Prediction of 60 day case-fatality after aneurysmal subarachnoid haemorrhage: results from the International Subarachnoid Aneurysm Trial (ISAT)

R. Risselada · H. F. Lingsma · A. Bauer-Mehren · C. M. Friedrich ·
A. J. Molyneux · R. S. C. Kerr · J. Yarnold · M. Sneade ·
E. W. Steyerberg · M. C. J. M. Sturkenboom

Abstract Aneurysmal subarachnoid haemorrhage (aSAH) is a devastating event with substantial case-fatality. Our purpose was to examine which clinical and neuro-imaging characteristics, available on admission, predict 60 day case-fatality in aSAH and to evaluate performance of our prediction model. We performed a secondary analysis of patients enrolled in the International Subarachnoid Aneurysm Trial (ISAT), a randomised multicentre trial to compare coiling with clipping in aSAH patients. Multivariable logistic regression analysis was used to develop a prognostic model to estimate the risk of dying within 60 days from aSAH based on clinical and neuro-imaging characteristics. The model was internally validated with bootstrapping techniques. The study population comprised of 2,128 patients who had been randomised to either endovascular coiling or neurosurgical clipping. In this population 153 patients (7.2%) died within 60 days. World Federation of Neurosurgical Societies (WFNS) grade was the most important predictor of case-fatality, followed by age, lumen size of the aneurysm and Fisher grade. The model discriminated reasonably between those who died within 60 days and those who survived (c statistic = 0.73), with minor optimism according to bootstrap re-sampling (optimism corrected c statistic = 0.70). Several strong predictors are available to predict 60 day case-fatality in aSAH patients who survived the early stage up till a treatment decision; after external validation these predictors could eventually be used in clinical decision making.

Keywords Aneurysm · Subarachnoid haemorrhage · Prediction model · Case-fatality

R. Risselada (✉) · M. C. J. M. Sturkenboom
Department of Medical Informatics, Erasmus MC,
Room Ee 2159, P.O. Box 2040, 3000 CA Rotterdam,
The Netherlands
e-mail: r.risselada@erasmusmc.nl

H. F. Lingsma · E. W. Steyerberg
Department of Public Health, Erasmus MC,
Rotterdam, The Netherlands

A. Bauer-Mehren
Research Unit on Biomedical Informatics (GRIB),
IMIM-Hospital del Mar, Universitat Pompeu Fabra,
Barcelona, Spain

C. M. Friedrich
Department of Bioinformatics, Fraunhofer Institute Algorithms
and Scientific Computing (SCAI), Sankt Augustin, Germany

A. J. Molyneux · J. Yarnold · M. Sneade
Oxford Neurovascular & Neuroradiology Research Unit,
University of Oxford, Oxford, UK

R. S. C. Kerr
Department of Neurosurgery, West Wing,
John Radcliffe Hospital, Oxford, UK

M. C. J. M. Sturkenboom
Department of Epidemiology, Erasmus MC,
Rotterdam, The Netherlands

Abbreviations

ACA	Anterior Cerebral Artery
AIC	Akaike's Information Criterion
aSAH	Aneurysmal subarachnoid haemorrhage
CI	Confidence interval
CT	Computed Tomography
GCS	Glasgow Coma Scale
ICA	Internal Carotid Artery
IQR	Inter quartile range
ISAT	International Subarachnoid Aneurysm Trial

ISRCTN	International Standard Randomised Controlled Trial Number
LP	Linear predictor
MCA	Middle Cerebral Artery
OR	Odds ratio
PC	Posterior Circulation
RCT	Randomised clinical trial
SAH	Subarachnoid haemorrhage
SE	Standard error
WFNS	World Federation of Neurosurgical Societies

Introduction

Subarachnoid haemorrhage (SAH) is a devastating event, which is marked by sudden onset of severe headache, causing substantial case-fatality. In 85% of the patients, the SAH is caused by rupture of an aneurysm (aSAH) [1, 2]. From those who survive the first month, approximately one-third remains dependent with respect to daily activities during their remaining lifetime [1]. Amongst patients who regain independency, quality of life remains reduced [3].

Early prediction of short term outcome in terms of case-fatality may support clinical decision making and may provide realistic and evidence based expectations to patients and relatives. Predictions may also be used to classify patients according to prognostic risk, which may be useful to compare outcome between different patient series, to study treatment results over time, or to stratify patients for randomised clinical trials (RCT).

Several other prognostic studies on outcome after aSAH have been performed, but most included relatively small numbers. Two included substantial numbers allowing analysis of the effects of multiple independent factors [4, 5]. However, these studies did not examine case-fatality, but arbitrarily dichotomized ordinal outcome scales (modified Rankin Scale or Glasgow Outcome Scale). Moreover, due to their design, these studies were unable to examine the effect of treatment on prediction of outcome.

Our aim was to develop a prognostic model for 60 day case-fatality, based on clinical features and neuro-imaging, regularly readily available on admission to a neurological or neurosurgical unit. These data were obtained from a large RCT conducted in mainly European countries.

Methods

Patients

Data were collected prospectively by the Medical Research Council funded International Subarachnoid Aneurysm

Trial (ISAT) (International Standard Randomised Controlled Trial, number ISRCTN49866681). Full details of the ISAT study are available elsewhere [6]. The aim of the trial was to determine whether treatment using endovascular coiling reduced the risk of patients being dependent or dead at 1 year by 25 percent when compared with neurosurgical treatment (clipping).

Predictors and outcome

We considered all patient characteristics that can be collected easily and reliably within the first hours after hospital admission and that were also present in the ISAT database. These included age, gender, previous occurrence of SAH, CT scan Fisher grading, lumbar puncture, World Federation of Neurosurgical Societies (WFNS) grading, number of intracranial aneurysms, location of the aneurysm, maximum lumen size of the aneurysm, vasospasm on angiography, and intended treatment by randomization. Fisher grading of blood visible on a plain CT scan runs from grade 1 ("no blood visible") up to grade 4 ("intraventricular or intraparenchymal blood"). Lumbar puncture was not performed in all participants. If it was performed it was graded 1 ("no blood in cerebrospinal fluid") or 2 ("xanthochromia or blood"); 0 otherwise ("no lumbar puncture"). WFNS scale runs from grade 1 ("Glasgow Coma Scale (GCS) 15 and no motor deficit") to grade 5 ("GCS 3-6 with or without motor deficit"). One category additional to the standard WFNS scale was created in ISAT for those in whom WFNS could not be assessed; 'grade 6'. The number of aneurysms was categorized in 1, 2, and 3 or more intracranial aneurysms. We discerned four aneurysm locations: Anterior Cerebral Artery (ACA), Internal Carotid Artery (ICA), Middle Cerebral Artery (MCA), and Posterior Circulation (PC). The maximum lumen size of the aneurysm was expressed in millimetres. Vasospasm was examined on angiography and categorized as 'none', 'mild', 'moderate', or 'severe'. Treatment was either neurosurgical clipping or endovascular coiling; we used treatment as allocated by the randomization procedure. We developed the model based on cases with a complete set of data. The outcome was 60 day case-fatality.

Model

We used univariate logistic regression analysis to estimate the association between single predictors and outcome, expressed as an odds ratio (OR). Predictors have a statistically significant effect if the 95 percent confidence interval (95% CI) does not include the value one. The prediction model was developed with multivariable logistic regression with backward stepwise selection. All potential predictors were entered into the model and those that met

Akaike's Information Criterion (AIC) were selected into the model. AIC compares models based on how well they fit the data, but penalizes for the complexity of the model. AIC requires that the increase in model χ^2 when entering a new predictor has to be larger than two times the degrees of freedom: $\chi^2 > 2\ df$. When considering a predictor with one df, such as gender, this implies that χ^2 has to exceed two, equivalent to $P < 0.157$. When considering a predictor with two df, $\chi^2 > 4$, or $P < 0.135$; and in case of four df, $\chi^2 > 8$, or $P < 0.092$ [7].

Performance

The performance of the model was assessed with respect to calibration and discrimination. Calibration is the ability of the model to produce unbiased estimates of the probability of the outcome. Calibration was examined with a goodness of fit test, which assesses agreement between predicted and observed risks over the full range of predicted probabilities [8].

Discrimination is the model's ability to separate patients with different outcomes. To quantify the discrimination, we used the concordance (c) statistic. For binary outcomes, c is identical to the area under the receiver operating characteristic curve [7]. The c statistic evaluates whether those with higher predicted risk are more likely to die within 60 days among all possible pairs of patients with different outcomes. A model with a c statistic of 0.5 has no discriminative power at all, for example a coin flip. A c statistic of 1.0 reflects perfect discrimination.

Model validation

The performance of a prediction model is generally worse in new patients than initially expected. This 'optimism' can be studied with internal validation techniques [7]. Internal validity of our model was assessed with standard bootstrapping procedures [7]. Bootstrapping involves drawing samples of patients with replacement from the study population. Each sample can be considered as if one is repeating the data collection with the same number of patients and under identical circumstances as the original. The multivariable logistic regression coefficients were re-estimated in 300 bootstrap samples. Each of these 300 models was evaluated on the original sample. The average difference in the c statistic was determined to indicate the optimism in the initially estimated discriminative ability [7]. A shrinkage factor was estimated from the bootstrap validation procedure and we shrunk the regression coefficients to provide better predictions for future patients [7].

All statistical analyses were performed using R software, version 2.8.1 (R foundation for statistical computing, Vienna, Austria).

Results

A total of 2,143 patients were recruited into the ISAT trial by 43 neurosurgical centres, mainly in Europe. CT scans of 14 patients were not performed or available, and in one patient no information on vasospasm was available. We excluded these cases from our analysis. Data on the outcome were available for all patients. Thus, we performed complete case analysis on 2,128 patients (99.3%) of whom 153 (7.2%) died within 60 days.

The distribution of patient characteristics of the study population is presented in Table 1. For reasons of small numbers in the 'severe' category of vasospasm, we aggregated data from the 'moderate' and 'severe' categories into one category. Univariate analysis showed a statistically significant relation with 60 day case-fatality for age, lumen size, Fisher grade, lumbar puncture, WFNS grade, and vasospasm. Sex, location and number of aneurysms and intended treatment were not significantly associated with 60 day case-fatality (Table 1). In the multivariable model with stepwise backward selection age, lumen size, Fisher grade, and WFNS grade met AIC and were included in the final model. In Table 2 the chi square statistics with corresponding P-values are presented as well as the point estimate of the OR. Age and WFNS grade were the most important predictors.

The goodness of fit test yielded a P-value of 0.86, suggesting that the model fitted the data in which it was developed well. The c statistic of the original model was 0.73, meaning that the model discriminates reasonably between patients who die within 60 days from onset of the SAH and those who survive this period.

Validation by means of 300 bootstrap samples resulted in a shrinkage factor of 0.85, which was applied to the betas of the model. The c statistic of the internally validated model was 0.70. Details of the final prognostic model for 60 day case-fatality are described in the appendix.

Discussion

We developed a prognostic model to predict the risk of 60 day case-fatality in individual patients after aSAH. Predictions were based on characteristics that are regularly readily available on admission to a neurological or neurosurgical unit and which were collected in a large clinical trial. The full model yielded a c statistic of 0.73.

Previously, several models to estimate the probability of unfavourable outcome after aSAH have been developed. Our model was similar to those; we included roughly the same predictors: age, clinical status, and lumen size [4, 5, 9]. However, our study is of added value because of the substantial size and the inclusion of both clipped and coiled

Table 1 Population characteristics and univariate association with 60 day case fatality

Predictor	Alive $n = 1{,}975$		Death $n = 153$						
	Median	IQR	Median	IQR	Beta	SE	OR	CI$_{min}$	CI$_{max}$
Age [10 years]	5.2	4.3–5.9	5.6	5.0–6.5	0.36	0.08	1.43	1.23	1.66
Lumensize [mm]	5	4–7	6	4–8	0.10	0.02	1.10	1.05	1.15
	n	%	n	%					
Sex									
Female	1,236	63	103	67			1		
Male	739	37	50	33	−0.21	0.18	0.81	0.57	1.15
Previous SAH									
Yes	129	7	6	4	−0.54	0.43	0.58	0.25	1.35
No	1,846	93	147	96			1		
Fisher grade									
1	112	6	2	1			1		
2	350	18	10	7	0.47	0.78	1.60	0.35	7.41
3	840	43	62	41	1.42	0.73	4.13	1.00	17.13
4	673	34	79	52	1.88	0.72	6.57	1.59	27.13
Lumbar puncture									
Xanthochromia or blood	217	11	7	5	−0.95	0.39	0.39	0.18	0.84
No blood	5	0	0	0	−5.19	20.8	0.01	0.00	∞
No puncture	1,753	89	146	95			1		
WFNS grade									
1	1,270	64	54	35			1		
2	495	25	51	33	0.89	0.20	2.42	1.63	3.60
3	120	6	13	8	0.94	0.32	2.55	1.35	4.80
4	55	3	19	12	2.09	0.30	8.12	4.51	14.63
5	13	1	7	5	2.54	0.49	12.66	4.86	33.02
6 (not assessable)	22	1	9	6	2.26	0.42	9.62	4.23	21.89
n of aneurysms detected									
1	1,555	79	116	76			1		
2	314	16	29	19	0.21	0.22	1.24	0.81	1.89
≥3	106	5	8	5	0.10	0.38	1.10	0.52	2.32
Location									
ACA	1,008	51	71	46	−0.16	0.19	0.85	0.59	1.23
ICA	638	32	53	35			1		
MCA	277	14	23	15	0.00	0.26	1.00	0.60	1.66
PC	52	3	6	4	0.33	0.45	1.39	0.57	3.38
Vasospasm									
None	1,575	80	109	71			1		
Mild	218	11	24	16	0.46	0.24	1.59	1.00	2.53
Moderate/severe	182	9	20	13	0.77	0.26	2.15	1.30	3.55
Intended treatment									
Clip	983	50	83	54			1		
Coil	992	50	70	46	−0.18	0.17	0.84	0.60	1.16

IQR, inter quartile range; *beta*, regression coefficient in the logistic regression model; *SE*, standard error; *OR*, odds ratio; *CI$_{min}$*, lower limit of the 95% confidence interval; *CI$_{max}$*, upper limit of the 95% confidence interval

patients. The studies by Hoh et al. [4] ($n = 515$) and Mocco et al. [9] ($n = 148$) contained relatively few patients. The small numbers of coiled patients (79 and 35, respectively) and the design of the study did not allow for taking the effect of treatment in consideration. The study of Rosengart et al. [5] ($n = 3{,}667$) was not able to do that

Table 2 Statistical parameters of the final model

Factor	X^2	df	P-value	OR	95% CI
WFNS grade	51	5	8×10^{-10}	Grade 1 = ref.	
2				1.87	1.23–2.83
3				1.70	0.87–3.32
4				4.87	2.60–9.14
5				7.00	2.54–19.28
6				5.75	2.41–13.73
Age [10 years]	17	1	4×10^{-5}	1.32	1.13–1.55
Lumen size [mm]	12	1	4×10^{-4}	1.08	1.03–1.13
Fisher grade	8	3	0.04	Grade 1 = ref.	
2				1.43	0.27–7.65
3				2.67	0.53–13.51
4				2.76	0.54–14.14

X^2 is the chi squared test statistic for the predictor in the final model; df = degrees of freedom

95% CI was calculated based on the S.E. of the estimates of the coefficients in the full model to avoid underestimation of uncertainty

either, since patients treated with Guglielmi or other detachable coils were excluded. All three studies used a dichotomized ordinal scale as an outcome, for which the cut off can be (arbitrarily) chosen in different studies. In a sense, examining case-fatality is also a dichotomization of an ordinal scale, though less arbitrary. Therefore, we are convinced that logistic regression is well suited for an outcome that is by its nature dichotomous, whereas for an ordinal outcome we would prefer specific modelling techniques.

Several limitations of this study should be acknowledged. This study used data from one large trial on a selected population of patients who survived the early stage up till a treatment decision and who were in equipoise regarding that decision on treatment with either endovascular coiling or neurosurgical clipping, which may limit external validity. The model may perform well in this development sample, but worse when applied to other groups of patients, for example, a less strictly selected population. Nonetheless, according to a recently published paper, the ISAT population proved to be quite similar to the population admitted with an aSAH to neurosurgical units in the United Kingdom [10]. Although in ISAT, a lower proportion of poor grade patients were enrolled. Validation of a prognostic model in independent patient series is therefore considered an essential next step [11]. However, since large samples of systematically collected data on aSAH are sparse, assessment of external validity is difficult. For now the external validity of our model remains to be established. This will be a topic of future research.

Although our model represents knowledge obtained from 2,128 SAH patients in equipoise regarding treatment,

statistical models can never replace the clinician with regard to decision making; they can only assist with this task. A prediction for an individual aSAH patient a particular situation is always subject to uncertainty. The model makes certain structural assumptions and statistical interaction terms were not included. It is hence possible that specific patterns of risk factors are inadequately reflected in the model predictions. Therefore, predictions should be regarded with care and not directly be applied for treatment limiting decisions.

Although the performance of the presented model was satisfactory, it might potentially be improved by including neuro-imaging biomarkers other than lumen size, location, Fisher grade on plain CT scan, and vasospasm on angiography. Biomarkers regarding anatomy and morphology might be considered, as well as aneurysm characteristics obtained from three and four dimensional angiography [12, 13]. Performance may also be improved by inclusion of subsequent information obtained after admission, including temporal course, neuro-imaging at later time points, eventual rebleeding of the aneurysm, delayed ischemic deficit, and other parameters such as hydrocephalus. The objective of the present study, however, was to investigate prognostic models that predict 60 day case-fatality with predictors available on admission.

Statistical testing for calibration has a number of drawbacks. First, the null hypothesis is of good calibration. Hence, if we test calibration in a small study, we have low power and will not reject the null hypothesis unless miscalibration is very severe. On the other hand, even a model with very good, though not perfect, calibration will fail the test in case of a sufficiently large sample. Moreover, reported goodness-of-fit tests are usually non-significant if they reflect apparent validation on the data that were also used to construct the model. Such non-significant results may contribute to the face validity of a model, but have no real scientific meaning [7].

In conclusion, we presented a prognostic model for predicting 60 day case-fatality after aneurysmal SAH. Our model contained age, lumen size, Fisher grade, and WFNS grade as predictors. After calibration and internal validation, our model showed reasonable performance, although external validity of our model remains to be established.

Appendix

Details of the prognostic model

The probability of dying within 60 days is calculated according to the logistic formula: $1/(1 + \exp^{-LP})$. The linear predictor (LP) takes the form of LP = intercept + regression coefficients × predictor values.

LP for 60 day case-fatality = $-5.812 + 0.2762 \times$ age + $0.3572 \times$ [Fisher grade II] + $0.9756 \times$ [Fisher grade III] + $1.008 \times$ [Fisher grade IV] + $0.6216 \times$ [WFNS grade 2] + $0.5261 \times$ [WFNS grade 3] + $1.574 \times$ [WFNS grade 4] + $1.934 \times$ [WFNS grade 5] + $1.738 \times$ [WFNS grade not assessable] + $0.07662 \times$ lumen size of aneurysm.

Coding of the predictors was as follows: age in decades, lumen size in millimetres; all other predictors, 1 if true and 0 if false.

References

1. Hop JW, Rinkel GJ, Algra A, van Gijn J. Case-fatality rates and functional outcome after subarachnoid hemorrhage: a systematic review. Stroke. 1997;28:660–4.
2. Van Gijn J, Kerr RS, Rinkel GJ. Subarachnoid haemorrhage. Lancet. 2007;369:306–18.
3. Hop JW, Rinkel GJ, Algra A, van Gijn J. Changes in functional outcome and quality of life in patients and caregivers after aneurysmal subarachnoid hemorrhage. J Neurosurg. 2001;95:957–63.
4. Hoh BL, Topcuoglu MA, Singhal AB, Pryor JC, Rabinov JD, Rordorf GA, Carter BS, Ogilvy CS. Effect of clipping, craniotomy, or intravascular coiling on cerebral vasospasm and patient outcome after aneurysmal subarachnoid hemorrhage. Neurosurgery. 2004;55:779–786; discussion 786–779.
5. Rosengart AJ, Schultheiss KE, Tolentino J, Macdonald RL. Prognostic factors for outcome in patients with aneurysmal subarachnoid hemorrhage. Stroke. 2007;38:2315–21.
6. Molyneux A, Kerr R, Stratton I, Sandercock P, Clarke M, Shrimpton J, et al. International Subarachnoid Aneurysm Trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised trial. Lancet. 2002;360:1267–74.
7. Steyerberg EW. Clinical prediction models. A practical approach to development, validation, and updating. New York: Springer; 2009.
8. Harrell FE Jr. Resampling model calibration; the calibrate function in library(Design) in R: a language and environment for statistical computing. Vienna (Austria): R foundation for statistical computing; 2006. Available: http://www.R-project.org/. Accessed 12 Jan 2010.
9. Mocco J, Ransom ER, Komotar RJ, Schmidt JM, Sciacca RR, Mayer SA, Connolly ES, Jr. Preoperative prediction of long-term outcome in poor-grade aneurysmal subarachnoid hemorrhage. Neurosurgery. 2006;59:529–538; discussion 529–538.
10. Langham J, Reeves BC, Lindsay KW, van der Meulen JH, Kirkpatrick PJ, Gholkar AR, et al. Variation in outcome after subarachnoid hemorrhage: a study of neurosurgical units in uk and ireland. Stroke. 2009;40:111–8.
11. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann Intern Med. 1999;130:515–24.
12. Hayakawa M, Katada K, Anno H, Imizu S, Hayashi J, Irie K, et al. Ct angiography with electrocardiographically gated reconstruction for visualizing pulsation of intracranial aneurysms: identification of aneurysmal protuberance presumably associated with wall thinning. AJNR Am J Neuroradiol. 2005;26:1366–9.
13. Zhang C, Villa-Uriol MC, De Craene M, Pozo J, Frangi A. Morphodynamic analysis of cerebral aneurysm pulsation from time-resolved rotational angiography. IEEE Trans Med Imaging. 2009;28:1105–16.