Targeting rtPA in acute ischemic stroke based on risk of intracranial hemorrhage or poor functional outcome: an analysis of the IST-3 trial

William N. Whiteley, PhD[1], **Douglas Thompson, BSc**[2], **Gordon Murray, PhD**[2], **Geoff Cohen, MA**[1], **Richard I. Lindley, FRCP**[3], **Joanna Wardlaw, FRCR FRCP**[4], and **Peter Sandercock, FRCPE**[1] **On behalf of the IST-3 collaborative group**

[1]Division of Clinical Neurosciences, University of Edinburgh, Bramwell Dott Building, Western General Hospital, Edinburgh EH4 2XU

[2]Edinburgh MRC Hub for Trials Methodology Research, Centre for Population Health Sciences, University of Edinburgh UK

[3]Neurological and Mental Health Division, The George Institute for Global Health, The University of Sydney, Sydney, Australia

[4]Neuroimaging Sciences, Edinburgh, UK & Division of Clinical Neurosciences, University of Edinburgh, UK

Abstract

Background and purpose—Intravenous rtPA, despite a risk of early symptomatic intracranial hemorrhage (SICH), is of net clinical benefit to acute stroke patients. We tested if predictive models could identify patients least likely to be harmed by SICH or those who gained no net benefit.

Methods—We used the IST-3 trial dataset, an international, multicentre, open treatment randomised trial of 0.9 mg/kg rtPA versus control in 3035 acute ischemic stroke patients. We compared the discrimination and calibration of previously developed predictive models for ICH and post-stroke poor outcome and developed a new model using variables selected by systematic

review. We calculated the absolute and relative risk reduction of death or dependency with rtPA in patients at a low, medium or high predicted risk of SICH or poor functional outcome calculated by model.

Results—Prediction models for SICH or poor outcome (HAT, SEDAN, GRASPS, Stroke TPI, DRAGON, THRIVE, our new model, and a model with NIHSS and age) had similar AUROCC to predict SICH (P for difference>0.05). The simplest model (with covariates NIHSS and age) predicted both SICH (AUROCC 0.63, 95%CI:0.58-0.68) and post stroke poor functional outcome (AUROCC 0.80, 95%CI: 0.77-0.82) similarly to complex models. There was no evidence that the effect of rtPA in patients at high predicted risk of SICH or poor functional outcome after stroke was less than in those at lower risk.

Conclusions—There is a clinically relevant net positive effect of rtPA in acute stroke patients at a high predicted risk of SICH or poor functional outcome.

INTRODUCTION

Intravenous thrombolytic therapy within 3 hours of acute ischemic stroke is associated with a 3-4% absolute increase in the risk of symptomatic intracranial hemorrhages (SICH)[1], that are either fatal or increase the risk of dependence.[2] Despite this early hazard, the net clinical effect of thrombolysis is substantial; for every 1000 patients treated under 3 hours, 90 more will be alive and independent post stroke.[1] However, clinicians may be unduly concerned about the early risk of SICH,[3] perhaps denying some patients the opportunity of clinical benefit from treatment.

Scores to predict very high SICH risk or negligible clinical benefit from intravenous thrombolysis might be help clinicians to select patients for treatment. In order to test the hypothesis that prediction scores improve selection and lead to greater net clinical benefit, we analysed data from the third international stroke trial (IST-3). We wished to use the best possible prediction models, so aimed to: undertake a systematic review of previous models, develop new models with novel statistical approaches, and then estimate the likely clinical impact with each model. In other words, we sought to assess whether the benefits and harms of thrombolysis vary in groups with different predicted prognosis.

METHODS

Ethics statement

The study was approved by the Multi-centre Research Ethics Committees, Scotland (reference MREC/99/0/78), and by local ethical committees. Patients or a valid proxy gave written consent to participate. This trial was registered (ISRCTN25765518).

IST-3 study design and participants

The details of the IST-3 study protocol,[4] statistical analysis plan,[5] and primary outcomes[6] have been published previously. In brief, ischemic stroke patients (with no upper age limit) who could start rtPA treatment within 6 hours of symptom onset, and in whom the randomising clinician was substantially uncertain about the risks and benefits of rtPA were

randomised 1:1 to standard care with an infusion of 0.9mg/kg rtPA or standard care without rtPA.

Measurement of baseline variables and clinical outcomes

For these analyses we used baseline clinical variables that had been measured and recorded by the treating clinician before randomisation, non-blinded information collected post randomisation, and findings from the brain scans which had been read by an expert panel blinded to clinical details and allocated treatment.

The trial event adjudication committee defined 'symptomatic post-rtPA ICH' (SICH) as a clinically significant deterioration or death within the first 7 days of treatment with evidence of either significant brain parenchymal hemorrhage (local or distant from the infarct) or significant hemorrhagic transformation of an infarct on brain imaging.[5] In addition, we extracted the variable 'any significant radiological post-rtPA ICH' by 7 days, measured by a blinded neuroradiology rater either on routine brain imaging 24-48 hours post randomisation or any scans performed in case of clinical deterioration (equivalent to 'parenchymal hemorrhage type 2' measured in previous trials of iv rtPA).[7] The primary measure of clinical outcome was the Oxford Handicap Scale (OHS) measured at 6 months after randomisation. We defined 'poor functional outcome' as an OHS of 3-6 (dead or dependent). We performed a post-hoc sensitivity analysis using a definition of poor functional outcome of OHS 5-6 (dead or dependent for all cares).

Identification of previously developed prediction models

We identified published clinical prediction scores by systematically searching the literature for models that aimed to predict post-rtPA SICH, or poor functional outcome after rtPA. (please see http://stroke.ahajournals.org). We hypothesised that a simple model containing only the variables NIHSS and age[8] would predict both SICH and poor functional outcome as well as the other scores.

Development of a new predictive model for SICH

We developed a new model to predict SICH in patients randomised to rtPA from the IST-3 trial. We created a binary logistic regression model with variables significantly associated with post-rtPA ICH in a systematic review.[7] We tested model assumptions of linearity and additivity and the effect of missing data. We internally validated the model with 150 bootstrap replicates and shrinkage of estimated regression coefficients to correct for overfitting.

Calibration, discrimination and classification of predictive models for SICH and poor functional outcome

We tested models performance in rtPA treated patients for the outcomes 'SICH', 'any significant radiological post-rtPA ICH' and 'poor functional outcome'. We measured discrimination with the area under receiver operator characteristic curves (AUROCC) which we compared non-parametrically. An AUROCC=1 indicates perfect discrimination, and AUROCC=0.5 indicates no better discrimination than chance. To test model calibration, we calculated the calibration slope and intercept by fitting a logistic regression model with

predicted risk as the only predictor (where a slope=1 and intercept=0 indicates a perfectly calibrated model) and compared the proportions of patients classified as low, medium and high risk with each model.

The choice of risk thresholds is controversial. In the absence of generally agreed thresholds, we used the mean of risk thresholds from previous studies to define: 'low', 'medium' and 'high' risk of post-rtPA intracranial hemorrhage and poor functional outcome. The means of the published thresholds were for ICH: 3%, 3-8% >8%; and for poor functional outcome: 35%, 35-56%, >56%). In a secondary analysis, we examined thresholds for a very high risk of SICH (>20%) and very high risk poor functional outcome (>70%).

Effect of rtPA in patients at high, intermediate and low risk of intracranial hemorrhage

We investigated the interaction between rtPA treatment and predictions of SICH or poor functional outcome on an absolute risk scale, by calculating the difference in the proportion of patients with poor functional outcome between patients treated with and without rtPA, in groups of patients at low, medium and high risk. Where possible, we re-calibrated the intercept of prediction models to the IST-3 dataset. To support this analysis we looked for interactions on a relative scale between treatment and predicted risk as a continuous variable, using ordinal logistic regression with the whole OHS as the dependent variable, after examining the proportional odds assumption.

We performed sensitivity analyses excluding those few patients randomised to rtPA who did not receive any; examining only those patients where the time to randomisation was <4.5 hrs; and only those treated after 4.5 hrs; and in addition made further adjustment for delay to treatment as a continuous variable. Post-hoc we repeated this analysis in patients randomised <3 hrs after stroke onset.

We used R version 2.13.1 for the statistical analysis. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

RESULTS

In patients treated with rtPA, 6.8% (104/1515) had an SICH within 7 days of randomisation; a further 2% (31) had a radiological hemorrhage by 7 days with no detectable clinical deterioration. The median time from randomisation to SICH was one day (IQR 1 to 2). By 6 months after randomisation, few patients who had suffered a SICH were independent in activities of daily living (8/104, 8%), compared to rtPA treated patients who did not have a SICH. (546/1411, 39%)

Patients who had an SICH (Table 1) were significantly (P<0.05) more likely to: have had a history of stroke or TIA; to have been taking an antiplatelet agent in the 48 hours prior to randomisation; to have had more neurological impairment or a higher blood glucose at randomisation, or to have had a visible infarct (in any location) or hyperdense artery on brain imaging. There was no detectable effect of delay to randomisation or delay to treatment on the odds of SICH.

Identification of previously developed prediction models

We identified 5 scores to predict post-rtPA ICH and 3 to predict post-rtPA poor functional outcome. We excluded four potentially relevant scores: two for which we were unable to calculate predicted risks from the published information, and two as they required baseline information that was not available in IST-3 (platelet count, and a diagnosis of cancer or renal failure), (please see http://stroke.ahajournals.org, Supplementary table I).

Development of model to predict SICH in IST-3

A logistic regression model for the prediction of SICH, developed in 1515 rtPA treated patients with variables significantly associated with intracranial hemorrhage from our previous systematic review [7] (age, NIHSS, glucose, prior hypertension, AF, antiplatelets, diabetes, leukoaraiosis and visible infarction), was able to discriminate modestly between patients with and without SICH (AUROCC corrected for optimism 0.65), and was well calibrated in this dataset (calibration slope corrected for optimism 1.12. intercept 0.32) (please see http://stroke.ahajournals.org, Supplementary table II).

There were no statistically or clinically significant two-way interactions between categorical and continuous variables, and there was no evidence of a non-linear relationship between any continuous variables with the odds of SICH. Multiple imputations for missing data made very little difference to the magnitude or direction of the estimates.

Calibration and discrimination of predictive models for SICH and poor functional outcome (Table 2)

All models to predict SICH or poor functional outcome discriminated modestly between patients who did and did not have an SICH (AUROCC range 0.56 to 0.68). The AUROCCs of all models were similar (P>0.05), apart from the dichotomised SPAN score which had significantly worse discrimination (P<0.05). Each previously developed model discriminated less well than in previous validation datasets. Models developed to predict SICH were better calibrated for the SICH outcome than those models developed to predict post-rtPA poor functional outcome, though all models (apart from the new score) over-predicted the risk of SICH. There were no important qualitative or quantitative differences in discrimination or calibration for any of the models when the outcome was 'radiological post-rtPA ICH' rather than 'symptomatic ICH' (please see http://stroke.ahajournals.org, Supplementary table III), or when we examined only those patients randomised with 3 hours of stroke onset. Each model classified a different proportion of the rtPA treated population at 'high', 'medium' and 'low' risk of SICH, differences that are potentially clinically relevant. (Figure 1)

All models discriminated moderately well between patients who did and did not have a poor functional outcome after stroke (AUROCC range 0.66 to 0.80). There were no significant differences in discrimination between models designed to predict poor functional outcome post-rtPA (Stroke TPI, NIHSS/age, DRAGON, new model, THRIVE) (differences all P>0.05), bar the SPAN score which was significantly worse than other models (P<0.001). All the models were well calibrated for death or dependence, whether or not they aimed to predict SICH or post stroke poor functional outcome. A sensitivity analysis examining a

different definition of functional outcome (OHS 5-6) made no difference to these conclusions.

Whilst the novel IST-3 score we developed had better discrimination than previous models to predict SICH, the absolute difference in the AUROCC between it and other models was small, and likely due to model over-fitting, and therefore we do not believe it will perform better than previously developed models in external validation.

Effect of rtPA in patients at high, intermediate and low risk of intracranial hemorrhage

In the 3035 patients in the IST-3 trial, we observed that the absolute risk reduction in poor functional outcome with rtPA treatment was greater both among patients at higher predicted risk of SICH (figure 2) and among patients at higher risk of poor functional outcome (figure 3). With the more statistically efficient ordinal logistic regression to measure treatment effect, there was no evidence of significant interactions between rtPA with continuous predicted risk of SICH or poor functional outcome on a relative scale. These conclusions were not changed by: excluding patients who were randomised but not treated with rtPA; patients who were randomised > 3 or > 4.5 hrs after stroke; making adjustment for delay to treatment as a continuous variable; or when examining higher thresholds of risk for SICH (>20%) or poor functional outcome (>70%). There was therefore no evidence to support a strategy of avoiding rtPA treatment in patients at a higher predicted risk of SICH or poor functional outcome in the IST-3 dataset.

DISCUSSION

The clinical effect of rtPA in patients at a higher predicted risk of SICH or poor functional outcome was at least as good, and possibly more so, than in patients with a lower risk. We found prediction scores discriminated only modestly well between patients who did and did not suffer a SICH, though discriminated moderately well between patients who did, and did not, have a poor functional outcome.

Our analyses suggest that clinical prediction scores are unlikely to play a role in selecting individual ischemic stroke patients for rtPA in routine practice. Patients (or their families) who want to know the probability of poor functional outcome or SICH could choose any one of these scores, accepting the uncertainty in absolute predicted risks for an individual. A simple score constructed with the fewest, most easily measured, clinical variables (for example NIHSS and age) would be the easiest to implement.

Our approach had a number of strengths. We selected comparator models from a systematic review, and measured the performance of models to predict important clinical outcomes in a large dataset. The IST-3 trial is broadly representative of current clinical practice as it included many elderly patients and patients with severe strokes and the rate of SICH was similar to that seen in clinical practice and previous clinical trials of rtPA[9]. This wider range of patients with differing prognoses from previous cohorts is strength of the analysis. We developed a new prediction model for intracranial haemorrhage minimising data dependent biases, and maximising the use of predictive information. Despite this, we were unable to

make very much better prediction than previously published models. We therefore did not validate this model in a new dataset.

There were no differences in our conclusions after sensitivity analyses restricted to patients treated less than 4.5 hours after stroke (the time threshold of the current EU licence for rtPA). The IST-3 trial had few missing baseline or outcome data (though glucose was not collected in the first 282 patients randomised); had a wide range in potentially predictive variables because of its wide inclusion criteria; and randomly allocated rtPA, so our conclusions about the use of scores to predict response to treatment are robust. Our conclusions are supported by recent work with observational data comparing treated and untreated acute stroke patients with a number of relative contraindications to rtPA (high glucose levels, extensive CT findings, etc.).[10]

We can identify limitations: IST-3 was an unblinded trial, though steps were taken to minimise bias. Overall IST-3 was a neutral trial in that there was no statistically significant difference in the dichotomous primary outcome, the proportion of patients dead or dependent after treatment with rtPA [OHS 0–2; adjusted OR 1.13, 95% CI 0.95–1.35]. However, the key secondary outcome, assessed by the more statistically efficient ordinal regression analysis showed clear evidence of a favourable shift in disability scores at both 6 and 18 months, [6, 11] and the effect of rtPA in IST-3 was similar to previous trials, after accounting for time to randomisation.[1] We tested the predictions of models constructed with easily measured baseline clinical and simple imaging variables. Future improvements in prediction are only likely if variables that we did not measure - such as advanced imaging methods, genotyping, or blood biomarkers related to the pathophysiology of post-rtPA ICH - better predict response to treatment.

CONCLUSION

Clinical prediction models were unable to identify patients least likely to be harmed by SICH or those who gained no net benefit from rtPA. These data suggest that intravenous rtPA has an absolute beneficial effect in patients at a high predicted risk of symptomatic intracranial hemorrhage or poor functional outcome.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Reference List

(1). Wardlaw JM, Murray V, Berge E, del Zoppo G, Sandercock P, Lindley RL, et al. Recombinant tissue plasminogen activator for acute ischaemic stroke: an updated systematic review and meta-analysis. The Lancet. 2012; 379:2364–72.

(2). Strbian D, Sairanen T, Meretoja A, Pitkaniemi J, Putaala J, Salonen O, et al. Patient outcomes from symptomatic intracerebral hemorrhage after stroke thrombolysis. Neurology. 2011; 77:341–8. [PubMed: 21715707]

(3). Brown DL, Barsan WG, Lisabeth LD, Gallery ME, Morgenstern LB. Survey of Emergency Physicians About Recombinant Tissue Plasminogen Activator for Acute Ischemic Stroke. Annals of Emergency Medicine. 2005; 46:56–60. [PubMed: 15988427]

(4). Sandercock P, Lindley R, Wardlaw J, Dennis M, Innes K, Cohen G, et al. Update on the third international stroke trial (IST-3) of thrombolysis for acute ischaemic stroke and baseline features of the 3035 patients recruited. Trials. 2011; 12:252. [PubMed: 22129158]

(5). IST-3 collaborative group. Statistical analysis plan for the third International Stroke Trial (IST-3); part of a thread of reports of the trial. Int J Stroke. 2012; 7:186–7. [PubMed: 22405276]

(6). IST-3 collaborative group. The benefits and harms of intravenous thrombolysis with recombinant tissue plasminogen activator within 6 h of acute ischaemic stroke (the third international stroke trial [IST-3]): a randomised controlled trial. The Lancet. 2012; 379:2352–63.

(7). Whiteley WN, Bruins Slot K, Fernandes P, Sandercock P, Wardlaw J. Risk Factors for Intracranial Hemorrhage in Acute Ischemic Stroke Patients Treated With Recombinant Tissue Plasminogen Activator: A Systematic Review and Meta-Analysis of 55 Studies. Stroke. 2012; 43:2904–9. [PubMed: 22996959]

(8). Konig IR, Ziegler A, Bluhmki E, Hacke W, Bath PMW, Sacco RL, et al. Predicting Long-Term Outcome After Acute Ischemic Stroke: A Simple Index Works in Patients From Controlled Clinical Trials. Stroke. 2008; 39:1821–6. [PubMed: 18403738]

(9). Wahlgren N, Ahmed N, Davalos A, Ford GA, Grond M, Hacke W, et al. Thrombolysis with alteplase for acute ischaemic stroke in the Safe Implementation of Thrombolysis in Stroke-Monitoring Study (SITS-MOST): an observational study. Lancet. 2007; 369:275–82. [PubMed: 17258667]

(10). Frank B, Grotta JC, Alexandrov AV, Bluhmki E, Lyden P, Meretoja A, et al. Thrombolysis in Stroke Despite Contraindications or Warnings? Stroke. 2013; 10 1161/STROKEAHA. 112.674622.

(11). IST-3 collaborative group. Effect of thrombolysis with alteplase within 6 h of acute ischaemic stroke on long-term outcomes (the third International Stroke Trial [IST-3]): 18-month follow-up of a randomised controlled trial. Lancet Neurol. 2013; 12:768–76. [PubMed: 23791822]
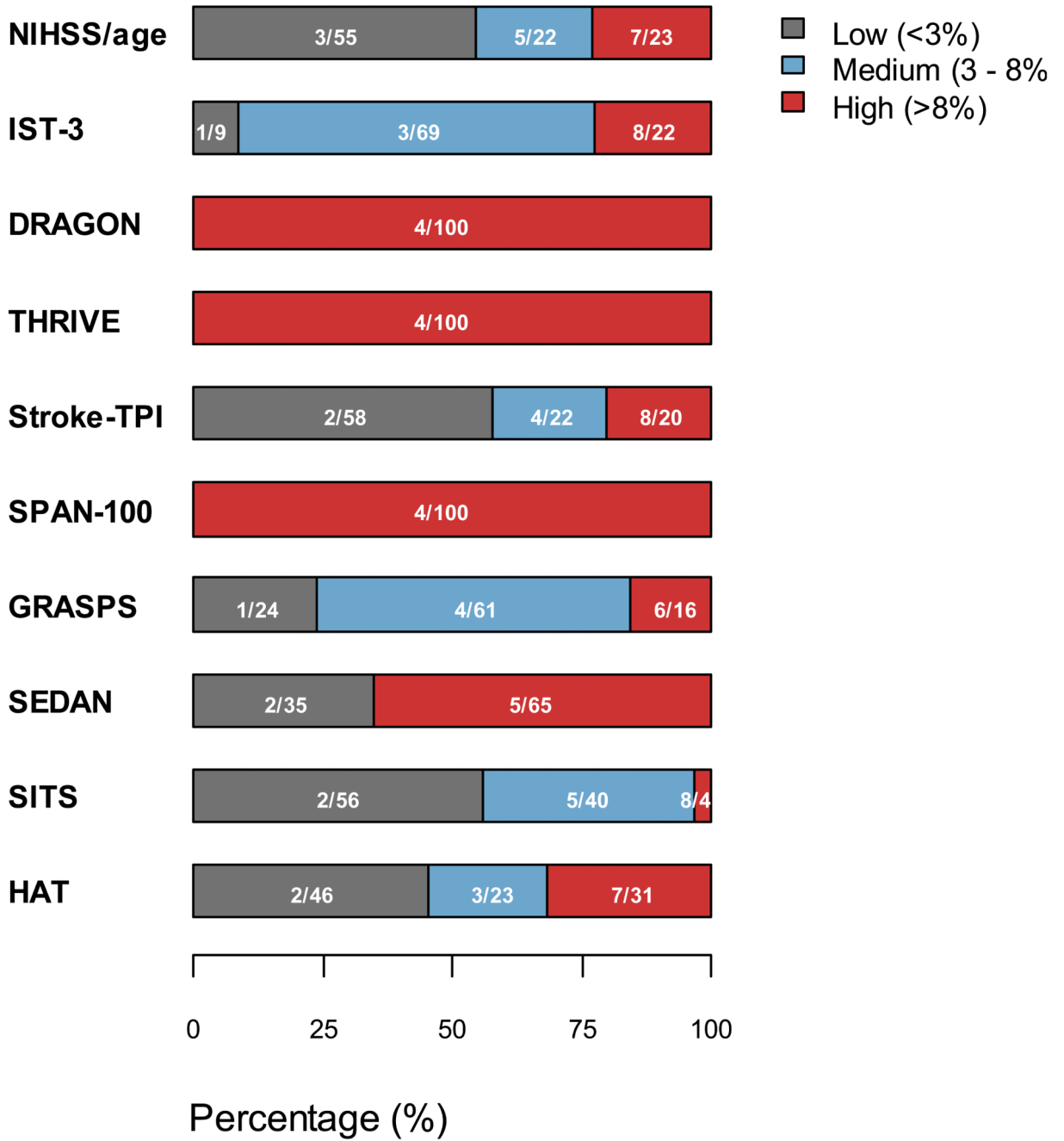
Figure 1.
Percentage of patients in predicted sICH categories (<3%, 3% to 8%, and >8%) across ten models. Numbers signify percentage with SICH in risk group/percentage of total patients in risk group

Risk strata by model (n/N)

Absolute risk reduction

Low risk (<3%)

HAT (581/1246)	0.01 [-0.04 , 0.07]
SEDAN (370/956)	0.03 [-0.04 , 0.09]
SITS (832/1523)	0.03 [-0.02 , 0.08]
GRASPS (206/649)	0.01 [-0.07 , 0.08]
IST3 (54/237)	-0.05 [-0.15 , 0.06]
Stroke TPI (703/1574)	-0.01 [-0.06 , 0.04]
NIHSS/age (731/1652)	0.00 [-0.05 , 0.05]

Medium risk (3 to 8%)

HAT (392/626)	0.02 [-0.06 , 0.10]
No patients classified by the SEDAN score	
SITS (815/1099)	-0.03 [-0.08 , 0.03]
GRASPS (1131/1657)	0.02 [-0.02 , 0.07]
IST3 (1143/1874)	0.01 [-0.04 , 0.05]
Stroke TPI (496/603)	0.07 [0.01 , 0.13]
NIHSS/age (561/677)	0.03 [-0.02 , 0.09]

High risk (>8%)

HAT (763/863)	0.04 [-0.01 , 0.08]
SEDAN (1366/1779)	0.02 [-0.02 , 0.06]
SITS (75/93)	0.10 [-0.06 , 0.26]
GRASPS (399/429)	0.04 [-0.00 , 0.09]
IST3 (532/612)	0.07 [0.02 , 0.12]
Stroke TPI (537/558)	0.05 [0.02 , 0.08]
NIHSS/age (647/688)	0.06 [0.03 , 0.09]

Overall effect (1939/3017)   0.02 [-0.02 , 0.05]

Harm | *Benefit*

-0.15   0.00   0.15   0.30

ARR

Figure 2.
Net clinical effect of rtPA among groups of patients at different risks of symptomatic intracranial hemorrhage. Absolute risk reduction (ARR) observed across three categories of predicted risk of SICH. A positive ARR indicates a reduced risk of death or dependency in the treated group, whilst a negative ARR suggests an excess risk in the rt-PA treated group. No patients were classified 'medium risk' by the SEDAN score.

Risk strata by model (n/N)   Absolute risk reduction

Low risk (<35%)

IST3 (1/7)	-0.20 [-0.55 , 0.15]
Stroke TPI (282/903)	-0.01 [-0.07 , 0.05]
DRAGON (123/506)	-0.05 [-0.12 , 0.03]
THRIVE (290/809)	0.02 [-0.05 , 0.08]
NIHSS/age (269/885)	-0.02 [-0.08 , 0.04]

Medium risk (35 to 56%)

IST3 (415/948)	-0.03 [-0.09 , 0.03]
Stroke TPI (389/630)	0.02 [-0.05 , 0.10]
DRAGON (547/1007)	0.02 [-0.04 , 0.09]
THRIVE (1039/1542)	0.01 [-0.04 , 0.05]
NIHSS/age (437/728)	0.02 [-0.05 , 0.09]

High risk (>56%)

IST3 (1313/1768)	0.04 [0.00 , 0.08]
Stroke TPI (1065/1202)	0.05 [0.02 , 0.09]
DRAGON (1064/1219)	0.05 [0.01 , 0.09]
THRIVE (605/658)	0.05 [0.01 , 0.09]
NIHSS/age (1233/1404)	0.05 [0.02 , 0.09]

Overall effect (1939/3017)   0.02 [-0.02 , 0.05]

Harm ⋮ *Benefit*

-0.30   -0.15   0.00   0.15
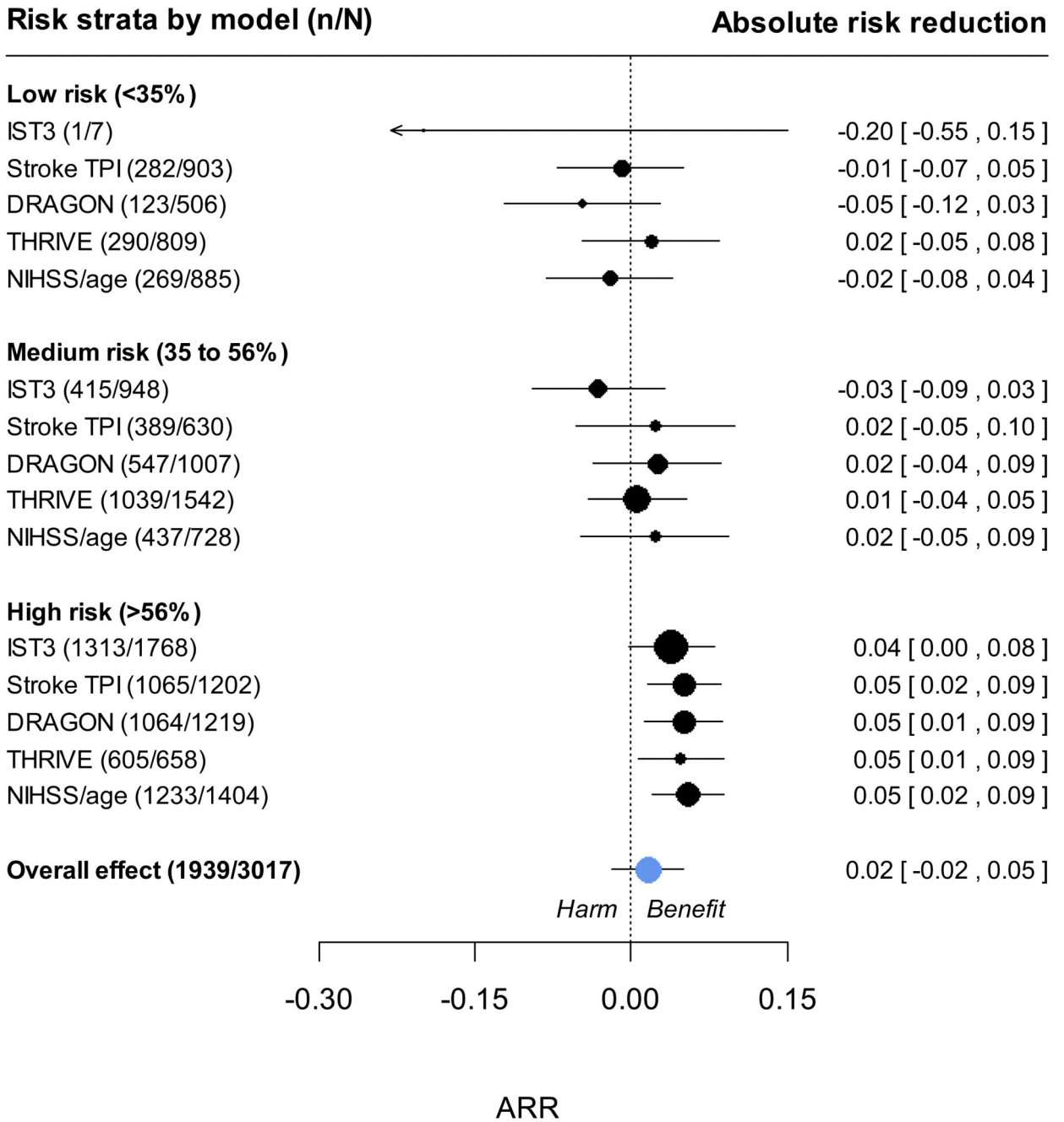
ARR

Figure 3.
Net clinical effect of rtPA among groups at different risks of poor functional outcome.
Absolute risk reduction (ARR) observed across three categories of predicted risk of poor
functional outcome. A positive ARR indicates an a reduced risk of death or dependency in
the treated group, whilst a negative ARR suggests an excess risk in the rt-PA treated group

Table 1

Baseline clinical variables and symptomatic intracranial hemorrhage in patients randomised to rtPA in IST-3, with univariate associations

	All (n = 1515)	sICH (n = 104)	No sICH (n = 1411)	OR	95% CI	P
Demographic risk factors						
Age (median, IQR)	81 (72 to 86)	81 (74 to 86)	81 (72 to 86)	1.01	0.99 to 1.03	0.3330
Male sex (n, %)	733, 48°%	60, 58%	673, 48%	1.50	1.00 to 2.24	0.0503
Weight (kg) (median, IQR)	70 (62 to 80)	70 (65 to 80)	70 (62 to 80)	1.01	0.99 to 1.02	0.2892
Stroke risk factors (n, %)						
Atrial Fibrillation	473, 31%	29, 28%	444, 31%	0.84	0.54 to 1.31	0.4472
Diabetes[1]	184, 12%	17, 16%	167, 12%	1.45	0.84 to 2.50	0.1783
Prior stroke or TIA[2]	354, 23%	34, 33%	320, 23%	1.65	1.08 to 2.53	0.0217
Prior hypertension[1]	975, 64%	64, 62%	911, 65%	0.90	0.59 to 1.36	0.6127
Treatment (n, %)						
Taking antiplatelets[3]	775, 51%	70, 67%	705, 50%	2.05	1.34 to 3.13	0.0009
Taking >1 antiplatelet[4]	80, 5%	10, 10%	70, 5%	2.21	1.10 to 4.46	0.0262
Taking warfarin or heparin[5]	50, 3%	1, 1%	49, 3%	0.29	0.04 to 2.11	0.2209
Stroke (median, IQR)						
NIHSS	11 (6 to 18)	15 (10 to 21)	11 (6 to 17)	1.06	1.04 to 1.09	<.00001
Blood glucose (mg/dl)[6]	126 (108 to 144)	126 (108 to 162)	126 (108 to 144)	1.00	1.00 to 1.01	0.0196
Systolic BP (mmHg)	156 (140 to 170)	160 (146 to 175)	155 (139 to 170)	1.01	1.00 to 1.02	0.0719
Diastolic BP (mmHg)[7]	80 (71 to 91)	82 (73 to 90)	80 (71 to 91)	1.00	0.99 to 1.02	0.6747
Delay to treatment (hrs)	4 (3 to 5)	4 (3 to 5)	4 (3 to 5)	0.91	0.77 to 1.07	0.2605
Imaging (n, %)						
Visible acute lesion on imaging	628, 41%	56, 54%	572, 41%	1.71	1.15 to 2.55	0.0085
Presence of leukoaraiosis[8]	765, 50%	57, 55%	708, 50%	1.25	0.83 to 1.87	0.2849
Hyperdense artery[8]	376, 25%	40, 38%	336, 24%	2.05	1.35 to 3.11	0.0007

sICH: symptomatic intracranial hemorrhage; NIHSS ; NIH stroke score

[1]
[2] missing values

[2] 3missing values

[3] 4 missing values

[4] 148 missing values

[5] 149 missing values

[6] 142 missing values

[7] 12 missing values

[8] 8 missing values

Table 2

Discrimination and calibration of models to predict intracranial hemorrhage and poor functional outcome after rtPA in IST-3 dataset

| Models | Score | n/N | Prediction of SICH | | | | n/N | Prediction of post-rtPA poor outcome | | | |
| | | | Discrimination | | Calibration | | | Discrimination | | Calibration | |
			AUROCC	95%CI	Intercept	slope		AUROCC	95%CI	Intercept	slope
HAT	0 to 5	87/1365	0.62	0.56 to 0.68	−0.29	0.39	856/1365	0.71	0.68 to 0.73	3.65	0.96
SEDAN	0 to 7	87/1365	0.62	0.56 to 0.69	−0.46	0.53	856/1365	0.74	0.71 to 0.76	3.19	0.99
SITS	0 to 11	85/1357	0.63	0.58 to 0.69	0.98	0.76	851/1357	0.66	0.63 to 0.69	4.45	0.79
GRASPS	45 to 101	87/1365	0.63	0.57 to 0.68	0.28	0.62	856/1365	0.77	0.74 to 0.79	3.76	1.56
SPAN-100 [1]	0 to 1	102/1507	0.56	0.52 to 0.61	−1.35	0.36	957/1507	0.66	0.64 to 0.68	2.19	1.26
STROKE TPI	GLM	87/1365	0.64	0.58 to 0.69	−1.73	0.33	856/1365	0.80	0.78 to 0.83	2.49	0.99
DRAGON2	0 to 10	87/1363	0.65	0.59 to 0.70	−3.68	0.47	855/1363	0.78	0.76 to 0.81	0.20	0.96
THRIVE	0 to 9	101/1504	0.60	0.55 to 0.66	−3.53	0.32	955/1504	0.76	0.74 to 0.79	0.26	1.07
NIHSS/age	GLM	102/1507	0.63	0.58 to 0.68	−4.46	0.27	957/1507	0.80	0.77 to 0.82	0.03	0.84
IST3 model	GLM	86/1361	0.68	0.63 to 0.74	0.01	1.32	854/1361	0.71	0.68 to 0.74	3.37	1.71

AUROCC: area under receiver operating characteristic curve ; DRAGON: Dense artery, Rankin score, Age, Glucose, Onset to treatment time, NIHSS score; GLM: generalised linear model. GRASPS: Glucose Race Age Sex Pressure Stroke Severity score; HAT: hemorrhage after thrombolysis score; NIHSS: National Institutes of Health Stroke Scale; SITS: Safe Implementation of Treatments in Stroke; SPAN: Stroke prognostication using age and NIHSS; THRIVE: Totalled Health Risk in Vascular Events; TPI: thrombolytic predictive instrument.

Discrimination is measured by the AUROCC. An AUROCC=1 indicates perfect discrimination, and AUROCC=0.5 indicates no better discrimination than chance. Calibration is measured by the intercept and slope of a calibration curve. A perfectly calibrated model has a slope=1 and intercept=0

[1] AUROCC for SPAN-100 is estimated from logistic regression analysis using the recommended dichotomy SPAN-100 100.