to select a point on the ROC curve that indicated maximum discrimination. However, a cohort of TIA patients is not really made up of two distinct populations but is more a population of people who have a whole range of different probabilities of stroke. Although these probabilities can be calculated with accuracy and reliability it is difficult to discriminate a fairly homogeneous cohort of high or low risk TIA patients. We have discussed the possible best cut-off points for high and low risk on the basis of data from the ROC curves and the odds ratios but acknowledge that more data need to be collected on risks and costs before better decision making processes can be applied.

As the validity of the prediction equations was best at lower cut-off points for high and low risk at five years (<30%), and as there is probably little practical importance of distinguishing a 30% risk from a 50% or 70% risk (they are all high risks), it is presently recommended that the equations be accepted as reliable up to about a 30% risk of an outcome event at five years, and that patients predicted to be at higher risk than that should be classified simply as "high risk". However, a lot more work is needed in this area. In particular, we need to seek whether there is any association between the predicted risk status of a patient and the risk of treatment and to decide over what time period the predicted risks should be examined and compared with the risks of treatment.

The prediction equations that have been adopted most widely in other disorders, such as the acute abdomen²⁰ and head injured patients,²¹ have all been validated repeatedly in a variety of test data sets. We are also examining the external validity of our prediction equations in the Dutch TIA trial²² and the European Carotid Surgery Trial⁷ cohorts of TIA patients. The next step will be to determine whether these equations, if and when widely validated, produce clinical benefits that are cost-effective when applied prospectively.¹⁰

In addition, it is relevant to consider an important finding, incidental to the aim of this study, that most of the outcome events occur in the TIA patients with lower predicted and "observed" risks. This is the same situation that is seen with other risk factors for stroke, such as blood pressure and cholesterol; most strokes occur in people with a "normal" blood pressure and cholesterol.^{23 24}

Conclusions

1) The process of risk stratification recognises that costly and risky investigations and treatments, such as carotid angiography and endarterectomy, may improve outcome in TIA patients who are at high risk of stroke, despite "best" medical therapy.

2) Costly and risky interventions will have little favourable effect however on the excel-

lent prognosis of TIA patients in the low-risk groups.

3) Our prediction equations, based on eight clinical variables that can be obtained at the bedside, are simple and reliable predictors of stroke and other serious vascular events for groups of TIA patients of similar predicted risk, and are better clinical tools than chance for stratifying individual TIA patients into categories of high and low risk of subsequent serious vascular events. The very similar predictive capacity of the equations in the "training" and "test" populations, particularly at low levels of risk (less than about 30% at five years), indicates that the equations have the potential to be generalised to different TIA populations and provide reasonably accurate prognostic stratification in those patients. However, the equations may perform less well if used in patient populations with a considerably lower prevalence of prognostic factors than the patients in our "training" and "test" cohorts.

4) The role of prediction equations in public health is to distinguish high and low risk patients so only the high risk patients are exposed to costly and risky (but effective) treatments, while the low risk patients are not exposed to these treatments that are potentially dangerous and consume medical resources unnecessarily.

5) Prediction equations should be used with caution; they should not lead to overemphasis on the need to identify and treat *only* high risk patients because most strokes, and other vascular outcome events, occur in patients who are predicted to be at lower risk (in the same way that most strokes occur in people with "normal" blood pressure and plasma cholesterol, who are considered to be at low risk). This emphasises the complementary benefits of the mass strategy of stroke prevention; simple, inexpensive and safe interventions with modest effects, such as aspirin and vascular risk factor control, need to be given to virtually all TIA patients, in the same way that simple and safe ways of lowering the mean blood pressure and serum cholesterol in the population are needed. The "mass" approach is likely to have a greater public health impact on stroke prevention than the "high risk" approach in which high technology (costly and risky) interventions are offered to the relatively few high risk patients.

Dr Hankey was supported by the Chest, Heart and Stroke Association and Mr Slattery by the Medical Research Council, UK. The data from the UK-TIA aspirin trial and the Oxfordshire Community Stroke Project were provided by the UK-TIA aspirin trial collaborators and Drs MS Dennis, JM Bamford and PAG Sandercock respectively. Dr M Correia assisted in the analysis of the UK-TIA aspirin trial prognostic factors.

- 1 Dennis M, Bamford J, Sandercock P, Warlow C. Prognosis of transient ischaemic attacks in the Oxfordshire Community Stroke Project. *Stroke* 1990;21: 848-53.
- 2 Hankey GJ, Slattery JM, Warlow CP. The prognosis of hospital-referred transient ischaemic attacks. *J Neurol Neurosurg Psychiatry* 1991;54:793-802.
- 3 Collins R, Peto R, MacMahon S, *et al*. Blood pressure, stroke and coronary heart disease. Part 2, short-term reductions in blood pressure: overview of randomised

- drug trials in their epidemiological context. *Lancet* 1990;335:827-38.
- 4 Peto R, Yusuf S, Collins P. Cholesterol lowering trial results in their epidemiological context (abstract). *Circulation* 1985;72(suppl III):III-451.
 - 5 Antiplatelet Trialists' Collaboration: Secondary prevention of vascular disease by prolonged antiplatelet treatment. *BMJ* 1988;296:320-31.
 - 6 Warlow CP. Ticlopidine, a new antithrombotic drug: but is it better than aspirin for long term use? *J Neurol Neurosurg Psychiatry* 1990;53:185-187.
 - 7 European Carotid Surgery Trialists' Collaborative Group. MRC European Carotid Surgery Trial: interim results for symptomatic patients with severe (70-99%) or with mild (0-29%) carotid stenosis. *Lancet* 1991;337:1235-43.
 - 8 North American Symptomatic Carotid Endarterectomy Trial (NASCET) collaborators: Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. *N Engl J Med* 1991;325:445-53.
 - 9 Hankey GJ, Slattery JM, Warlow CP. Transient ischaemic attacks: which patients are at high (and low) risk of serious vascular events? *J Neurol Neurosurg Psychiatry* 1992;55:640-52.
 - 10 Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules: applications and methodological standards. *N Engl J Med* 1985;313:793-9.
 - 11 Sackett DL, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology: a basic science for clinical medicine*, 2nd ed., Boston, Little, Brown, 1991.
 - 12 Kernan WN, Horwitz RJ, Brass LM, Viscoli CM, Taylor KJW. A prognostic system for transient ischaemic attacks or minor stroke. *Ann Intern Med* 1991;114:552-7.
 - 13 UK-TIA Study Group. The United Kingdom transient ischaemic attack (UK-TIA) aspirin trial: final results. *J Neurol Neurosurg Psychiatry* 1991;54:1044-54.
 - 14 Dennis MS, Bamford JM, Sandercock PAG, Warlow CP. Incidence of transient ischaemic attacks in Oxfordshire, England. *Stroke* 1989;20:333-9.
 - 15 Dennis MS, Bamford JM, Sandercock PAG, Warlow CP. A comparison of risk factors and prognosis for transient ischaemic attacks and minor ischaemic strokes. The Oxfordshire Community Stroke Project. *Stroke* 1989;20:1494-9.
 - 16 Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Statistics in Medicine* 1984;3:143-152.
 - 17 Kaplan EL, Meier P. Nonparametric estimations from incomplete observations. *Journal of the American Statistical Association* 1958;53:457-81.
 - 18 Harrell FE Jr, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA* 1982;247:2543-6.
 - 19 Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-35.
 - 20 de Dombal FT. Analysis of symptoms in the acute abdomen. *Clinics in Gastroenterology* 1985;14:31-43.
 - 21 Titterton DM, Murray GD, Murray LS, et al. Comparison of discrimination techniques applied to a complex data set of head injured patients. *J Royal Stat Soc* 1981;A144:145-175.
 - 22 The Dutch TIA Trial Study Group. A comparison of two doses of aspirin (30 mg vs. 283 mg a day) in patients after a transient ischaemic attack or minor ischaemic stroke. *N Engl J Med* 1991;325:1261-6.
 - 23 Rose G. Strategy of prevention: Lessons from cardiovascular disease. *BMJ* 1981;282:1847-50.
 - 24 Rose G, Day S. The population mean predicts the number of deviant individuals. *BMJ* 1990;301:1031-4.