ORIGINAL RESEARCH

THE “FRAGILITY” OF STROKE THROMBOLYSIS

Daniel M Fatovich  MBBS FACEM PhD
Centre for Clinical Research in Emergency Medicine, Department of Emergency Medicine, Royal Perth Hospital, Western Australia, Australia; Harry Perkins Institute of Medical Research, University of Western Australia, Perth, Western Australia, Australia
Correspondence: Professor D M Fatovich, Department of Emergency Medicine, Royal Perth Hospital, GPO Box X2213, Perth WA 6847, Australia. Daniel.Fatovich@health.wa.gov.au

Abstract
The “fragility index” for a clinical trial reports how robust the results of the trial are. It calculates the minimum number of patients who would need to have a different primary outcome to change the measured p value from < 0.05 to > 0.05. A low fragility index means that only a small number of patients need to have a different outcome for the trial to lose traditional statistical significance. This increases the probability that the trial result was due to the play of chance. The application of this principle is described here for placebo-controlled trials of thrombolysis with alteplase (tPA) in acute stroke.

Tasman Medical Journal 2020; 2(1): 6-10

Introduction
One of the negative consequences of the premature adoption of novel therapies is that subsequent research assumes its value and enhances it with faster and wider implementation. Stroke thrombolysis has promoted new systems of care for stroke, including the development of stroke centres, Telestroke services and stroke ambulances with CT scanners. The mantra behind this is the phrase “time is brain”, which was borrowed from “time is muscle” for acute myocardial infarction. The basis for the phrase which propelled this juggernaut was the NINDS trial, published in 1995. In addition, a 2006 paper presenting a mathematical calculation based on rodent brain tissue promoted another mantra: 2 million neurons are lost per minute. This was
fuel for a campaign using marketing-based medicine.¹

Many doctors are simply unaware that of the 12 stroke thrombolysis trials that led to this situation, only two are reported as positive. Of the 10 that failed to find benefit for stroke thrombolysis, 4 had to be stopped early for harm (increased mortality). That leaves only 2 supportive studies (NINDS and ECASS-III) for our current situation.

**NINDS**

The National Institute of Neurologic Disorders and Stroke (NINDS)² study was the trial that started this controversy. This is the only placebo-controlled trial reporting a benefit for acute ischaemic stroke, if the trial drug (tPA) is given within 3 hrs of onset of symptoms. The trial consisted of two parts. Part 1 (n = 291) found no difference in the National Institutes of Health Stroke Severity score criterion (improvement by 4 points), or resolution of deficits at 24 hrs. Part 2 (n = 333) was an extension to Part 1, added for unexplained reasons. This “part” had an originally planned primary outcome of 20% improvement in functional and stroke scores, but this was changed post-hoc to a dichotomous favourable (modified Rankin Scale 0-1) vs unfavourable (modified Rankin Scale 2-6) outcome at 3 months. The original planned primary outcome was not reported, and it is reasonable to question this decision. The results were that 39% of tPA patients achieved the new endpoint vs 26% on placebo, an absolute difference of 13% (relative difference 33%). However, this was counterbalanced by symptomatic intracranial haemorrhage (0.6% placebo, 6.4% tPA; relative difference 1067%), of which 45% in the tPA group died.

In NINDS, there was an imbalance in baseline stroke severity, as the placebo group had more severe strokes (27.5% vs 18.3% tPA) and more mild strokes in the tPA group (19% vs 4.2% placebo).³ It is well known that stroke severity is a major determinant of outcome, and that this imbalance could explain the result.⁴ The study methodology also required selective trial enrolment, with equal numbers in the 0-90 min time frame, and the 91-180 min time frame. In a re-analysis using the original data and controlling for stroke severity, the authors found no benefit with tPA and no support for the time is brain hypothesis.⁵ This re-analysis result is consistent with the totality of the stroke thrombolysis trial results. Unfortunately, there has been no replication study that validates the original result. Yet replication is a fundamental part of the scientific process.

In summary, the NINDS study (part 1) was negative. The extension study (part 2) was subject to unexplained protocol changes, and reported major drug-related adverse effects, including a 10-fold increase in intracranial (including fatal) haemorrhage that offset the modest advantage in outcome in survivors.

**ECASS**

In 2008, ECASS-III was published, reporting a 7% absolute improvement of tPA vs placebo, in the 3-4.5 hr time frame.⁶ Again, there was an imbalance in baseline stroke severity, with Shy reported on an updated online result that meant there was a significant difference between the groups, with placebo again having more severe strokes (placebo 14.1% vs tPA 7.7%).⁷ This results in no benefit for tPA in the adjusted analysis. The original version of the paper reported that there was no baseline difference between the groups. Similarly, if the primary endpoint in ECASS-III is recategorized from a modified Rankin Scale (mRS) 0-1 vs 2-6 to a more realistic mRS 0-2 vs 3-6, the supposed benefits of tPA disappears.⁸

The inter-rater reliability of outcome assessments using the mRS are moderate at best.⁹¹⁰ A clinical trial has internal validity if and only if the imbalance between groups and bias in the assessment of outcome and chance, have been excluded as possible explanations for the difference in outcomes.¹¹ Newer concepts such as the fragility index (FI) reinforce this.

**FRAGILITY INDEX APPLIED TO NINDS AND ECASS**

The FI is a method to understand how statistically reproducible a study is. It is the minimum number of patients who would need
to have a different outcome to change the p value from < 0.05 to > 0.05, i.e. from statistically significant to insignificant. The 0.05 threshold as a measure of statistical significance is problematic, but this is not discussed in the present analysis. A low FI means that only a small number of patients would have to have their outcome change for the trial to lose statistical significance. It is simply calculated by repeatedly applying Fisher’s exact test, while successively reallocating patients, one at a time, from the favourable outcome group to the other (control) group. Thus the size of the trial remains unchanged. This can be done with online FI calculators, which add events to the group with a smaller number of events (either placebo or active). The FI can be applied only to trials in which there is a significant difference by treatment group in the primary outcome.

The FI of the NINDS trial is 3. This means that only 3 patients would have to have had a different outcome to change the study from positive to negative. The FI of ECASS-III is 1. When these results are considered in light of the methodological flaws described above and the lack of internal validity, it defies credibility to understand how one unreplicated study had the effect of changing international practice. It has been argued that the case for thrombolysis is built on a house of cards. Data from stroke registries and community studies, used for post-marketing monitoring in the real world of clinical practice, supports the case that treatment of stroke with tPA is worse than placebo.

**DISCUSSION**

**IST-3 Trial and meta-analyses**

The IST-3 trial, published in 2012, is the largest clinical trial of stroke thrombolysis (n = 3035), up to 6 hrs after stroke onset. This was a negative trial for its primary endpoint. The first 276 patients were double-blind and in this group no advantage of tPA over placebo was observed. Hence the FI cannot be calculated. The subsequent patients were open-label and despite the study’s negative result, it was concluded to be positive by using a secondary outcome measure (ordinal shift in the Oxford Handicap Score) that has itself been the subject of controversy. In the 3-4.5 hr subgroup (n = 1177), no difference in functional outcomes in those randomized to tPA vs placebo (32% tPA vs 38% placebo (OR 0.73) was reported. However, this was obtained by using a 99% confidence interval (0.50-1.07). Readers will be aware that the conventional reporting approach is the 95% CI. Use of a 95% CI results in a statistically significant association between placebo and good functional outcome.

Emberson et al (2014) published a meta-analysis (MA) which concluded that tPA significantly improves the overall odds of a good stroke outcome when delivered within 4.5 hrs of onset. This was irrespective of age or stroke severity and despite an increased risk of fatal intracranial haemorrhage. About half their included patients were from IST-3, the largest stroke thrombolysis trial, which was negative on its primary outcome. IST-3 found no consistent relationship between time-to-treatment and outcome. It is difficult to understand how the negative finding of IST-3 did not move the MA needle in a more negative direction. If most clinical trials including IST-3 were negative, but the MA is positive, trials other than IST-3 are driving the outcome.

Although Emberson et al promoted their MA as using individual patient data, this does not negate the problem of heterogeneity. Trials must be clinically and statistically homogeneous to be validly pooled. One needs a p value for test of heterogeneity > 0.1 to remove heterogeneity as a problem. Emberson et al did not report an assessment of heterogeneity and this was not provided when requested. Their time-to-treatment analysis is also invalid because there is no evidence that time-to-treatment functions as a continuous variable. Attempting to compare patient outcome based on time-to-treatment, therefore, degenerates into a comparison between heterogeneous study populations. Despite this, the authors forced all data into a regression equation, which implied that thrombolysis is beneficial up to five hours after stroke. This relationship is likely due to confounding factors.
The Emberson et al paper also had multiple conflicts of interest. It has been shown that industry supported studies are associated with more favourable conclusions than independent sources. The Institute of Medicine has published guidelines on MA and recommends exclusion of “individuals with a clear financial conflict of interest”. The multiple conflicts of interest listed in the Emberson et al paper show that these guidelines were not followed. A 2016 MA by Donaldson et al reported no conflicts of interest, and reached a more skeptical conclusion that promoted replication studies.

In 2018, the author of the original phrase “time is brain” wrote: “It is no longer reasonable to believe that the effect of time on the ischaemic process represents an absolute paradigm. It is increasingly evident that the volume of injured tissue within a given interval after the estimated time of onset shows considerable variability in large part due to the beneficial effect of a robust collateral circulation.” The effect of any intervention in stroke is due to factors such as the precise automated measurement of a small ischaemic core, a large reversible penumbra, and the use of thrombectomy in highly selected patients with detailed consent, coupled with absence of treatment detriment. The endovascular trials have demonstrated that it is possible to identify a cohort of patients who will benefit from revascularization, independent of time of onset, when perfusion imaging is paired with an effective treatment modality. This has not been achieved for stroke thrombolysis using tPA.

It is time to recognise the “fragility” of the stroke thrombolysis literature and ask again whether treatment with tPA has a real benefit, when the harms are proven.

**Review:** Reviewed internally by TMJ

**Ethics approval:** Not required (literature-based)

**Conflicts of interest:** None declared

**REFERENCES**


