

ASTRAL, DRAGON and SEDAN scores predict stroke outcome more accurately than physicians

G. Ntaios^a, F. Gioulekas^b, V. Papavasileiou^c, D. Strbian^d and P. Michel^e

^aDepartment of Medicine, University of Thessaly, Larissa; ^bSub-directorate of Informatics, Larissa General University Hospital, Larissa, Greece; ^cDepartment of Neurosciences, Stroke Service, Leeds Teaching Hospitals NHS Trust and School of Medicine, University of Leeds, Leeds, UK; ^dDepartment of Neurology, Helsinki University Central Hospital, Helsinki, Finland; and ^eStroke Centre, Neurology Service, Centre Hospitalier Universitaire Vaudois and University of Lausanne, Lausanne, Switzerland

Keywords:

ASTRAL, DRAGON, SEDAN

Received 7 February 2016

Accepted 9 June 2016

European Journal of Neurology 2016, **0**: 1–7

doi:10.1111/ene.13100

Background and purpose: ASTRAL, SEDAN and DRAGON scores are three well-validated scores for stroke outcome prediction. Whether these scores predict stroke outcome more accurately compared with physicians interested in stroke was investigated.

Methods: Physicians interested in stroke were invited to an online anonymous survey to provide outcome estimates in randomly allocated structured scenarios of recent real-life stroke patients. Their estimates were compared to scores' predictions in the same scenarios. An estimate was considered accurate if it was within 95% confidence intervals of actual outcome.

Results: In all, 244 participants from 32 different countries responded assessing 720 real scenarios and 2636 outcomes. The majority of physicians' estimates were inaccurate (1422/2636, 53.9%). 400 (56.8%) of physicians' estimates about the percentage probability of 3-month modified Rankin score (mRS) > 2 were accurate compared with 609 (86.5%) of ASTRAL score estimates ($P < 0.0001$). 394 (61.2%) of physicians' estimates about the percentage probability of post-thrombolysis symptomatic intracranial haemorrhage were accurate compared with 583 (90.5%) of SEDAN score estimates ($P < 0.0001$). 160 (24.8%) of physicians' estimates about post-thrombolysis 3-month percentage probability of mRS 0–2 were accurate compared with 240 (37.3%) DRAGON score estimates ($P < 0.0001$). 260 (40.4%) of physicians' estimates about the percentage probability of post-thrombolysis mRS 5–6 were accurate compared with 518 (80.4%) DRAGON score estimates ($P < 0.0001$).

Conclusions: ASTRAL, DRAGON and SEDAN scores predict outcome of acute ischaemic stroke patients with higher accuracy compared to physicians interested in stroke.

Introduction

Several prognostic scores were recently introduced for outcome prediction in acute ischaemic stroke patients, which could be helpful in the clinical and research setting [1]. Amongst others, the ASTRAL score predicts functional outcome in the general ischaemic stroke population [2], and the DRAGON and SEDAN scores

predict functional outcome and symptomatic intracranial haemorrhage (sICH) respectively in patients treated with intravenous thrombolysis (IVT) [3,4]. These three scores were validated in large external populations and it was shown that they can offer a reliable estimate of patient outcome, with the areas under the curve in the validation cohorts of the original publications ranging between 0.77 and 0.902 [5–14]. Their association with outcome is summarized in Data S1. Stroke prognostic scores received criticism that they aim at the non-specialist physician and that they may be less useful than a stroke specialist's prognostication.

Correspondence: G. Ntaios, Department of Medicine, University of Thessaly, Biopolis, Larissa 41110, Greece (tel.: +302413502888; fax: +302413501557; e-mail: gntaios@med.uth.gr).

G. Ntaios and F. Gioulekas are considered first co-authors.

Our aim was to investigate whether the ASTRAL, DRAGON and SEDAN scores make a more accurate prediction of stroke outcome compared with physicians.

Methods

An online anonymous survey was developed based on open-source software [15]. Invitations for participation in the survey were sent to physicians with interest in stroke medicine through (i) the European Stroke Organization and the World Stroke Organization newsletters, (ii) the stroke societies of the authors' countries and (iii) personal contacts of the authors.

The survey consisted of a random set of six real scenarios with predefined structure (Data S2). The patients' parameters presented in the scenarios were more extensive than the parameters needed to calculate the three scores, i.e. the scenarios included also information about specific patient symptoms/signs and comorbidities. The patients presented in the scenarios were retrieved from the Acute Stroke Registry and Analysis of Lausanne (ASTRAL registry) which was also used for the development of the ASTRAL score [16]. In an effort to prevent bias, the present study included only patients entered in the ASTRAL registry after the development of the ASTRAL score. The study was performed according to the ethical guidelines of the commission for research on humans of the Canton of Vaud, subcommission III.

Details about the scores have been published previously [2–8]. Each survey participant was offered a different set of six scenarios selected from the ASTRAL registry with the use of a random number generator: the first three scenarios of each set referred to the general ischaemic stroke population (i.e. regardless of whether the patient was thrombolysed) and the participant was asked to predict the percentage probability of a 3-month modified Rankin score (mRS) > 2 (this estimate was compared to the ASTRAL score's prognosis for the same scenario). The last three scenarios of each set referred to IVT patients and the participant was asked to predict (i) the percentage probability of sICH defined with the ECASS-II criteria [17] (this estimate was compared to the SEDAN score's prognosis for the same scenario), (ii) the percentage probability of 3-month mRS 0–2 (this estimate was compared to the DRAGON score's prognosis for the same scenario) and (iii) the percentage probability of 3-month mRS 5–6 (this estimate was compared to the DRAGON score's prognosis for the same scenario). Participants were explicitly asked (i) to use solely their clinical judgement and avoid any prognostic tool and

(ii) to complete a set of demographic questions focused on their own age, specialty status, hospital setting and stroke treatment experience. There was no time limit for the completion of the survey and participants were not reimbursed or offered any other incentive.

An estimate was considered accurate if it was within the 95% confidence interval (CI) of the actual outcome. Any interaction between the accuracy of the physicians' estimates with the values of the scores was also investigated. This was performed by linear regression using natural cubic splines to generate trend lines. The comparison of the estimates was performed with the χ^2 test. Categorical covariates are summarized as median and interquartile range (IQR). Statistical analysis was performed with the R v.3.2.0 [18] statistical package.

Results

Between 20 November 2014 and 1 February 2015, 244 physicians [median age 41.0 (IQR 14.0) years] from 32 countries and four continents participated in the survey. 83 (34.0%) were females; 194 (79.5%) were practising in university hospitals and 50 (20.5%) in community hospitals; 88 (36.1%) were specialty physicians for >10 years, 107 (43.8%) for ≤10 years and 49 (20.1%) were in specialization training. 123 (50.4%) were treating >10 stroke patients weekly, 88 (36.1%) between 5 and 10, and 33 (13.5%) <5 patients. The majority of participants were neurologists ($n = 193$, 79.1%), 24 (9.8%) internists and 15 (6.1%) geriatricians.

In total, 720 different real scenarios (39.4% females) were used in the survey with median age 68.8 (IQR 19.4) years and median National Institutes of Health Stroke Scale (NIHSS) 9 (IQR 19.4).

ASTRAL score versus physicians for the prediction of 3-month mRS > 2 in the general ischaemic stroke population

In all, 244 physicians responded to 704 scenarios (431 different real-life patients). 609 (86.5%) ASTRAL score estimates and 400 (56.8%) physicians' estimates were accurate ($P < 0.0001$) (Fig. 1). There was an interaction between physicians' accuracy and patients' ASTRAL score ($P < 0.001$), i.e. the degree of accuracy of the physicians was higher in patients with very low and very high ASTRAL score (Fig. 2, upper left panel). The majority of physicians' inaccurate estimates ($n = 231$, 76.0%) represented an overestimation of the probability of unfavourable outcome.

Figure 1 Probability of mRS > 2 in the general ischaemic stroke population stratified by ASTRAL score: actual outcome and 95%CI range (blue dots); physicians' outcome estimates (white circles; size corresponds to number of responses, range 1–10); ASTRAL score outcome estimates (red squares).

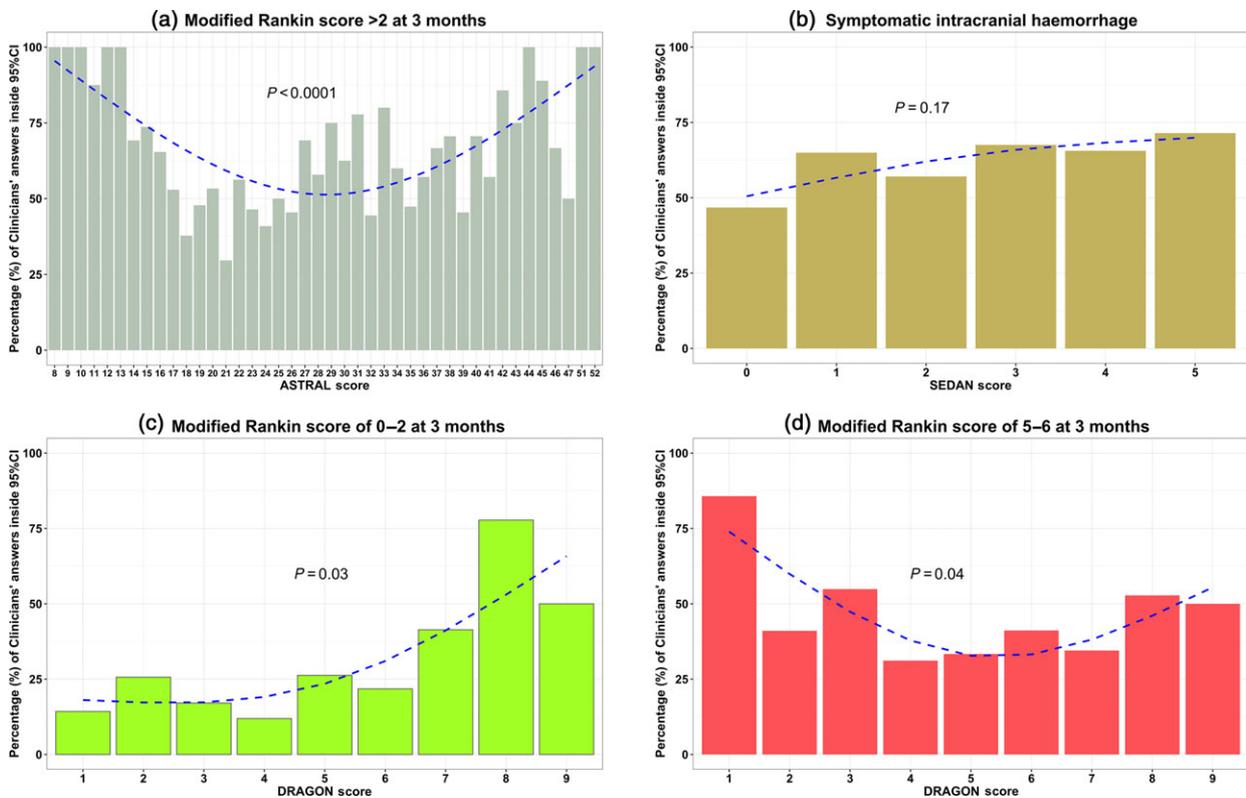
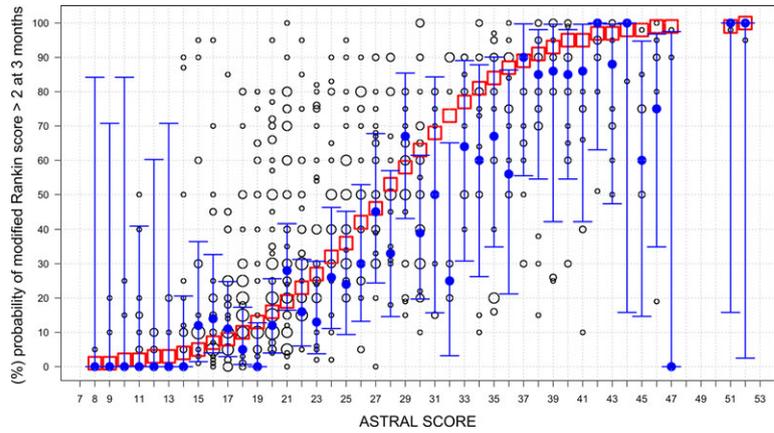


Figure 2 Proportion of clinicians' accurate estimates stratified by patients' prognostic scores. Dashed lines and *P* values correspond to trends of accuracy of physicians' estimates with the patients' score values.

SEDAN score versus physicians for the prediction of sICH in IVT patients

Two hundred and twenty physicians responded to 644 scenarios (380 different real-life IVT patients). 583 (90.5%) SEDAN score estimates and 394 (61.2%) clinicians' estimates were accurate (*P* < 0.0001) (Fig. 3). There was no interaction between physicians' accuracy and patients' SEDAN score, i.e. the degree of physicians' accuracy was constant across the whole range of the SEDAN score (Fig. 2, upper right panel).

The majority of physicians' inaccurate estimates (*n* = 219, 87.6%) represented an overestimation of the probability of sICH.

DRAGON score versus physicians for the prediction of 3-month mRS ≤ 2 in IVT patients

Two hundred and twenty physicians responded to 644 scenarios (380 different real-life IVT patients). Two hundred and forty (37.3%) DRAGON score estimates and 160 (24.8%) physicians' estimates were accurate

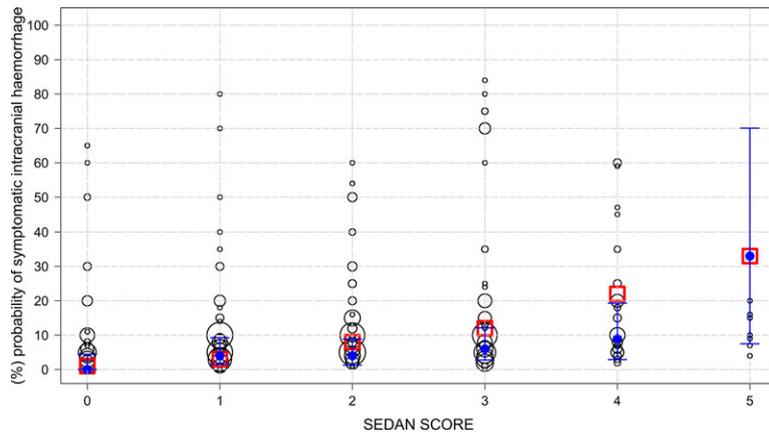


Figure 3 Probability of symptomatic intracranial haemorrhage in thrombolysed patients stratified by SEDAN score: actual outcome and 95%CI range (blue dots); physicians' outcome estimates (white circles; size corresponds to number of responses, range 1–82) and SEDAN score outcome estimates (red squares).

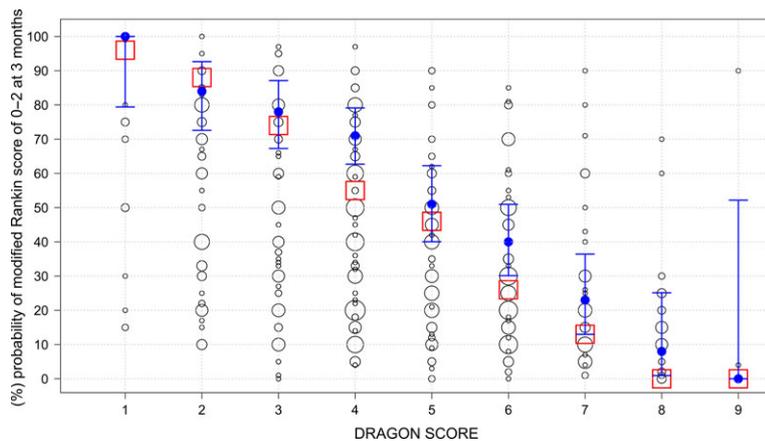


Figure 4 Probability of mRS ≤ 2 in thrombolysed patients stratified by DRAGON score: actual outcome and 95%CI range (blue dots); physicians' outcome estimates (white circles; size corresponds to number of responses, range 1–22); DRAGON score outcome estimates (red squares).

($P < 0.0001$) (Fig. 4). Physicians' accuracy increased with increasing values of the DRAGON score (i.e. with decreasing risk of favourable outcome) ($P = 0.03$, Fig. 2, lower left panel). The majority of physicians' inaccurate estimates ($n = 415$, 85.7%) represented an underestimation of the probability of unfavourable outcome.

DRAGON score versus physicians for the prediction of 3-month mRS ≥ 5 in IVT patients

Two hundred and twenty physicians responded to 644 scenarios (380 different real-life IVT patients). 518 (80.4%) DRAGON score estimates and 260 (40.4%) physicians' estimates were accurate ($P < 0.0001$) (Fig. 5). There was an interaction between physicians' accuracy and patients' DRAGON score ($P = 0.04$), i.e. the degree of physicians' accuracy was highest in patients with the lowest DRAGON score (Fig. 2, lower right panel). The majority of physicians' inaccurate estimates ($n = 247$, 64.3%) represented an overestimation of the probability of unfavourable outcome.

Accuracy of clinicians' estimates

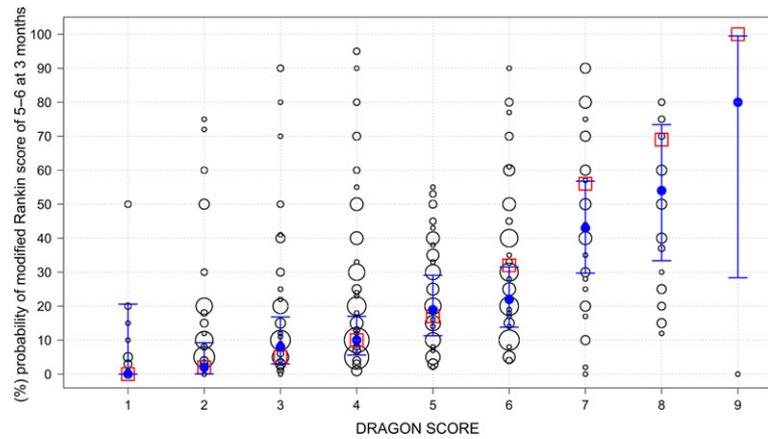
Overall, the majority of physicians' estimates were inaccurate (1422/2636, 53.9%). In sensitivity analysis, prognostic score estimates were more accurate than clinicians' estimates regardless of their age, sex, specialization, specialty status, experience and hospital setting (Table S1).

Discussion

The present study compared three well-validated prognostic scores to a large set of physicians interested in stroke medicine aiming to investigate if the accuracy of their outcome prediction was similar. It was found that stroke outcome was predicted much more accurately by prognostic scores compared with physicians interested in stroke medicine.

Perhaps some would be surprised by the results that prognostic scores predict stroke outcome more accurately than stroke physicians, even the more experienced ones, especially taking into account that the patients' parameters presented in the scenarios were

Figure 5 Probability of mRS ≥ 5 in thrombolysed patients stratified by DRAGON score: actual outcome and 95%CI range (blue dots); physicians' outcome estimates (white circles; size corresponds to number of responses, range 1–33); DRAGON score outcome estimates (red squares).



more extensive than the parameters needed to calculate the three scores. However, this should not come as a surprise because these scores represent the output of methodologically solid mathematical models which were developed and validated on the basis of dozens of demographic, clinical, laboratory and radiological parameters of thousands of patients.

This is the second study that shows that a stroke prognostic score is more accurate than physicians interested in stroke: recently, the Clinician Judgment versus Risk Score to Predict Stroke Outcomes (JURaSSiC) study showed that the iScore performed more accurately in outcome prediction of acute ischaemic stroke patients than clinicians with expertise in stroke [19]. The present study expands these results to include three more well-validated risk scores which may allow for the prediction of several stroke outcomes, i.e. functional outcome in the general ischaemic stroke population and in IVT patients, as well as sICH in the latter population.

Stroke prognostic scores were criticized for being aimed for non-specialist physicians and that they would be of no use for stroke specialists [20]. However, the present study and the JURaSSiC study [19] show that probably this is not the case given that scores are much more accurate than physicians; of importance, this finding remained across several physician subgroups such as stroke expertise, specialization, specialty status, hospital setting and others. Also, they were criticized that they do not incorporate technology like brain and arterial imaging; recently, the addition of multimodal imaging did not increase the prognostic accuracy of the ASTRAL score [21] but it improved the accuracy of the DRAGON score [9,10]. Still, one should keep in mind that these scores were developed to be used at admission and, in most centres worldwide, advanced multimodal imaging is not available at admission.

The finding that the ASTRAL, DRAGON and SEDAN scores (as well as the iScore [19]) make a more accurate prediction of stroke outcome than physicians interested in stroke medicine should remind the latter to be careful when discussing outcome prediction in the acute phase of stroke. It also justifies the use of these scores whenever an estimate of a patient's outcome is needed, and independently of the physicians' experience and self-perceived likelihood of patient outcome. This may apply to the clinical setting, e.g. informing patients and relatives, planning the long-term living setting of the patient and, perhaps, even for treatment decisions: patients with NIHSS < 4 were excluded from the IVT randomized trials and, as a result, many patients with an NIHSS < 4 are not treated with thrombolysis even though meta-analyses showed that these patients may also benefit [22]. In this context, the use of stroke prognostic scores could identify patients with an NIHSS < 4 but with a moderate outcome prognosis who could possibly benefit from thrombolysis and thus allow stroke physicians to bypass this 'rigid' NIHSS threshold and tailor treatment to each patient's characteristics. Perhaps just as importantly, the use of prognostic scores may be warranted in the research setting, e.g. scores could be used in acute stroke treatment trials as selection criteria leading to exclusion of patients whose pre-treatment likelihood of a favourable outcome is too high or too low to allow detection of a treatment effect. This could reduce 'noise' and therefore the estimated population sample necessary to show a treatment effect, especially if the study aims at achieving a rigid cut-off such as a functional outcome defined as mRS 0–2. In a shift-analysis approach, however, treatment effects across functional outcomes may be discovered, and pre-randomization prognostication may be less useful [23]. Also, scores may be used in non-randomized studies

to control for case-mix variation. In addition, the implementation of risk scores could be a useful tool towards increasing thrombolysis rates: our results showed that most stroke physicians overestimated the ICH risk after IVT, which frequently discourages physicians from offering this treatment modality in eligible acute stroke patients. In this context, the implementation of the SEDAN score for more accurate ICH prediction could reassure hesitant treating physicians and prevent them from depriving their patients of this treatment. Finally, this result could also have implications for stroke care organizations given that anyone can have 90% accuracy in predicting ICH after thrombolysis using the SEDAN score.

The strengths of this study include the large number of scenarios used, the fact that scenarios were different for and randomly allocated to each participant, the large and representative set of responders originating from 32 different countries, the considerable stroke expertise of participating physicians and, finally, the fact that the survey responders based their estimates on a broader set of patient parameters than those included in the scores. A possible bias of our study was that the scenarios were derived from the ASTRAL registry which was also used for the development of the ASTRAL score; however, the patients included in this analysis were different from those included in the ASTRAL score development, and the ASTRAL score performed similarly well when validated externally in several other stroke registries [6,7]. Another limitation is that the patient dataset originated from a single hospital registry characterized by specific stroke expertise and resources, which obviously influenced stroke outcome; it is possible that outcomes in these patients may have been different from what would be expected by participating physicians serving in other hospitals with different levels of stroke expertise and resources. Similar studies using datasets from different stroke registries would be useful to confirm the conclusions.

The ASTRAL, DRAGON and SEDAN scores predict outcome of patients with acute ischaemic stroke with higher accuracy compared to a large representative set of physicians interested in stroke medicine. This highlights the importance of using these scores in clinical practice and research to help clinicians make better-informed decisions and researchers design studies more efficiently.

Disclosure of conflicts of interest

Michel, Ntaios, Papavasileiou, Strbian: involved in development and/or validation of ASTRAL,

DRAGON, SEDAN and/or the simple variables model. Gioulekas: no disclosures.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Data S1. Scores and outcomes.

Data S2. Scenarios structure.

Table S1. Sensitivity analysis.

References

1. Ntaios G, Papavasileiou V, Michel P, Tattlisumak T, Strbian D. Predicting functional outcome and symptomatic intracranial hemorrhage in patients with acute ischemic stroke: a glimpse into the crystal ball? *Stroke* 2015; **46**: 899–908.
2. Ntaios G, Faouzi M, Ferrari J, Lang W, Vemmos K, Michel P. An integer-based score to predict functional outcome in acute ischemic stroke: the ASTRAL score. *Neurology* 2012; **78**: 1916–1922.
3. Strbian D, Engelter S, Michel P, *et al.* Symptomatic intracranial hemorrhage after stroke thrombolysis: the SEDAN score. *Ann Neurol* 2012; **71**: 634–641.
4. Strbian D, Meretoja A, Ahlhelm FJ, *et al.* Predicting outcome of IV thrombolysis-treated ischemic stroke patients: the DRAGON score. *Neurology* 2012; **78**: 427–432.
5. Strbian D, Seiffge DJ, Breuer L, *et al.* Validation of the DRAGON score in 12 stroke centers in anterior and posterior circulation. *Stroke* 2013; **44**: 2718–2721.
6. Papavasileiou V, Milionis H, Michel P, *et al.* ASTRAL score predicts 5-year dependence and mortality in acute ischemic stroke. *Stroke* 2013; **44**: 1616–1620.
7. Liu G, Ntaios G, Zheng H, *et al.* External validation of the ASTRAL score to predict 3- and 12-month functional outcome in the China National Stroke Registry. *Stroke* 2013; **44**: 1443–1445.
8. Strbian D, Michel P, Seiffge DJ, *et al.* Symptomatic intracranial hemorrhage after stroke thrombolysis: comparison of prediction scores. *Stroke* 2014; **45**: 752–758.
9. Turc G, Aguetaz P, Ponchelle-Dequatre N, *et al.* External validation of the MRI-DRAGON score: early prediction of stroke outcome after intravenous thrombolysis. *PLoS One* 2014; **9**: e99164.
10. Turc G, Apoil M, Naggara O, *et al.* Magnetic resonance imaging-DRAGON score: 3-month outcome prediction after intravenous thrombolysis for anterior circulation stroke. *Stroke* 2013; **44**: 1323–1328.
11. Giralt-Steinhauer E, Rodriguez-Campello A, Cuadrado-Godia E, *et al.* External validation of the DRAGON score in an elderly Spanish population: prediction of stroke prognosis after IV thrombolysis. *Cerebrovasc Dis* 2013; **36**: 110–114.
12. Ovesen C, Christensen A, Nielsen JK, Christensen H. External validation of the ability of the DRAGON score to predict outcome after thrombolysis treatment. *J Clin Neurosci* 2013; **20**: 1635–1636.
13. Asuzu D, Nystrom K, Amin H, *et al.* Comparison of 8 scores for predicting symptomatic intracerebral

- hemorrhage after IV thrombolysis. *Neurocrit Care* 2015; **22**: 229–233.
14. Seiffge DJ, Karagiannis A, Strbian D, *et al.* Simple variables predict miserable outcome after intravenous thrombolysis. *Eur J Neurol* 2014; **21**: 185–191.
 15. <https://gioulekas.pythonanywhere.com/Study/default/index>. (accessed 03/06/2015)
 16. Michel P, Odier C, Rutgers M, *et al.* The Acute STroke Registry and Analysis of Lausanne (ASTRAL): design and baseline analysis of an ischemic stroke registry including acute multimodal imaging. *Stroke* 2010; **41**: 2491–2498.
 17. Hacke W, Kaste M, Fieschi C, *et al.* Randomised double-blind placebo-controlled trial of thrombolytic therapy with intravenous alteplase in acute ischaemic stroke (ECASS II). Second European-Australasian Acute Stroke Study Investigators. *Lancet* 1998; **352**: 1245–1251.
 18. R: A Language and Environment for Statistical Computing. R Core Team; R Foundation for Statistical Computing. Vienna; Austria 2014. www.R-project.org.
 19. Saposnik G, Cote R, Mamdani M, *et al.* JURaSSiC: accuracy of clinician vs risk score prediction of ischemic stroke outcomes. *Neurology* 2013; **81**: 448–455.
 20. Caplan LR. Scores of scores. *JAMA Neurol* 2013; **70**: 252–253.
 21. Ntaios G, Papavasileiou V, Faouzi M, Vanacker P, Wintermark M, Michel P. Acute imaging does not improve ASTRAL score's accuracy despite having a prognostic value. *Int J Stroke* 2014; **9**: 926–931.
 22. Emberson J, Lees KR, Lyden P, *et al.* Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. *Lancet* 2014; **384**: 1929–1935.
 23. Lees KR, Bath PM, Schellinger PD, *et al.* Contemporary outcome measures in acute stroke research: choice of primary outcome measure. *Stroke* 2012; **43**: 1163–1170.