

Transient ischaemic attacks: which patients are at high (and low) risk of serious vascular events?

Graeme J Hankey, James M Slattery, Charles P Warlow

Abstract

The aims of this study were to determine the important prognostic factors at presentation which identify patients with transient ischaemic attacks (TIA) who are at high risk (and low risk) of serious vascular events and to derive a prediction model (equation) for each of the major vascular outcome events.

A cohort of 469 TIA patients referred to a University hospital, without prior stroke, were evaluated prospectively and followed up over a mean period of 4.1 years (range 1-10 years). The major outcome events of interest were 1) stroke 2) coronary event and 3) stroke, myocardial infarction or vascular death (whichever occurred first). Prognostic factors and their hazard ratios were identified by means of the Cox proportional hazards multiple regression analysis. The significant adverse prognostic factors (in order of strength of association) for stroke were an increasing number of TIAs in the three months before presentation, increasing age, peripheral vascular disease, left ventricular hypertrophy and TIAs of the brain (compared with the eye); the prognostic factors for coronary event were increasing age, ischaemic heart disease, male sex, and a combination of carotid and vertebrobasilar TIAs at presentation; and for stroke, myocardial infarction or vascular death they were increasing age, peripheral vascular disease, increasing number of TIAs in the three months before presentation, male sex, a combination of carotid and vertebrobasilar TIAs at presentation, TIAs of the brain (compared with the eye), left ventricular hypertrophy and the presence of residual neurological signs after the TIA. Prediction models (equations) of both the relative risk and absolute risk of each of the major outcome events were produced, based on the presence or level of the significant prognostic factors and their hazard. Before it can be concluded that our equations accurately predict prognosis and can be generalised to other populations, their predictive power needs to be validated in other, independent samples of TIA patients (which we are currently doing).

(*J Neurol Neurosurg Psychiatry* 1992;55:640-652)

Stroke is an enormous problem for the affected individual, carers and society.^{1,2} Besides good

nursing and rehabilitation, current treatments for acute stroke have not been shown to have a major impact on case fatality and disability. There is much greater potential for reducing the burden of stroke on the community by stroke prevention. As about half of all strokes can be attributed to high blood pressure³⁻⁷ and up to a third to cigarette smoking,⁷⁻¹² the first priorities in stroke prevention are to reduce the mean blood pressure of the population,¹³ to detect and control hypertension^{14,15} and to discourage cigarette smoking in the community. Another potential avenue of stroke prevention is the 10-15% of stroke patients in the community who have experienced a transient ischaemic attack (TIA) before their stroke;^{3,5,16} if these people can be identified and treated immediately after their TIA then it may be possible to prevent, or at least delay, the occurrence of stroke. Measures of secondary prevention after TIA include management of vascular risk factors¹⁷ (such as hypertension,¹⁵ smoking¹⁸ and hypercholesterolaemia,¹⁹) longterm antithrombotic treatment^{20,21} and, in the appropriate patient, carotid endarterectomy.^{22,23} However, although patients with TIAs, as a group, have an increased risk of stroke and other serious vascular events (such as, coronary events), their prognosis as individuals is extremely variable; some have a remarkably benign clinical course while others proceed to repeated or fatal vascular insults.^{24,25} It is therefore not appropriate to investigate and treat all TIA patients in the same way; higher risk and more expensive investigations and treatments (such as carotid angiography and endarterectomy) should be reserved for high risk patients in addition to lower risk and less expensive treatments that should be directed to a much wider range of patients. The crucial question is: "how can we identify the high (and low) risk patient?"

What influences prognosis?

The prognosis of TIA patients is likely to be influenced by the cause of the TIA and the presence and level of prognostic factors (including treatments) in any given patient. TIAs may be a symptom of any one of a heterogeneous group of disorders (such as large vessel atheromatous thrombo-embolism, small vessel lipohyalinosis or microatheroma, cardiac embolism, arteritis, hypercoagulable states), each of which may confer a different prognosis. However, many previous studies of TIA prognosis have tended to "lump" all TIA patients together, irrespective of underlying pathogenesis. Nevertheless, even in prognostic

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

GJ Hankey
JM Slattery
CP Warlow

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

Received 4 September 1991
and in final revised form
11 November 1991.
Accepted 27 November
1991

studies limited to more homogeneous cohorts pathogenically, such as patients with TIA due only to presumed atherothromboembolism²⁶ or patients with TIA due to atherothromboembolism, lipohyalinosis or cardiogenic embolism,²⁵ the prognosis was quite variable (among individual TIA patients), due mainly to differences in the prevalence and level of prognostic factors.

Which prognostic factors are important?

In the past, several studies²⁷⁻³¹ have tried to identify adverse prognostic factors without taking into account potential confounding by other variables. As a result, a stronger predictive role may have been assigned to factors that were merely "markers" for other factors of greater importance. Many techniques can be employed to allow for the influence of confounding variables. Matching and stratification can consider only a few variables at a time, and then only by sacrificing statistical power. Multiple regression can be used to adjust (control) simultaneously for the effects of many variables to determine the "independent" effects of one of them. Predictors are those variables that have an independent correlation with the dependent variable (for example, outcome) after controlling for potential confounding variables.

Only seven prognostic studies³²⁻⁴¹ and one randomised treatment trial⁴² have searched for prognostic factors in TIA patients by means of multiple regression analysis and two prognostic studies have described partial correlations after regression analysis with adjustment only for age,⁴³ or age (< 65, ≥ 65) and sex.²⁶ The results are difficult to compare because of the considerable heterogeneity among the studies in terms of their aims, the patients studied, and the primary measures of outcome. In only one of these studies⁴¹ was a prediction model (equation) derived but, as it was based on a retrospective study of a rather small number of outcome events (38 patients with stroke or death), it is perhaps not surprising that the results were not successfully validated in another (test) sample of patients.

The aims of this study were to identify, from a large prospectively studied cohort of TIA patients, the important prognostic factors at clinical presentation and to use these factors to derive a prediction model (equation) for important vascular outcome events.

As this study was performed in several steps and each step was determined by the outcome of the preceding step, we have elected to present and discuss the methods and results together, one step at a time.

The patient characteristics, methods, definitions, treatments and survival analysis of this cohort of 481 hospital-referred TIA patients, without prior stroke, have been described elsewhere.²⁵ Briefly, an "inception cohort" of 481 TIA patients who presented to one of us (CPW) between 1977 and 1986 was assembled prospectively and followed up completely until the patients died or until the end of 1986.

Methods and results

Statistical analysis was confined to the 469 patients (98%) in whom the cause of the presenting TIA was considered to be atheromatous thromboembolism, lipohyalinosis or cardiogenic embolism.

The major outcome events of interest were: 1) first-ever stroke (n = 63), with censoring of non-stroke deaths; 2) coronary event (n = 58), censoring non-coronary event deaths, and 3) stroke, myocardial infarction or vascular death (n = 118), whichever occurred first, censoring non-vascular deaths (for definitions see.²⁵) Stroke was of obvious interest; coronary event was considered because it is the most common cause of death in TIA patients and shares pathogenic factors and treatment options with stroke; and the composite outcome event stroke, myocardial infarction or vascular death was analysed because it combines events that are due mainly to atheromatous thromboembolism, it is the most frequent and therefore the most statistically stable outcome event, it is the most likely outcome event to be independent of the causes of censorship (such as non-vascular deaths and end of study censorship), and it is being used increasingly in other studies (because of the above reasons).²⁰ Although we recorded other subdivisions or combinations of outcome events, these are not presented because either the number of outcome events was too small (for example, early stroke, within 100 days of entry [n = 18] and ipsilateral ischaemic stroke [n = 35]), the inclusion of death from non-vascular events, such as cancer and suicide, impaired the sensitivity of the analysis, or the analysis of the outcome event produced no additional useful information (such as, disabling stroke and vascular death).

The analysis of potential prognostic factors for the three major outcome events was performed in four steps (and some sidesteps), followed by the development of the prediction model (equation) and then finally the internal validation of the model.

Step 1 (table 1)

Single variable regression analysis (descriptive statistics)

We initially calculated raw measures of association between each potential prognostic factor and the three major outcome events: a) stroke; b) coronary event and c) stroke, myocardial infarction or vascular death, without adjusting for the effects of other prognostic variables. The significant ($p < 0.05$) raw associations are presented in two ways in table 1 for the composite outcome event stroke, myocardial infarction or vascular death; the non-significant ($p > 0.05$) raw associations are available (unpublished), as are all raw associations for the other outcome events.

Firstly, the frequency of occurrence of each major outcome event was calculated according to the presence and absence (dichotomous variables), or level (continuous variables), of each baseline variable. The data concerning continuous baseline variables (such as age) are listed in an ordered categorical form. The

Table 1 a) Raw data of the proportion of patients with each significant ($p < 0.05$) prognostic variable who suffered the composite vascular outcome event: stroke, myocardial infarction or vascular death and b) Hazard ratios of the outcome event for each significant ($p < 0.05$) prognostic factor

Factors	Categories	Patients No	Stroke, myocardial infarction or vascular death			
			%	Hazard ratio	95% confidence interval or HR	p Value
DEMOGRAPHIC FACTORS						
Age (years)	< 60	186	17.2			
	60-70	165	26.1	1.67*	1.3 to 2.1*	0.001
	> 70	118	36.4			
Sex	Male	317	28.7			
	Female	152	17.8	0.54	0.4 to 0.8	0.005
VASCULAR DISEASES AND RISK FACTORS						
Peripheral vascular disease	No	390	21.3			
	Yes	79	44.3	2.83	1.9 to 4.2	0.001
Left ventricular hypertrophy (ECG)	No	415	22.4			
	Yes	46	47.8	2.57	1.6 to 4.1	0.001
Carotid bruit: left and/or right	No	328	20.1			
	Yes	141	36.9	2.24	1.6 to 3.2	0.001
Carotid bruit: left and/or right	No	328	20.1			
	Unilateral	81	37.0			
	Bilateral	60	36.7	2.11	1.3 to 3.4	0.002
Symptomatic carotid bruit	No	361	22.2			
	Yes	108	35.2	1.89	1.3 to 2.8	0.001
Asymptomatic carotid bruit	No	390	23.1			
	Yes	75	36.0	1.89	1.2 to 2.9	0.004
Arcus Senilis	No	241	19.5			
	Yes	221	30.8	2.03	1.4 to 2.9	0.001
Systolic blood pressure (mmHg)	lower quartile: 100-135	118	19.5			
	middle half: 135-175	227	26.9	1.11*	1.04 to 1.17*	0.001
	upper quartile: 175-245	124	27.4			
Hypertension	No	271	22.1			
	Yes	198	29.3	1.49	1.04 to 2.1	0.031
Ischaemic heart disease	No	370	23.0			
	Yes	99	33.3	1.56	1.04 to 2.3	0.031
TIA CHARACTERISTICS						
Carotid and Vertebrobasilar TIA at presentation	No	433	22.9			
	Yes	36	52.8	2.74	1.7 to 4.5	0.001
No of TIAs in last month (brain and eye)	lower quartile: 0-1	276	22.1			
	middle half: 1-2	73	27.4	1.42*	1.2 to 1.7*	0.001
	upper quartile: > 2	120	30.8			
No of TIAs in last 3 months (brain and eye)	lower quartile: 0-1	197	19.8			
	middle half: 1-4	151	25.2	1.20*	1.1 to 1.3*	0.001
	upper quartile: > 4	121	33.9			
Carotid TIA brain and VB TIA only**	Yes	21	5.2	2.53	1.4 to 4.7	0.003
	Yes	148	16.9	0.52	0.3 to 0.8	0.004
TIA eye (AFx) only**	No	158	19.0			
	Yes	311	28.3	1.67	1.1 to 2.5	0.015
TIA brain at presentation	No	443	23.9			
	Yes	26	46.2	2.04	1.1 to 3.7	0.020
INVESTIGATIONS						
ICA stenosis (%), symptomatic side	lower quartile: 0-10	65	6.2			
	middle half: 10-69	107	23.4	1.22*	1.1 to 1.3*	0.001
	upper quartile: 70-100	57	31.6			

* Hazard ratios of continuous variables for an increment (or decrement) of 10 units (e.g. the hazard of a stroke, MI or vascular death is 1.67 times greater for a 70 year old than a 60 year old.)

**The hazard ratio of each category of the site of the TIA was calculated relative to the other sites combined.

Footnote: the non-significant factors ($p > 0.05$) analysed were: diastolic blood pressure, cardiac failure, diabetes mellitus, cigarette smoking, number of cigarettes smoked each day, body mass index, family history of ischaemic heart disease or cerebrovascular disease, cardiomegaly, alcohol intake, atrial fibrillation, valvular heart disease, migraine, TIA brain at presentation, carotid TIA brain and eye only, carotid TIA brain only, carotid TIA brain and eye and vertebrobasilar TIA, vertebrobasilar TIA only, vertebrobasilar TIA and TIA eye only, ipsilateral TIA brain and eye, duration of longest TIA, residual neurological signs, TIA due to atheromatous thromboembolism or lipohyalinosis, TIA due to cardiac embolism, TIA due to thromboembolism and cardiac embolism, haematocrit, plasma triglycerides, plasma cholesterol, cranial CT evidence of infarction on the symptomatic and asymptomatic sides, internal carotid artery (ICA) stenosis on the asymptomatic side, aspirin therapy, and carotid endarterectomy. The results of the single variable regression analysis of these are available (unpublished).

results are expressed as the percentage of patients with a certain factor (for example, peripheral vascular disease) and without the factor who suffered the outcome event of interest (such as stroke) at any time during follow up. Although these raw data do not account for the varying period of follow up or time to the outcome event for each patient, nor for the effects of other variables, these "hard" numbers often help the clinician to understand and accept the more obscure types of statistical analysis.

Secondly, a hazard ratio (and its 95% confidence interval and p value) for each outcome event was calculated for each potential prognostic factor. The varying period of follow up, or time to the outcome event, for each patient was accounted for but not the effects of other variables. This analysis allows for comparison of our data set with other studies that have concentrated on investigating

the prognostic power of single variables. The "hazard" is the instantaneous probability of the measure of outcome, such as stroke, happening in unit time, at a specific time, conditional on survival until that time. The "hazard" ratio is the ratio of the probability of the outcome event occurring at some point in time (conditional on survival until that time) in patients with a given (prognostic) factor to the probability of the outcome event in patients without that factor. For continuous variables, the hazard ratio is presented for changes of 10 units in the level of the variable (for example, 10 years for age, 10 mmHg for blood pressure, 10% for carotid stenosis). These data were obtained from a Cox proportional hazards analysis (see below)⁴⁴⁻⁴⁷ and the 95% confidence interval of the hazard ratio was calculated using standard techniques: exponential (beta coefficient $\pm 1.96 \times$ standard error)⁴⁸ (see step 3).

Step 2 (table 2)

Selection of variables for multiple (variable) regression analysis

Variables were selected for inclusion in the multiple regression analysis on the basis of: a) the strength of their association with the outcome event (step 1); b) the biological plausibility of the association; c) what was already known, or had been suggested about the association; d) the number of patients and outcome events with which the variable was associated, and e) convenience (table 2). When several variables were closely associated, one was chosen which contained the information common to all of these variables.

Measurements of carotid stenosis were not included in the initial Cox model because information was available only on the 229 patients (49%) who had angiography of the carotid system on the symptomatic side. The only other variable with more than 2% of its

values missing was body mass index (24 missing values out of 469). Nine other potential prognostic variables had less than 2% of their values missing, of which only one, left ventricular hypertrophy (8 missing values out of 469) is used in our prediction models (see below). Values for missing variables, other than carotid stenosis, were imputed by means of the BMDP program PAM⁴⁹ before survival analysis because most computer programmes that carry out the Cox regression model analysis automatically exclude all patients in a series missing even one piece of relevant data.

Step 3 (table 3)

Cox proportional hazards multiple regression analysis

A Cox proportional hazards multiple regression analysis was performed by means of a BMDP computer programme (P2L).⁴⁴⁻⁴⁷ This method of statistical analysis for survival data was preferred because it is more powerful and versatile than the log rank test, which has long been the standard tool for comparing survival curves.⁴⁷

To understand and interpret the results of the Cox analysis it is essential to be familiar with the concept of a "hazard", the assumptions made and the limitations of the model.⁴⁷ A hazard at time t, is written as h(t), and is the probability in unit time of a patient suffering the outcome event of interest at time t, given that he or she has survived up to time t. The Cox analysis assumes that the hazard of a patient with a set of prognostic variable values z at time t is related to a baseline hazard function h₀(t) by the expression: h(t; z) = h₀(t) x exp(β'z), where β (beta) is a vector of unknown regression coefficients. The model allows us to estimate β (beta) and so estimate the hazard of any patient relative to the baseline hazard function h₀ and hence, more usefully, the hazard of any patient relative to another patient (the hazard ratio).

Table 3 shows only the statistically significant (p < 0.05) prognostic variables for the three major outcome events, after adjusting for the effects of all the other variables listed in

Table 2 Factors included in the Cox proportional hazards multiple regression analysis

Dichotomous variables (yes/no)	Continuous variables
UNAVOIDABLE DEMOGRAPHIC FACTORS	
Sex (male/female)	Age (years)
VASCULAR DISEASES AND RISK FACTORS	
Left ventricular hypertrophy (ECG)	Systolic blood pressure (mmHg)
Peripheral vascular disease	Body mass index
Diabetes mellitus	Cigarette smoking (Number/day)
Carotid bruit: any	Alcohol intake (units/day)
Arcus Senilis	
Cardiomegaly (CXR)	
Ischaemic heart disease	
Atrial fibrillation	
Migraine	
TIA CHARACTERISTICS	
Carotid and vertebralbasilar TIA at presentation	Duration of longest TIA (minutes)
Vertebrobasilar TIA only	No of TIAs in last 3 months
Amaurosis fugax (AFx) only	
TIA brain and AFx in same patient	
Residual neurological signs	
INVESTIGATIONS	
Low haematocrit (bottom 10%)	ICA stenosis (%), symptomatic side*
High haematocrit (top 10%)	Plasma Cholesterol (mmol/l)
Infarct on CT, asymptomatic side	
Infarct on CT, symptomatic side	
TREATMENT	
Aspirin therapy	
Carotid endarterectomy	

*only included in separate analysis of the subset of 229 patients who had carotid angiography.

Table 3 Results of Cox proportional hazards multiple regression analysis of each variable in table 2, adjusted for all other variables, showing the significant (p < 0.05) prognostic factors for each of the three major vascular outcome events

Factor	Beta coefficient	Coefficient/standard error	Hazard ratio	Hazard ratio** (for multiples of 10)
STROKE, MYOCARDIAL INFARCTION OR VASCULAR DEATH (n = 118)				
Age*	0.06	3.98	1.06	1.73
No of TIAs in last 3 months*	0.01	3.24	1.01	1.15
Male Sex	0.75	2.86	2.12	
Peripheral vascular disease	0.68	2.73	1.98	
Carotid and vertebralbasilar TIAs	0.70	2.35	2.00	
Residual signs	0.73	2.10	2.08	
Amaurosis fugax only	-0.69	-2.59	0.50	
STROKE (n = 63)				
No of TIAs in last 3 months*	0.01	2.63	1.01	1.11
Age*	0.04	2.09	1.04	1.49
CORONARY EVENT (n = 58)				
Age*	0.08	3.96	1.08	2.26
Male Sex	1.24	2.98	3.46	
Carotid and vertebralbasilar TIAs	1.06	2.59	2.90	

*Continuous variables

**Hazard ratios of continuous variables for increment of 10 units (for example, for 10 TIAs in last 3 months, or 10 years of age, or 10 mm Hg systolic BP)

table 2. Estimates of the regression parameters (the beta coefficient) were produced for each variable, together with a standard error and p value. For factors in which the ratio of the beta coefficient to the standard error exceeds 1.96 or is less than -1.96, the p value is less than 0.05. The beta coefficients were converted into hazard ratios by calculating the exponential (antilogarithm) of the coefficient.

Sidestep 3.

The influence of carotid stenosis.

A separate multiple regression analysis was performed (of all of the variables listed in table 2 for each of the major outcome events) in the smaller subset of the 229 patients for whom angiographic measurements of diameter stenosis of the origin of the internal carotid artery on the symptomatic side were available. Although the number of patients and outcome events were smaller in this selected subgroup and therefore, the results were more prone to sampling error, it was quite clear that in this group of patients, after adjustment for the effects of the other variables, the degree of carotid stenosis (as a continuous variable) was the most significant adverse prognostic factor for stroke (hazard ratio 1.69 for each 10% increment of carotid stenosis) and stroke, myocardial infarction or vascular death (hazard ratio 1.46). It was also a significant

prognostic factor for coronary event (hazard ratio 1.31). As these 229 patients were not a representative sample of the 469 patients (they were a select group who were considered potential candidates for carotid endarterectomy on clinical grounds),²⁵ the results have not been tabulated in this paper but are available (unpublished).

Step 4 (table 4)

Stepwise Cox proportional hazards multiple regression analysis

The next stage of the analysis was a forward stepwise selection procedure aimed at deriving the constellation of prognostic factors and their associated hazard ratios that best predicted the occurrence of each outcome event. The "stepwise regression" process is a method of forcing single factors and combinations of factors to compete against each other in a sort of statistical tournament. The most commonly used stepwise procedure first selects the best individual prognostic factor and then, one by one, adjoins complementary factors to complete the model.

From the multiple regression analysis of all variables (step 3), it was decided that the variables shown to be significant ($p < 0.05$) "independent" prognostic factors for stroke, myocardial infarction or vascular death (the most statistically robust outcome event) should

Table 4 Prediction models for relative risk of each outcome event derived from stepwise Cox proportional hazards multiple regression analysis of survival data of 469 patients with TIA
Selection criteria: variables entered if $p < 0.03$ and removed if $p > 0.05$

STROKE, MYOCARDIAL INFARCTION OR VASCULAR DEATH (n = 118 events)					
Factor	Beta coefficient	Coefficient/standard error	p-value	Hazard ratio	Hazard ratio 95% confidence interval
**Age	0.0598	5.29	< 0.001	1.06 (1.82)	1.04 to 1.09
**Peripheral vascular disease	0.8355	3.92	< 0.001	2.31	1.52 to 3.50
**Number of TIAs in last 3 months	0.0151	3.65	< 0.001	1.02 (1.16)	1.01 to 1.02
**Male sex	0.6820	3.03	0.002	1.98	1.27 to 3.07
**Carotid and VB TIAs	0.7078	2.64	0.008	2.03	1.20 to 3.43
**TIA brain (vs AFx only)	0.5572	2.35	0.019	1.75	1.10 to 2.78
*Left ventricular hypertrophy (ECG)	0.5404	2.22	0.026	1.72	1.07 to 2.76
**Residual neurological signs	0.6566	2.01	0.044	1.93	1.02 to 3.65
<i>No additional variables selected using the less stringent criteria: enter if $p < 0.10$ and remove if $p > 0.15$</i>					
STROKE (n = 63 events)					
<i>Model with significant prognostic variables for stroke, MI or vascular death forced in</i>					
Number of TIAs in last 3 months	0.0158	3.37	0.001	1.02 (1.17)	1.01 to 1.03
Age	0.0452	2.99	0.003	1.05 (1.57)	1.03 to 1.08
Peripheral vascular disease	0.7568	2.57	0.010	2.13	1.20 to 3.80
Left ventricular hypertrophy (ECG)	0.6805	2.11	0.035	1.98	1.05 to 3.72
TIA brain (vs AFx only)	0.7179	2.08	0.038	2.05	1.04 to 4.03
Residual neurological signs	0.7389	1.80	0.072	2.09	0.94 to 4.69
Carotid and VB TIAs	0.5265	1.43	0.151	1.69	0.82 to 3.48
Male sex	0.3600	1.23	0.217	1.43	0.81 to 2.54
<i>Additional variables selected using the less stringent criteria: enter if $p < 0.10$ and remove if $p > 0.15$: Ischaemic heart disease, carotid bruit</i>					
CORONARY EVENT (n = 58 events)					
<i>Model with significant prognostic variables for stroke, MI or vascular death forced in</i>					
Age	0.0810	4.68	< 0.001	1.08 (2.25)	1.05 to 1.12
*Ischaemic heart disease	0.8726	3.06	0.002	2.39	1.37 to 4.18
Male sex	1.0200	2.92	0.003	2.77	1.40 to 5.50
Carotid and VB TIAs	1.0680	2.90	0.004	2.91	1.41 to 5.99
Peripheral vascular disease	0.5944	1.86	0.063	1.81	0.97 to 3.39
Number of TIAs in last 3 months	0.0097	1.50	0.117	1.01 (1.10)	0.998 to 1.02
TIA brain (vs AFx only)	0.4403	1.34	0.180	1.55	0.82 to 2.95
Residual neurological signs	0.4791	0.96	0.339	1.62	0.60 to 4.31
Left ventricular hypertrophy (ECG)	0.3321	0.94	0.347	1.39	0.70 to 2.78
<i>No additional variables selected using the less stringent criteria: enter if $p < 0.10$ and remove if $p > 0.15$</i>					

**Statistically significant prognostic variables for stroke, MI, vascular death (see table 3) that were forced into model.

*Additional prognostic variables selected using stepwise procedure (enter if $p < 0.03$, remove if $p > 0.05$) after the significant variables had been forced into the model.

Hazard ratios for increases of 10 units in continuous variables, such as age and number of TIAs, are shown in brackets. (e.g. for an increase in age of one decade, the hazard ratio for stroke, MI or vascular death increases by a factor of 1.82) (e.g. for every 10 TIAs in the last 3 months the hazard ratio for stroke, MI or vascular death increases by a factor of 1.16).

Note: these beta coefficients and hazard ratios that have been derived from the stepwise analysis, adjusting for variables in the model, are different from those in table 3, derived after adjusting for all variables.

be forced into the prediction model for stroke, myocardial infarction or vascular death. Thereafter, each of the other factors listed in table 3 was provisionally entered into the model and the best additional predictor (for example, left ventricular hypertrophy) was selected. In turn, a variable was entered into (or removed from) the regression equation on the basis of a computed significance probability. Variables were entered into the model only if the *p* value was less than 0.03 and were removed if the *p* value exceeded 0.05. The variables chosen in the final stepwise regression model for the composite outcome event of stroke, myocardial infarction or vascular death were then forced into the models for stroke and for coronary event. Otherwise, a variable selected in the model for stroke, for instance, which is not in the model for the composite outcome event may result in a worse predicted outcome for stroke than for stroke, myocardial infarction or vascular death, which would not seem plausible.

Prediction model (equation) of relative risk

The relative risk of a patient (with certain prognostic factors) suffering an outcome event compared with another patient (who does not have these factors) can be derived by multiplying the hazard ratios of the relevant important prognostic factors that have been derived from the Cox analysis (table 4). For example, a male patient with 10 TIAs of the brain involving both carotid and vertebrobasilar territories and bearing all of the adverse prognostic factors for stroke, myocardial infarction or vascular death had a 114 times greater risk (hazard ratios for each factor: $1.82 \times 2.31 \times 1.16 \times 1.98 \times 2.03 \times 1.75 \times 1.72 \times 1.93$) than a woman who was 10 years younger, who did not have peripheral vascular disease (PVD), left ventricular hypertrophy (LVH), or residual signs and who had no TIAs of a different nature before an isolated episode of amaurosis fugax (table 4). It is clear that as the number of adverse prognostic factors increases, the patient's prognosis worsens.

Sidestep 4a

The influence of carotid stenosis

Although it was not possible to enter the variable "percentage diameter carotid stenosis on the symptomatic side" into the Cox proportional hazards multiple regression analysis (because data were present in only 229 patients), it was possible to analyse the extent to which carotid stenosis enhanced the predictive power of the model (that had already been developed from the other factors [table 4]) by means of a computer package called "EGRET".⁵⁰ The addition of information about carotid stenosis to the original prediction model (table 4) did not significantly improve the predictive power of the model. This does not mean that carotid stenosis is not an important prognostic factor on its own; it means that a knowledge of the degree of carotid stenosis does not add significantly to the accuracy of the patient's predicted outcome that is based on a knowledge of the

presence or level of the other important prognostic factors (such as age, PVD, LVH) in the prediction models.

Sidestep 4b

Alternative models

Four alternative models were generated to test the sensitivity of our model to the selection procedure used.

The first was to enter all the variables listed in table 2 into the model and allow the model to include and exclude variables freely (without forcing the significant variables from the initial analysis into the model). The results were very similar to those presented. This is not surprising as the pre-specified variables for entry in the model that is presented were not only significant but also, we believe, medically plausible.

The second was to "redefine" the arbitrary entry ($p < 0.03$) and removal ($p > 0.05$) criteria to $p < 0.10$ and $p > 0.15$ respectively. Using the less stringent criteria, more variables were selected for stroke (table 4) but the 95% confidence intervals of their hazard ratios were wide, reflecting greater uncertainty and a higher chance of sampling error.

The third was to force only the significant variables from the multiple regression analysis of each outcome event (table 3) into the respective models. Different variables were selected in different models making it theoretically possible that the predicted risk of stroke in a given patient could be higher than the predicted risk of stroke, myocardial infarction or vascular death, which is not realistic.

The fourth was to force the significant variables for stroke and coronary event (table 3) into the model for stroke, myocardial infarction or vascular death to ensure consistency among the models. Only ischaemic heart disease (IHD) was added to the model and its influence was very small (the beta coefficient for IHD was -0.088 and the hazard ratio was 1.09).

From a clinical and statistical point of view, these alternative methods and resulting models were therefore not considered as sound as the models presented.

Step 5 (tables 5a, 5b, 5c, 5d; fig 1)

Prediction model (equation) of absolute risk.

There are two reasons why results from studies such as this are often presented primarily in terms of relative risks rather than absolute risks. The first is that it is necessary only to make an assumption about the form that the relative risk function takes (that is, the way the prognostic variables modify the hazards or risks) to be able to pick out which prognostic variables are best able to predict outcome events. The second is that, if attempts are made to calculate absolute risks from data on a group of patients, it must be assumed that other patients to whom the model will be applied are generally similar to the group from which it was derived (in terms of unknown prognostic factors) rather than making the weaker assumption that they are similar in their response to the presence or absence of a

Table 5a Prediction models of absolute risk: method of calculating the survival free of stroke, myocardial infarction or vascular death

STROKE, MYOCARDIAL INFARCTION OR VASCULAR DEATH (n = 118 events)			
Prognostic factors		Mean (for the whole cohort)	Beta coefficient (from table 4)
Age	(years)	62.52	0.0598
Peripheral vascular disease	(0 = no, 1 = yes)	0.17	0.8355
Number of TIAs in last 3 months	(number)	6.59	0.0151
Male sex	(1 = male, 2 = female)	1.32	-0.6820*
Carotid and VB TIAs	(0 = no, 1 = yes)	0.08	0.7078
TIA brain (vs AFx only)	(0 = no, 1 = yes)	0.68	0.5572
Residual neurological signs	(0 = no, 1 = yes)	0.06	0.6566
Left ventricular hypertrophy (ECG)	(0 = no, 1 = yes)	0.10	0.5404
The "average" patient had the above mean values of the prognostic factors used in the model (i.e. was 62.52 years old etc.).			
Survival free of stroke, MI or vascular death			
for the "average" patient with mean values of prognostic variables (as above)			
(derived from Cox proportional hazards multiple regression analysis, not Kaplan-Meier technique)			
at 1 year, S(1) = 0.9396 (i.e. 93.96%)			
at 5 years, S(5) = 0.7615 (i.e. 76.15%)			
Hazard ratio (X) for 60 year old man with one TIA brain and no prognostic factors (h 60) compared with "average" patient (h mean)			
Log (h 60 / h mean)**		= (0.0598 (60-62.52) + 0.8355 (0 - 0.17) + 0.0151 (1-6.59) - 0.6820 (1-1.32) + 0.7078 (0-0.08) + 0.5572 (1-0.68) + 0.5572 (1-0.68) + 0.6566 (0-0.06) + 0.5404 (0-0.10)) = (-0.151 - 0.142 - 0.0844 + 0.2182 - 0.0566 + 0.178 - 0.0394 - 0.054) = -0.131	
h 60 / h mean		= exp (-0.131)	
X		= 0.88	
Survival free of stroke, MI or vascular death			
for a 60 year old man with one TIA and none of the prognostic factors			
at one year S(1) ^x		= $\frac{0.9396 \times 0.88}{1}$ = 0.9468	
at five years S(5) ^x		= $\frac{0.7615 \times 0.88}{1}$ = 0.7875	

*The negative coefficient for sex means male sex is an adverse prognostic factor when coded as: 1 = M, 2 = F.
 **"h 60" is the hazard for a 60 year old man with one TIA brain and no prognostic factors.
 ***"h mean" is the hazard for the "average" patient.

Table 5b Prediction models of absolute risk: for any individual TIA patient with known values of the prognostic variables in the prediction model

Survival free of STROKE, MYOCARDIAL INFARCTION OR VASCULAR DEATH

at one year S(1) = **0.9468^x**

at 5 years S(5) = **0.7875^x**

where S(1) = 0.9468 and S(5) = 0.7875 were derived as in table 5a

where X = exp (0.0598 [Age - 60] - 0.6820 [Female: yes = 1, no = 0] + 0.8355 [PVD: yes = 1, no = 0] + 0.7078 [Carotid and VB TIA: yes = 1, no = 0] + 0.0151 [Number of TIAs in last 3 months - 1] - 0.5572 [AFx only: yes = 1, no = 0] + 0.6566 [Residual signs: yes = 1, no = 0] + 0.5404 [LVH: yes = 1, no = 0])

Survival free of STROKE

at one year S(1) = **0.9584^x**

at 5 years S(5) = **0.8789^x**

where S(1) = 0.9584 and S(5) = 0.8789 were derived using the same methods as for stroke, MI, vascular death in table 5a

where X = exp (0.0452 [Age - 60] - 0.3600 [Female: yes = 1, no = 0] + 0.7568 [PVD: yes = 1, no = 0] + 0.5265 [Carotid and VB TIA: yes = 1, no = 0] + 0.0097 [Number of TIAs in last 3 months - 1] - 0.7179 [AFx only: yes = 1, no = 0] + 0.7389 [Residual signs: yes = 1, no = 0] + 0.6805 [LVH: yes = 1, no = 0])

Survival free of CORONARY EVENT

at one year S(1) = **0.9874^x**

at 5 years S(5) = **0.9158^x**

where S(1) = 0.9874 and S(5) = 0.9158 were derived using the same methods as for stroke, MI, vascular death in table 5a

where X = exp (0.0810 [Age - 60] - 1.020 [Female: yes = 1, no = 0] + 0.5944 [PVD: yes = 1, no = 0] + 0.1.068 [Carotid and VB TIA: yes = 1, no = 0] + 0.0097 [Number of TIAs in last 3 months - 1] - 0.4403 [AFx only: yes = 1, no = 0] + 0.4791 [Residual signs: yes = 1, no = 0] + 0.3321 [LVH: yes = 1, no = 0] + 0.8726 [Ischaemic heart disease: yes = 1, no = 0])

Footnote:

To calculate X for an individual patient, insert the appropriate value of each of the patient's prognostic factors into the equation.

particular set of prognostic factors. Nevertheless, estimates of absolute risk are more useful to a clinician than relative risks and they can be easily calculated from estimates of the baseline hazard. Hence, we also present prediction models (equations) of absolute risk, though with an additional caveat concerning the assumption which must be made regarding

the representativeness of our clinical sample.

To generate a prediction equation of absolute risk of survival free of an outcome event, a baseline survival function is required as a reference point. The baseline survival is usually defined as the survival for someone with zero values of the prognostic variables. An alternative and usually more meaningful reference

Table 5c The "bottom line" for clinicians

PREDICTION EQUATION for survival free of STROKE, MYOCARDIAL INFARCTION or VASCULAR DEATH			
Indicate if applicable to patient			CUMULATIVE SCORE
	START		
	(Age - 60)	MULTIPLIED BY	6
	Female	SUBTRACT	68
	Amaurosis fugax (AFx) only	ADD	56
	More than one TIA in last 3 months	ADD	(1.5 × [n - 1])
	Carotid and vertebralbasilar TIAs at presentation	ADD	71
	Peripheral vascular disease (PVD)	ADD	84
	Residual neurological signs	ADD	66
	Left ventricular hypertrophy (ECG)	ADD	54
	DIVIDE "y" BY 100 and		TOTAL = y
	EXPONENTIATE (e ^{100y})	=	X
	(note: e ^{-x} = 1 / e ^x)		

PROBABILITY of survival free of STROKE, MI or VASCULAR DEATH

At one year = 0.95^x

At five years = 0.79^x

Example: 65 year old woman with 5 episodes of AFx only in last 3 months and ECG shows LVH.

(65 - 60) × 6 = 5 × 6	=	30
Female	SUBTRACT	68 - 38
Amaurosis fugax only	SUBTRACT	56 - 94
More than one TIA in last 3 months	ADD (1.5 × [5-1])	- 88
	[1.5 × 4 = 6]	
Left ventricular hypertrophy (ECG)	ADD	54 - 34
DIVIDE (- 34) BY 100 and		
EXPONENTIATE (e ^{-0.34})		
e ^{-0.34} = 1 / e ^{0.34}	(e ^{0.34} = 1.4)	
= 1 / 1.4	=	0.7

PROBABILITY of survival free of STROKE, MI or VASCULAR DEATH

At one year = 0.95^{0.7} = 0.96 or 96%

At five years = 0.79^{0.7} = 0.85 or 85%

Note: the variables are listed in the order in which they are usually collected clinically whereas in tables 4 and 5b they are in order of the statistical strength of the association.

Table 5d The "bottom line" for clinicians

PREDICTIVE EQUATION for survival free of STROKE			
Indicate if applicable to patient			CUMULATIVE SCORE
	START		
	(Age - 60)	MULTIPLIED BY	4.5
	Female	SUBTRACT	36
	Amaurosis fugax (AFx) only	SUBTRACT	72
	Carotid and vertebralbasilar TIAs	ADD	53
	More than one TIA in last 3 months	ADD	(1.6 × [n - 1])
	Peripheral vascular disease (PVD)	ADD	76
	Left ventricular hypertrophy (ECG)	ADD	68
	Residual neurological signs	ADD	74
	DIVIDE "y" BY 100 and		TOTAL = y
	EXPONENTIATE (e ^{100y})	=	X

PROBABILITY of survival free of STROKE

At one year = 0.96^x

At five years = 0.88^x

PREDICTIVE EQUATION for survival free of CORONARY EVENT

PREDICTIVE EQUATION for survival free of CORONARY EVENT			
Indicate if applicable to patient			CUMULATIVE SCORE
	START		
	(Age - 60)	MULTIPLIED BY	4.5
	Female	SUBTRACT	102
	Amaurosis fugax only	SUBTRACT	44
	Carotid and vertebralbasilar TIAs	ADD	107
	More than one TIA in last 3 months	ADD	(0.97 × [n - 1])
	Peripheral vascular disease	ADD	59
	Left ventricular hypertrophy (ECG)	ADD	33
	Residual neurological signs	ADD	48
	Ischaemic heart disease (IHD)	ADD	87
	DIVIDE "y" BY 100 and		TOTAL = y
	EXPONENTIATE (e ^{100y})	=	X

PROBABILITY of survival free of CORONARY EVENT

At one year = 0.99^x

At five years = 0.92^x

function is the survival for someone with average values of the prognostic variables. This can be estimated means of the programme BMDP 2L which uses methods presented in Cox and Oakes (1984).⁴⁵ Even this function may not always be clinically sensible; for example, the mean value for sex was 1.32 because 32% of the cohort were female (coded 2) and 68% were male (coded 1). Thus we

have transposed the baseline survival to that of a 60 year old man with one TIA of the brain and none of the other prognostic factors in the model.

To estimate the survival probability for a patient having any combination of prognostic factors, the first step is to multiply each of the patient's prognostic variables by the appropriate beta coefficient for that variable and to

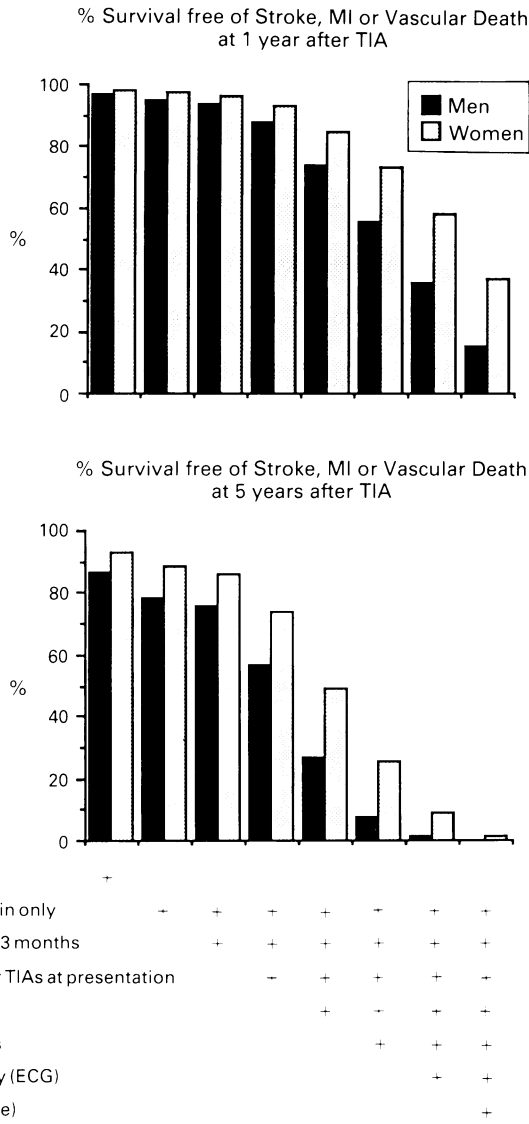


Figure 1 Bar graphs showing the probability of survival free of stroke, myocardial infarction or vascular death at one and five years after presentation with a TIA in 60 year old men and women and the impact of the important prognostic factors.

add these individual products together (tables 5c, 5d). In doing this, it is necessary to subtract 60 from the age of the chosen patient and one from the number of TIAs that the patient had suffered in the last three months. The exponential of the number derived above is an estimate of the hazard ratio of the chosen patient compared with the baseline patient. The estimate of the probability that the chosen patient will be free of the vascular outcome event of interest at one year is the one year probability for the baseline patient raised to the power of the hazard ratio. A “recipe” style approach to using the prediction models is given in tables 5c and 5d together with a calculated example (table 5c). We have not been able to identify a suitable method of calculating the 95% confidence interval of these estimates. Figure 1 illustrates the probability of survival free of stroke, myocardial infarction or vascular death for patients with different combinations of prognostic factors.

Step 6 (table 6)

Internal validation of the prediction model

Having selected statistically significant prognostic factors for a prediction model of outcome, such as this, it is natural (and important) to ask how successful the model will be in trying to predict an event using these factors. The acid test of any clinical score is its success when applied to patients who were not in the group from which it was derived (for example, its external validity)—and this is in hand—but as a first check we should establish its internal validity and answer the question: “If we had been able to apply this model to the group of patients from whom it was derived when they first came to our attention, would we have been satisfied with the predictions at some future time?” Obviously, if the answer is “No” then it is not worthwhile to test the model on other data sets.

Table 6 Internal validation of the prediction model

OUTCOME at 5 years	PREDICTED RISK		TOTAL
	< 20%	≥ 20%	
Stroke, myocardial infarction or vascular death	23	80	103
No event after 5 years follow up	82	68	150
Censored by the end of the study (before 5 years follow up completed)	76	127	203
Censored by (non-vascular) death	8	5	13
TOTAL	189	280	469
Odds ratio of an event within five years in those with a risk of ≥ 20% compared with those with a risk of < 20%			
= (80/68) / (23/82)			
= 4.2 (95% CI: 2.4 to 7.4)			
OUTCOME at 5 years	PREDICTED RISK		TOTAL
Stroke	< 20%	≥ 20%	
No event after 5 years follow up	26	31	57
Censored by the end of the study (before 5 years follow up completed)	126	29	155
Censored by (non-stroke) death	150	55	205
TOTAL	25	27	52
TOTAL	327	142	469
Odds ratio of a stroke within five years in those with a risk of ≥ 20% compared with those with a risk of < 20%			
= (31/29) / (26/126)			
= 5.2 (95% CI: 2.7 to 9.9)			
OUTCOME at 5 years	PREDICTED RISK		TOTAL
Coronary Event	< 20%	≥ 20%	
No event after 5 years follow up	18	30	48
Censored by the end of the study (before 5 years follow up completed)	132	30	162
Censored by (non-coronary) death	157	72	229
TOTAL	16	14	30
TOTAL	323	146	469
Odds ratio of a coronary event within five years in those with a risk of ≥ 20% compared with those with a risk of < 20%			
= (30/30) / (18/132)			
= 7.3 (95% CI: 3.7 to 14.5)			

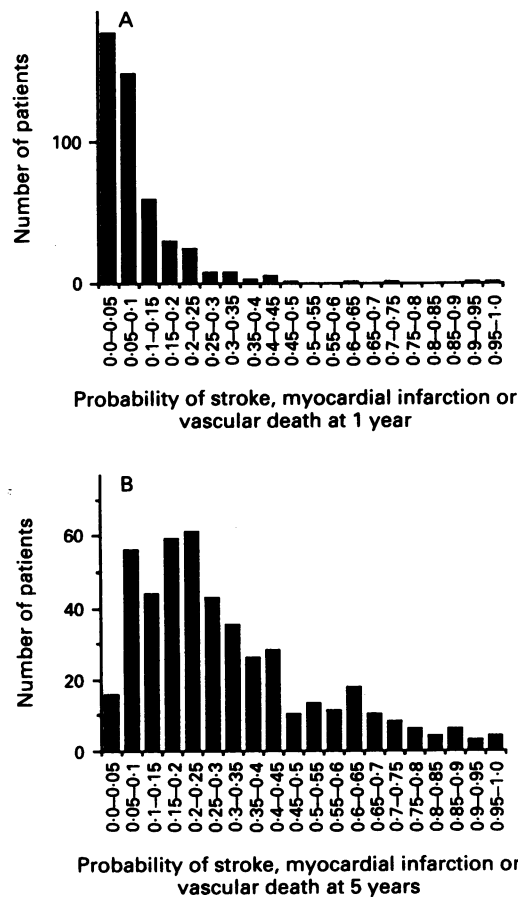


Figure 2 a) Predicted risk profile of the 469 TIA patients for stroke, myocardial infarction or vascular death at 1 year. b) Predicted risk profile of the 469 TIA patients for stroke, myocardial infarction or vascular death at 5 years.

To test the internal validity of the prediction model, the model was used to estimate the probability of survival (free of the event of interest) for each patient (fig 2) and the results were then examined to see how well they discriminated survivors from non-survivors (table 6). Patients were classified arbitrarily as "high risk" (or "low risk") if the model suggested a greater than (or less than) 20% probability of suffering the vascular outcome event of interest at a given time (we have arbitrarily chosen one and five years) after the presenting TIA. The cut-off for risk status was chosen as 20% because about 20% of TIA patients suffer a serious vascular event (such as stroke) within five years (that is, this is a rough average). Table 6 shows, for each major outcome event, the number of "high" and "low" risk patients who suffered the event, who did not suffer the event, and who were censored by either the end of the study (that is, were still alive at the end of the study) or by death. The odds ratios (OR) show that, for uncensored patients, the group identified by the model as high risk at five years for stroke, myocardial infarction or vascular death had about a 4.2 times greater odds of suffering a stroke, myocardial infarction or vascular death than the low risk group. The odds ratio was similar for stroke alone (OR: 5.2) and coronary event alone (OR: 7.3) at five years.

Discussion

The main aim of this study was to produce a prediction model (equation) of the prognosis of TIA patients based on simple clinical information obtained at the bedside which would help clinicians identify TIA patients at differential risk who require differential treatment.

Before accepting a prediction model, clinicians need to consider whether: 1) the clinical data were collected in a standardised prospective study with complete follow up; 2) the computational analysis was appropriate and applied accurately; 3) the clinicians worked closely with the statistician at all stages of the statistical analysis; 4) the prediction equation has been presented clearly and means what it says; 5) the prediction equation is likely to prove medically acceptable to clinicians (that is, it is simple, and plausible clinically) and beneficial to patients, and 6) whether the prediction equation has been validated in an independent set of patients.

In addition to meeting the first two criteria, we believe the data in this study were analysed appropriately by means of the Cox proportional hazards multiple regression analysis.⁴⁴ From a large set of variables, a smaller subset of variables was selected that independently and significantly contributed to the overall variation in outcome. They were then arranged in order of the strength of their contribution. Using these prognostic variables (and their regression coefficients) a mathematical expression (model) was developed relating these "independent" variables to outcome, for each outcome event.

The disadvantages of the Cox analysis are that the internal workings of these "mathematical tap dancing" methods are a "black box" to most clinicians and the prediction models are based on assumptions (such as the proportionality of hazards and the multiplicative risk structure) that may not be met by the data. The major assumption is that the effect of (or relative risk associated with) a prognostic variable does not change with time. For example, if the prognostic factor is a binary variable, the survival curve of one group must always be above the survival curve of the other group; they cannot cross. When the assumption of time independence is not true, as is the case of survival of patients treated, and not treated, with carotid endarterectomy,^{22 23} the Cox model selects factors associated with earlier death (or other outcome) compared with factors associated with later death (or other outcome). The results in our study should not be affected by this idiosyncrasy because patient follow up was prolonged over an average of 4.1 years, but a problem may arise in data bases containing a preponderance of patients with very short follow up.

The major potential limitation of the stepwise procedure is that it tends to examine too few combinations and consequently, the first combination selected (initial model) may not be the best model. We have attempted to overcome this problem by deriving alternative models which could possibly "outperform" the initial model, as occurs in some sports where

the best individual player/athlete (or factor) in the competition need not be a member of the best team (or combination of factors). For example, carotid stenosis was the most significant factor (the best athlete) in the small subset of patients undergoing angiography ($n = 229$) but when it was added to the final prediction model in all patients (the best team), it did not contribute significantly to the predictive ability of the model made up by the other factors (or team members). Exploring alternative models may also resolve seeming differences between prognostic models for the same disease produced by different medical centres; the alternative model of one centre may closely match the reported model of another centre. Of course, no analysis can compensate for a failure to measure important but as yet unrecognised prognostic factors that were not examined in the study. If such factors are identified in future studies (for example, perhaps racial factors, carotid ulceration and/or carotid stenosis measured by carotid ultrasound, elevated plasma fibrinogen, or impaired cerebral blood flow reactivity to hypercapnoea) then it will be important to try adding these to the model to see if it is improved.

Analysis of a small number of outcome events is prone to emphasise, by chance, some particular quirk of the data set. Such "over-fitting" in a model is particularly likely to occur when the original data sets are small because random variations due to the play of chance are more likely to distort any underlying trends. For example, in view of the recent publications of the effect of carotid endarterectomy in reducing the risk of ipsilateral ischaemic stroke in patients with carotid TIA and severe symptomatic carotid stenosis,^{22 23} we were interested to see what factors may predict ipsilateral stroke. However, only 35 of the 53 carotid ischaemic strokes (66%) were ipsilateral to the presenting TIA in our cohort and the model was discarded because of the small number of outcome events. To avoid sampling error, it has been suggested that the original data set should contain at least five (and perhaps ten⁵¹) patients with the outcome event of interest for every prognostic variable examined in the equation. In other words, if ten variables are to be analysed as possible prognostic factors for stroke, there should be at least 50 patients with stroke (and 50 survivors free of stroke) in the original data set.⁵² In our study the most robust outcome event was stroke, myocardial infarction or vascular death because there were 118 events, justifying the analysis of up to 23 variables.

Internal validity of the prediction model

The prediction model that is presented is shown in table 6 to be a satisfactory predictor of outcome events in uncensored patients in this cohort. However, it should be emphasised that although these results (of the internal validation) are encouraging, they are something of an upper bound to what can be expected when we attempt to externally validate the model in other groups of patients. The 95% confidence intervals of the odds ratios of

the internal validation (table 6) are strictly those of the internal validation and cannot be extrapolated further.

Although the calculations of the prediction models are referred to as probabilities they should be regarded in the same way as a Kaplan-Meier survival estimate and interpreted with the same qualifications, particularly when censorship may not be independent of the outcome event. As the predictors for stroke and for coronary event clearly select out vascular death, and to some extent, non-vascular death (table 6), our assumption concerning independence of the event and censorship by death may be tenuous. For instance, the calculation of the probability of stroke assumes that those who died from non-stroke vascular causes were not an abnormally high risk group for stroke. This is almost certainly not true since it is unlikely that stroke and other vascular events are biologically independent. However, death accounted for relatively few of the censorships in the study (table 6) and, provided the results of this study are not applied to populations with very different vascular death rates, it is unlikely that the practical value of the study is diminished by this. Nevertheless, the calculated "probability" of stroke, myocardial infarction or vascular death is probably the most accurate because it fails to account only for patients with non-vascular deaths (which are likely to be biologically independent events from the vascular events). Whether the models for stroke and coronary event tell us anything more than the model for stroke, myocardial infarction or vascular death will not be known until an attempt has been made to validate the models externally. Perhaps we may have exacerbated the problem by forcing in terms which predicted vascular death.

Plausibility of the prediction model

Our prediction model is not only internally valid statistically but it is also biologically and clinically plausible. The prognostic factors we have identified have also been suggested in other studies: increasing age,^{39 41 42} male sex,³⁸ TIA of the brain compared to TIA of the eye (amaurosis fugax) only,^{24 31 36} multiple TIAs,³⁷ multiple TIAs involving both the carotid and vertebral territories at presentation,³⁷ and evidence of peripheral vascular disease,^{37 42} ischaemic heart disease^{37 39 42} and residual neurological signs.²² Left ventricular hypertrophy has been identified as an adverse prognostic factor for stroke and other vascular events in asymptomatic people^{53 54} and also in the TIA patients in the Dutch TIA trial⁵⁷ (P Koudstaal, personal communication). The degree of diameter stenosis of the origin of the internal carotid artery on the symptomatic side was also a strong prognostic factor but it was not statistically significant (at the 0.05 level) when added to the model; in other words, much of the prognostic information of the angiographic findings for the three major outcome events can be obtained from the clinical features already in the model. Of course, it would have been interesting to

examine the predictive power of carotid stenosis for ipsilateral ischaemic stroke, particularly in view of the findings of the randomised trials of carotid endarterectomy.^{22 23}

Some readers may be surprised that the model did not select effective treatments, such as aspirin or carotid endarterectomy, as important prognostic factors. In the first instance, however, treatment is really not a good variable to examine because the decision to treat a patient or not is often based on the clinician's assessment of who is at high risk and who is not. Consequently, treatment may be seen as a marker of risk (a possible confounding factor) and not be selected by the model. Also, aspirin and carotid endarterectomy are known to have only modest overall effects on outcome and are likely to be overshadowed by stronger influences on outcome, such as age.

As with any measurement instrument in medicine, prediction equations will not always be correct; the 95% confidence intervals of the hazard ratio estimates have been included in table 4 to give some idea of the likely range of error arising from imperfections in the system itself.

Simplicity of the model

The prediction equations produced are quite simple to use because they contain only a few key variables. From a knowledge of the prevalence and level of the adverse prognostic factors possessed by any given patient, it is possible to estimate the relative risk of an outcome event by simply multiplying the hazard ratios of the relevant factors, and the absolute risk by exponentiating (with the aid of a pocket calculator that can perform exponential functions) the sum of beta coefficients of the relevant factors, as shown in tables 5c and 5d.

Implications of the clinical prediction system

The aim and implications of the results of the clinical prediction models are that they are likely to influence the management of TIA patients; patients deemed at high risk are likely to be more suitable candidates for higher (as well as lower) risk investigations and treatments and patients considered low risk should be candidates for only lower risk interventions. Also, it is possible that the identification of any significant adverse prognostic factors that are modifiable and might be causal should enable further treatment to be targeted to reducing the presence and level of the factor in these high risk patients.

Although all of the prognostic factors selected by our models are merely markers of high risk (for example, age, increasing number of TIAs) and none are treatable factors (such as high blood pressure), this does not mean that treatable factors (such as high blood pressure) should not be treated merely because they are not selected by the model. The model should be used to determine who is at high risk and who needs treatment, not which treatment to use.

Also, it cannot be assumed that the factors which predict risk of stroke in asymptomatic

populations are the same as those which predict recurrent stroke. This is because exposure to prognostic factors for stroke is likely to be far more heterogeneous in stroke-free individuals than in people who have already suffered a stroke. A hypothetical example may be that high blood pressure in asymptomatic people is a strong prognostic factor for stroke but is not selected as a prognostic factor for recurrent stroke because all stroke patients have high blood pressure. Instead, factors not common to all stroke patients but associated with recurrent stroke, are likely to be selected out.

Evaluation of clinical prediction systems

Before accepting and adopting a prediction system, clinicians must evaluate its applicability to their patients. If our prediction equations are used prematurely and are not externally valid then patients are likely to suffer.⁵⁵ We are planning to test the model in the UK-TIA aspirin trial,⁵⁶ Oxfordshire Community Stroke Project,²⁴ Dutch TIA trial⁵⁷ and European Carotid Surgery Trial²² cohorts of TIA patients. Whilst awaiting validation in other populations of TIA patients, the proposed prediction systems should initially be regarded as "just another test" to be weighed against more conventional pieces of clinical information. Later, when the prediction system has been validated in other data sets and it pronounces an individual patient is at "high risk" of stroke and other serious vascular events, the implications are that an attempt should be made to reduce the reversible elements of that risk and to consider potentially effective therapies, some of which may also carry an associated intrinsic risk. The only patients in whom potentially effective therapies are likely to be excluded on the basis of the results of the prediction equation should be those who are deemed to be at low risk of serious vascular events and they should only be denied high risk treatments.

As any decision making in patient management should involve the patient, it follows that the doctor should share with the patient any additional information derived from the prediction model. The model should be used with clinical commonsense; it is no substitute for an accurate history and physical examination which, in fact, provide most of the variables in the model.

Dr Hankey was supported by a grant from the Chest, Heart and Stroke Association. Mr Slattery was supported by a grant from the Medical Research Council. We thank Professor JD Miller, Professor J van Gijn, Dr PAG Sandercock and Dr MS Dennis for their comments on the manuscript.

- Martin J, Meltzer H, Elliot D. Office of population censuses and surveys (social survey division). Report 1. *The prevalence of disability among adults*. London: HMSO, 1988.
- Harris AI. *Handicapped and impaired in Great Britain*. Office of population censuses and surveys (social survey division) London: HMSO, 1971.
- Schoenberg BS, Schoenberg DG, Pritchard DA, Liliendfeld AM, Whisnant JP. Proportion of ischaemic stroke attributable to various forms of cardiovascular disease, diabetes mellitus and transient ischaemic attacks. *J Neurol* 1985;232 (suppl):98.
- Wolf PA, Kannel WB, McGee DL. Prevention of ischaemic stroke: Risk factors. In: Barnett HJM, Stein BM, Mohr JP, Yatsu FM, eds. *Stroke: pathophysiology, diagnosis and*

- management. Edinburgh: Churchill Livingstone, 1987; 967-88.
- 5 Sandercock PAG, Warlow CP, Jones LN, Starkey IR. Predisposing factors for cerebral infarction: the Oxfordshire community stroke project. *BMJ* 1989;298:75-80.
 - 6 Tuomilehto J, Bonita R, Stewart A, Nissinen A, Salonen JT. Hypertension, cigarette smoking, and the decline in stroke incidence in Eastern Finland. *Stroke* 1991;22:7-11.
 - 7 Shaper AG, Phillips AN, Pocock SJ, Walker M, Macfarlane PW. Risk factors for stroke in middle aged British men. *BMJ* 1991;302: 1111-5.
 - 8 Bonita R, Scragg R, Stewart A, Jackson R, Beaglehole R. Cigarette smoking and risk of premature stroke in men and women. *BMJ* 1986;293:6-8.
 - 9 Colditz GA, Bonita R, Stampfer MJ, et al. Cigarette smoking and risk of stroke in middle-aged women. *N Engl J Med* 1988;318:937-41.
 - 10 Wolf PA, D'Agostino RB, Kannel WB, Bonita R, Belanger AJ. Cigarette smoking as a risk factor for stroke: the Framingham Study. *JAMA* 1988;259:1025-9.
 - 11 Shinton R, Beevers G. Meta-analysis of relation between cigarette smoking and stroke. *BMJ* 1989;298:789-94.
 - 12 Donnan GA, McNeil JJ, Adena MA, Doyle AE, O'Malley HMO, Neill GC. Smoking as a risk factor for cerebral ischaemia. *Lancet* 1989;ii:643-7.
 - 13 Rose G, Day S. The population mean predicts the number of deviant individuals. *BMJ* 1990;301:1031-4.
 - 14 MacMahon S, Peto R, Cutler J, et al. Blood pressure, stroke, and coronary heart disease. Part 1, prolonged differences in blood pressure: prospective observational studies corrected for the regression dilution bias. *Lancet* 1990;335:765-74.
 - 15 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.
 - 16 Dennis M, Bamford J, Sandercock P, Warlow C. Incidence of transient ischaemic attacks in Oxfordshire, England. *Stroke* 1989;20:333-9.
 - 17 Dunbabin DW, Sandercock PAG. Preventing stroke by the modification of risk factors. *Stroke* 1990;21 (suppl IV): IV-36-IV-39.
 - 18 Rose G, Hamilton PJS, Colwell L, Shipley MJ. A randomised controlled trial of anti-smoking advice: 10 year results. *J Epidemiol Community Health* 1982;36:102-8.
 - 19 Peto R, Yusuf S, Collins R. Cholesterol lowering trial results in their epidemiological context (abstract). *Circulation* 1985;72 (suppl III) III-451.
 - 20 Antiplatelet Trialists' Collaboration. Secondary prevention of vascular disease by prolonged antiplatelet treatment. *BMJ* 1988;296:320-31.
 - 21 Warlow CP. Ticlopidine, a new antithrombotic drug: but is it better than aspirin for long term use? *J Neurol Neurosurg Psychiatry* 1990;53:185-7.
 - 22 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.
 - 23 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.
 - 24 Dennis M, Bamford J, Sandercock P, Warlow C. Prognosis of transient ischaemic attacks in the Oxfordshire Community Stroke Project. *Stroke* 1990;21:848-53.
 - 25 Hankey GJ, Slattery JM, Warlow CP. The prognosis of hospital-referred transient ischaemic attacks. *J Neurol Neurosurg Psychiatry* 1991;54:793-802.
 - 26 Heyman A, Wilkinson WE, Hurwitz BJ, et al. Risk of ischaemic heart disease in patients with TIA. *Neurology* 1984;34:626-30.
 - 27 Marshall J. The natural history of transient ischaemic cerebrovascular attacks. *Quart J Med* 1964;33:309-24.
 - 28 Baker RN, Ramseyer JC, Schwartz WS. Prognosis in patients with transient cerebral ischaemic attacks. *Neurology* 1968;18:1157-65.
 - 29 Acheson J. Factors affecting the natural history of "focal cerebral vascular disease". *Q J Med* 1971;40:25-46.
 - 30 Muuronen A, Kasté M. Outcome of 314 patients with transient ischaemic attacks. *Stroke* 1982;13:24-31.
 - 31 Poole CJM, Ross Russell RW. Mortality and stroke after amaurosis fugax. *J Neurol Neurosurg Psychiatry* 1985;48:902-5.
 - 32 Whisnant JP, Cartledge NEF, Elveback LR. Carotid and vertebral-basilar transient ischaemic attacks: effect of anticoagulants, hypertension and cardiac disorders on survival and stroke occurrence—a population study. *Ann Neurol* 1978;3:107-15.
 - 33 Whisnant JP, Wiebers DO. Clinical epidemiology of transient cerebral ischaemic attacks (TIA) in the anterior and posterior cerebral circulation. In: Sundt TM Jr, ed. *Occlusive cerebrovascular disease, diagnosis and surgical management*. Philadelphia: Saunders, 1987:60-5.
 - 34 Conneally PM, Dyken ML, Fuddy DE, et al. Cooperative study of hospital frequency and character of transient ischaemic attacks. VIII. Risk factors. *JAMA* 1978;240:742-6.
 - 35 Wilkinson WE, Heyman A, Pfeffer RI, Burch JG. A questionnaire for TIA symptoms: a predictor of subsequent stroke. In: M Reivich, HI Hurvig, eds. *Cerebrovascular disease*. New York: Raven Press, 1983: 77-92.
 - 36 Hurwitz BJ, Heyman A, Wilkinson WE, Haynes CS, Udey CM. Comparison of amaurosis fugax and transient cerebral ischaemia: a prospective clinical and arteriographic study. *Ann Neurol* 1985;18:698-704.
 - 37 Candelise L, Vigotti M, Fieschi C, et al. Italian Multicenter Study on reversible cerebral ischaemic attacks: VI—Prognostic factors and follow up results. *Stroke* 1986;17:842-7.
 - 38 Calandre L, Bermejo F, Balseiro J. Long-term outcome of TIAs, RINDs and infarctions with minimum residuum. A prospective study in Madrid. *Acta Neurol Scand* 1990;82:104-8.
 - 39 Howard G, Toole JF, Frye-Pierson J, Hinshelwood LC. Factors influencing the survival of 451 transient ischaemic attack patients. *Stroke* 1987;18:552-7.
 - 40 Evans GW, Howard G, Murros KE, Rose LA, Toole JF. Cerebral infarction verified by cranial computed tomography and prognosis for survival following transient ischaemic attack. *Stroke* 1991;22:431-6.
 - 41 Kernan WN, Horwitz RJ, Brass LM, Viscoli CM, Taylor KJW. A prognostic system for transient ischaemic attack or minor stroke. *Ann Intern Med* 1991;114:552-7.
 - 42 American-Canadian Co-operative Study Group: Persantin aspirin trial in cerebral ischaemia. Part III: Risk factors for stroke. *Stroke* 1986;17:12-8.
 - 43 Simonsen N, Christiansen HD, Heltberg A, Marquardsen J, Pedersen HE, Sorenson PS. Long-term prognosis after transient ischaemic attacks. *Acta Neurol Scand* 1981;63:156-68.
 - 44 Cox DR. Regression Models and life-tables. *J R Stat Soc* 1972;B34:187-220.
 - 45 Cox DR, Oakes D. *Analysis of survival data*. London: Chapman and Hall, 1984.
 - 46 Hopkins A. P2L. Survival analysis with covariates-Cox models. In: Dixon WJ, Brown MB, Engelman L, et al, eds. *Biomedical data program (BMDP) statistical software*. Berkeley, California: University of California Press, 1985;576-94.
 - 47 Tibshirani R. A plain man's guide to the proportional hazards model. *Clinical and Investigative medicine* 1982;5:63-8.
 - 48 Morris JA, Gardner M. Calculating confidence intervals for relative risks, odds ratios, and standardised ratios and rates. In: Gardner MJ, Altman DG, eds: *Statistics with confidence*. Belfast: Universities Press, *BMJ*, 1989: 50-63.
 - 49 Frane J. PAM Description and Estimation of missing data. In: Dixon WJ, Brown MB, Engelman L, et al, eds. *BMDP Statistical software*. Berkeley, California: University of California Press, 1985;217-34.
 - 50 Epidemiological Resources. *EGRET: epidemiological graphics, estimation and testing*. Seattle, Washington: Statistics and epidemiology research corporation, 1988.
 - 51 Sackett DL, Haynes RB, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. Boston: Little, Brown, 1985.
 - 52 Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules: applications and methodological standards. *N Engl J Med* 1985;313:793-9.
 - 53 Levy D, Garrison RJ, Savage DD, Kannel WB, Castelli WP. Prognostic implications of echocardiographically determined left ventricular mass in the Framingham Heart Study. *N Engl J Med* 1990;322:1561-6.
 - 54 Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: a risk profile from the Framingham Study. *Stroke* 1991;22:312-8.
 - 55 Seymour DG, Green M, Vaz FG, Coles EC. Risk prediction in medicine and surgery: ethical and practical considerations. *J Roy Coll Physicians* 1990;24:173-7.
 - 56 UK-TIA Study Group. The United Kingdom transient ischaemic attack (UKTIA) aspirin trial: final results. *J Neurol Neurosurg Psychiatry* 1991;54:1044-54.
 - 57 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.