

Are your MRI contrast agents cost-effective?
Learn more about generic Gadolinium-Based Contrast Agents.



AJNR









Reliability of the Modified TIC1 Score among Endovascular Neurosurgeons

D.M. Heiferman, N.C. Pecoraro, A.W. Wozniak, K.C. Ebersole, L.M. Jimenez, M.R. Reynolds, A.J. Ringer and J.C. Serrone

This information is current as of April 18, 2024.

AJNR Am J Neuroradiol 2020, 41 (8) 1441-1446
doi: <https://doi.org/10.3174/ajnr.A6696>
<http://www.ajnr.org/content/41/8/1441>

Reliability of the Modified TICI Score among Endovascular Neurosurgeons

 D.M. Heiferman,  N.C. Pecoraro,  A.W. Wozniak,  K.C. Ebersole,  L.M. Jimenez,  M.R. Reynolds,  A.J. Ringer, and  J.C. Serrone



ABSTRACT

BACKGROUND AND PURPOSE: The modified TICI score is the benchmark for quantifying reperfusion after mechanical thrombectomy. There has been limited investigation into the reliability of this score. We aim to identify intra-rater and inter-rater reliability of the mTICI score among endovascular neurosurgeons.

MATERIALS AND METHODS: Four independent endovascular neurosurgeons (raters) reviewed angiograms of 67 patients at 2 time points. κ statistics assessed inter- and intrarater reliability and compared raters'-versus-proceduralists' scores. Reliability was also assessed for occlusion location and by dichotomizing modified TICI scores (0–2a versus 2b–3).

RESULTS: Interrater reliability was moderate-to-substantial, weighted $\kappa = 0.417$ –0.703, overall $\kappa = 0.374$ ($P < .001$). The dichotomized modified TICI score had moderate-to-substantial interrater agreement, κ statistics = 0.468–0.715, overall $\kappa = 0.582$ ($P < .001$). Intrarater reliability was moderate-to-almost perfect, weighted $\kappa = 0.594$ –0.81. The dichotomized modified TICI score had substantial-to-almost perfect reliability, $\kappa = 0.632$ –0.82. Proceduralists had fair-to-moderate agreement with raters, weighted $\kappa = 0.348$ –0.574, and the dichotomized modified TICI score had fair-to-moderate agreement, $\kappa = 0.365$ –0.544. When proceduralists and raters disagreed, proceduralists' scores were higher in 79.6% of cases. M1 followed by ICA occlusions had the highest agreement.

CONCLUSIONS: The modified TICI score is a practical metric for assessing reperfusion after mechanical thrombectomy, though not without limitations. Agreement improved when scores were dichotomized around the clinically relevant threshold of successful revascularization. Interrater reliability improved with time, suggesting that formal training of interventionalists may improve reporting reliability. Agreement of the modified TICI scale is best with M1 and ICA occlusion and becomes less reliable with more distal or posterior circulation occlusions. These findings should be considered when developing research trials.

ABBREVIATIONS: mTICI = modified TICI; TIMI = Thrombolysis in Myocardial Infarction

Following the paradigm-shifting studies that demonstrated the efficacy of endovascular mechanical thrombectomy for acute, large-vessel occlusion in 2015, timely and thorough revascularization in cases of large-vessel occlusions has become a tenet of acute ischemic stroke care. The modified TICI (mTICI) score has been the preferred grading system for completeness of reperfusion across major studies.^{1–9} However, despite the clinical importance that reperfusion

has shown for patients' likelihood of an independent functional outcome, there has been limited investigation into the reliability of the mTICI score.¹⁰ We sought to investigate the reliability of the mTICI score among experienced neuroendovascular surgeons.

MATERIALS AND METHODS


Case Selection

Following institutional review board approval (IRB LU No. 210370), a retrospective review was conducted of all patients who underwent an endovascular mechanical thrombectomy for acute ischemic stroke at Loyola University Medical Center between January 2015 and March 2016. Patients were excluded for not having complete pre- and postprocedural anterior-posterior and lateral-projection DSAs of the entire intracranial cavity available for review. Data were gathered on the thrombus location and categorized as the following: ICA, first-segment middle cerebral artery (M1), second-segment middle cerebral artery (M2), or

Received March 9, 2020; accepted after revision May 18.

From the Department of Neurological Surgery (D.M.H., N.C.P., M.R.R., J.C.S.), Clinical Research Office (A.W.W.), Loyola University Stritch School of Medicine and Loyola University Medical Center, Maywood, Illinois; Department of Neurological Surgery (K.C.E.), University of Kansas Medical Center, Kansas City, Kansas; and Mayfield Brain & Spine (L.M.J., A.J.R.), Cincinnati, Ohio.

Please address correspondence to Daniel M. Heiferman, MD, Department of Neurological Surgery, Loyola University Stritch School of Medicine, 2160 South First Ave, Maywood, IL 60153; e-mail: daniel@heiferman.com

 Indicates article with supplemental on-line table.

<http://dx.doi.org/10.3174/ajnr.A6696>

TICI - Survey 1

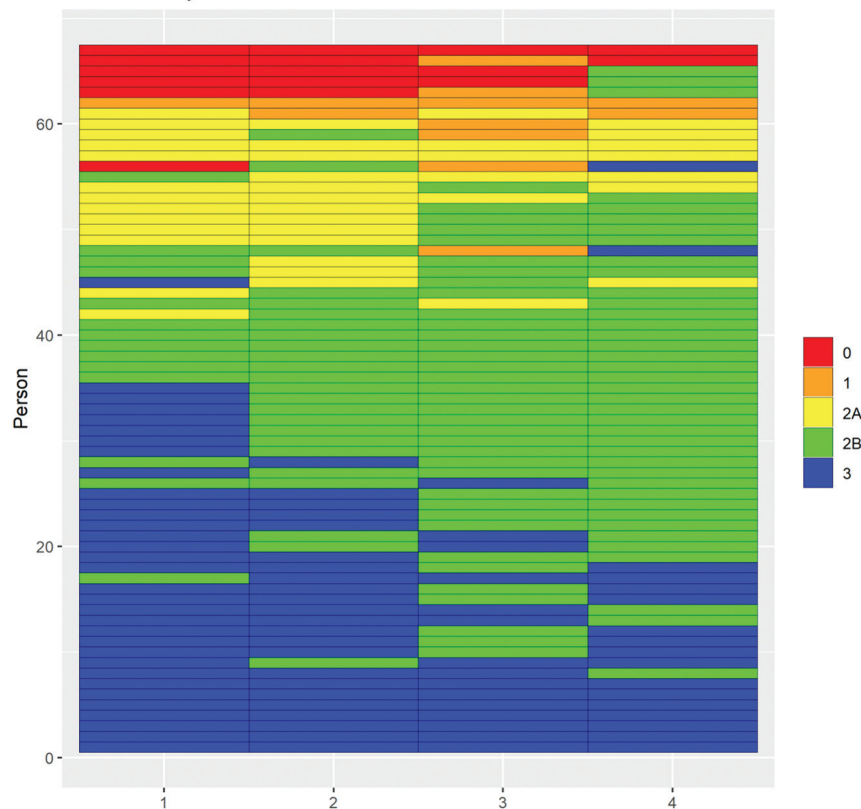


FIGURE. TICI scores for each of the 67 patients from each of the raters and both time points. Each row represents 1 patient. Column 1 refers to rater 1, column 2 refers to rater 2, and so forth. The color legend shows how each rater classified the patient during his or her respective review. For example, patient 1 was classified as a 3 by all raters, illustrating 100% agreement among all raters. Patient 20 was classified as a 3 twice during time point 1 and as a 2b twice during time point 2, illustrating only 50% agreement among raters overall at time point 1. Patient 56 had different classifications by all raters showing only 25% agreement.

“other” site of vessel occlusions. We also recorded the proceduralists’ mTICI assessments when the procedures were performed, which we will refer to as the proceduralists’ score.

Simulated Core Laboratory Survey

DSAs were de-identified and uploaded onto a secure Web site supplied by the university in compliance with all research and legal regulations. Anterior-posterior and lateral views of the initial control DSA and the final recanalization DSA were viewable on a separate page for each patient. Each case was given a score by each rater according to the mTICI score, which was provided to the raters on an introductory slide defined as

0 = no reperfusion

1 = antegrade reperfusion past the initial occlusion, but limited distal branch filling with little-or-slow distal reperfusion

2a = antegrade reperfusion of less than half of the previously occluded target artery ischemic territory

2b = antegrade reperfusion of more than half of the previously occluded target artery ischemic territory

3 = complete antegrade reperfusion of the previously occluded target artery ischemic territory, with absence of visualized occlusion in all distal branches.

Four fellowship-trained endovascular neurosurgeons with 2–18 years of experience from 3 different institutions who perform mechanical thrombectomies in their current practice participated in the simulated core laboratory case review. Raters, as we will refer to them, had access to only the DSAs and were blinded to demographics, presentation, and outcomes. No rater had performed any of the procedures included in the study. The images were placed in random order for the first review. Unknown to the raters, they were asked again 1 month later to review the same DSAs in a different random order for intrarater reliability assessment. Raters were unable to access survey answers after submission.

Statistical Methods

This study was powered to detect $\kappa = 0.80$ agreement between pair-wise raters with the null hypothesis of chance agreement of .50. A sample size of 67 resulted in 92.6% power when α was set to .05. Overall agreement was defined as the binary proportion of pair-wise scores that matched. Pair-wise inter- and intrarater reliability was assessed with percentage agreement, and weighted and unweighted Cohen κ statistics. In addition, trends of the raters’ scores compared with the proceduralists’ TICI scores were assessed

with pair-wise agreement and κ statistics. Overall interrater reliability was assessed using the overall Fleiss κ statistic.¹¹ All interrater reliability statistics were assessed at the second time point. Agreement ranged from 0 (no agreement) to 1 (perfect agreement). κ statistics ranged from < 0 to 1. Landis and Koch¹² guidelines were used to categorize levels of agreement for κ statistics: < 0 , no agreement; > 0 –0.2, slight agreement; > 0.2 –0.4, fair agreement; > 0.4 –0.6, moderate agreement; > 0.6 –0.8, substantial agreement; and > 0.8 –1, almost perfect agreement. Additional analyses compared reliability measures by occlusion location and by dichotomizing the TICI scores into poor recanalization (0, 1, 2a) and good recanalization (2b, 3). The Fisher exact test was used to compare percentage agreement measures among locations of occlusions. All analyses were conducted in SAS 9.4 (SAS Institute).

RESULTS

Of the 67 patients included in the study, 20 had ICA, 33 had M1, seven had M2, and 7 had other occlusions (3 of the vertebral/basilar artery, 1 posterior cerebral artery, 1 anterior cerebral artery, and 2 of the third-segment MCA). Pair-wise agreement between the raters ranged from 44.8% to 67.2%, the unweighted κ statistics

Table 1: Agreement by site of occlusion

	ICA (n = 20)	M1 (n = 33)	M2 (n = 7)	Other (n = 7)
No raters in agreement (No.) (%)	0 (0)	0 (0)	0 (0)	1 (14.3)
2 Raters in agreement (No.) (%)	2 (10)	10 (30.3)	2 (28.6)	2 (28.6)
3 Raters in agreement (No.) (%)	14 (70)	10 (30.3)	5 (71.4)	4 (57.1)
All raters in agreement (No.) (%)	4 (20)	13 (39.4)	0 (0)	0 (0)
Overall κ (range)	0.315 (0.128–0.502)	0.440 (0.272–0.608)	0.094 (–0.067–0.256)	0.050 (–0.14–0.241)

Fisher exact test, $P = .01$.

Table 2: Intrarater reliability for TICI scores^a

	Overall Agreement	Weighted κ	Overall κ	Dichotomous Agreement	Dichotomous κ
Rater 1 intrarater agreement	71.6 (59.3–82)	0.751 (0.64–0.863)	0.596 (0.443–0.748)	92.5 (83.4–97.5)	0.819 (0.667–0.971)
Rater 2 intrarater agreement	79.1 (67.4–88.1)	0.81 (0.706–0.914)	0.707 (0.567–0.846)	92.5 (83.4–97.5)	0.824 (0.677–0.972)
Rater 3 intrarater agreement	62.7 (50–74.2)	0.603 (0.464–0.741)	0.446 (0.284–0.607)	91 (81.5–96.6)	0.731 (0.53–0.931)
Rater 4 intrarater agreement	70.1 (57.7–80.7)	0.594 (0.435–0.752)	0.529 (0.361–0.697)	86.6 (76–93.7)	0.632 (0.423–0.84)

^a Data are calculated from data from the first and second survey.

ranged from 0.268 to 0.538, and the weighted κ statistics ranged from 0.417 to 0.703, indicating moderate-to-substantial interrater reliability. Higher weighted κ 's compared with unweighted κ 's indicated that many of the differences in agreement were small on the ordinal TICI scale. However, the overall κ for the 4 raters was 0.374 ($H_0: \kappa = 0, P < .001$), suggesting that there was only fair overall agreement (On-line Table).

The Figure illustrates each of the raters' scores for each individual patient at time 1. The patients are stacked vertically from most profusion (on average) to least profusion (on average). As shown by the weighted κ statistic, many of the disagreements were within 1 step of each other. However, there were only 17 patients (25%) for whom all raters were in complete agreement. Agreement and κ statistics were highest in the M1 and ICA occlusions and lowest in the M2 and other groups. Those with M1 occlusion had a significantly higher proportion of complete agreement (39%) followed by ICA occlusions (20%). There were no cases of complete agreement for those who had M2 or other occlusions (Table 1, $P = .010$).

Pair-wise interrater agreement for the dichotomous mTICI score ranged from 80.6% to 88.1%, and κ statistics ranged from 0.468 to 0.715, showing moderate-to-substantial agreement. The overall κ for our 4 raters for the dichotomized mTICI score was 0.582 ($H_0: \kappa = 0, P < .001$) (On-line Table). The interrater reliability improved on the second survey for both the ordinal and dichotomized analyses (0.37–0.43 and 0.58–0.65, respectively).

Intrarater reliability was higher than interrater reliability. Pair-wise intrarater agreement ranged from 62.7% to 79.1%. Unweighted κ statistics ranged from 0.446 to 0.707 (Table 2), showing moderate-to-substantial agreement, while weighted κ 's ranged from 0.594 to 0.81, showing moderate-to-almost perfect agreement. When the outcome was dichotomized, agreement ranged from 86.6% to 92.5%, with κ statistics from 0.632 to 0.824 showing substantial-to-almost perfect intrarater reliability.

When we compared the simulated core lab raters' scores with the proceduralists' mTICI scores, pair-wise agreement ranged from

52.2% to 58.2%. Unweighted κ 's ranged from 0.163 to 0.388, and weighted κ 's ranged from 0.348 to 0.574, showing fair-to-moderate agreement. For dichotomized mTICI scores, κ 's ranged from 0.365 to 0.544, showing fair-to-moderate agreement. Of the 268 comparisons (4 raters' scores each compared with 67 proceduralists' score), 138 (51.5%) were in agreement and 130 (48.5%) were not in agreement. Of the 130 not in agreement, 100 (76.9%) proceduralists' scores were higher than the raters' scores, with only 30 (23.1%) with the proceduralists' scores lower than the raters' scores.

DISCUSSION

In the era of endovascular mechanical thrombectomy for acute ischemic stroke, a number of revascularization scores have been developed and some further modified. These scores aim to quantify the success of revascularization or reperfusion and have been important reporting and prognostic metrics.

In 1992, Mori et al¹³ were the first to re-purpose the Thrombolysis in Myocardial Infarction (TIMI) scale from the cardiac literature for cerebral revascularization. These investigators broke down the partial filling grade 2 of TIMI into grades 2 (< 50% filling) and 3 (> 50% filling).¹³ Subsequently, the TICI scale was proposed by Higashida et al,¹⁴ in 2003, which focused on the revascularization assessment of territory reperfusion compared with arterial recanalization and changed the partial reperfusion grades 2 and 3 into grades 2a (less than two-thirds territory filling) and 2b (slowed-but-complete territory filling), respectively. Higashida et al reserved grade 3 for complete reperfusion.¹⁵ The Interventional Management of Stroke (IMS) II investigators established the mTICI scale, which continued to focus on reperfusion, but simplified the scale to define 2a as < 50% territory filling and 2b as > 50% filling.¹⁰ The mTICI scale has been widely used in the most recent major mechanical thrombectomy trials.^{1–9}

A further gradation of the partial perfusion grade was suggested by Noser et al,¹⁶ with 2c representing near-complete territory filling with delayed contrast runoff, which is used at

many centers but has not been universally adopted. Other revascularization scales have been proposed but have not found widespread use, including the Thrombolysis in Brain Ischemia (TIBI) scale that is based on transcranial duplex measurements, the Arterial Occlusion Lesion (AOL) scale that assesses recanalization, and the Qureshi scale that factors in the site of occlusion.^{14,15,17,18}

Our study found that raters had fair overall interrater agreement when analyzing the entire mTICI scale, which improved to moderate agreement when the responses were dichotomized to either $\leq 2a$ or $\geq 2b$. Five previous studies have researched interrater reliability of the cerebral conventional angiography revascularization score. Bar et al,¹⁹ in 2012, published the reliability of the TIMI scale applied to cerebral revascularization across 2 reviewers assessing 43 cases. They found a weighted $\kappa = 0.4$, which is nearly equivalent to our findings ($\kappa = 0.374$). Gaha et al,²⁰ in 2014, published their reliability assessment of the original TICI scale, finding an overall $\kappa = 0.45$ across 9 observers, and when dichotomized, agreement increased, with $\kappa = 0.62$.

In 2013, Suh et al²¹ looked specifically at the effect of changing from a two-thirds territory to one-half territory threshold between 2a and 2b grades of the TICI and mTICI scales, respectively. Interobserver variability was assessed as good for the TICI and mTICI scales (intraclass correlation coefficient = 0.73 and 0.67, respectively). The TICI 2a–2b threshold variability led to different grading in $\approx 20\%$ of cases, and the mTICI score (using the one-half territory threshold) was better at predicting clinical outcome. Volny et al,²² in 2017, assessed reliability using the addition of the 2c “near-complete reperfusion” to the mTICI scale. Sixty-one patients were assessed by 3 reviewers of different specialties and levels of training, who also compared different combinations of consensus grading with those of different specialties and levels of training. When compared against a criterion standard of a consensus grading between a neurointerventional fellow and attending, they found fair reliability for a stroke physician ($\kappa = 0.36$), moderate reliability for a neuroradiologist ($\kappa = 0.48$), and moderate reliability for a neurointerventional fellow ($\kappa = 0.56$). They also found that different combinations of reviewer consensus grading increased to almost perfect agreement and concluded that mTICI 2c is a feasible adjunct.²²

The most recent mTICI reliability study by Fahed et al,²³ in 2018, assessed 305 patients in the Contact Aspiration vs Stent Retriever for Successful Revascularization (ASTER) trial by 2 blinded neurointerventional radiologists, and these scores were also compared with those of the proceduralist who performed the procedure. Scores were analyzed both on an ordinal scale as well as dichotomized to $\leq 2a$ or $\geq 2b$. They found moderate agreement for the nondichotomized mTICI score and substantial agreement with the dichotomized mTICI score.²³ These findings largely mirror our findings.

We also found a trend of proceduralists’ scores being higher than independent raters’ scores, which was also found by multiple prior studies.^{23–26} For cases not in agreement, the proceduralists’ score was higher than the raters’ scores in 77% of cases. This speaks to a consistent bias that interventionalists must be aware of and attempt to overcome by rigorous objective assessment or internal blinded review.

Finally, we found that cases of M1 occlusions had the highest agreement, followed by ICA occlusions, while M2 and other vessel occlusions had poor agreement. These findings may be explained by an effort to score only the final cerebral reperfusion result, disregarding the initial perfusion findings. In fact, a posterior cerebral artery occlusion was the 1 case to be scored differently by all 4 raters. Although the scales explicitly use the initial area of the brain not receiving antegrade perfusion as the denominator to calculate the reperfusion result, it is a simpler process to always use an M1 occlusion area as the denominator by removing a step of interpretation (the determination of the initial nonperfusing brain area). Because M1 occlusions are the most common occlusion location requiring mechanical thrombectomy, many interventionalists may gravitate toward this standard consciously or subconsciously. Another explanation is that many descriptions of the mTICI scale use an M1 occlusion as an example, describing a 2b reperfusion result “eg, when greater than 50% of the MCA territory is filling.” Also, the best predictor of good outcome is a lower final infarct volume (regardless of how much brain was potentially at risk at the beginning of the case). Other explanations of this finding include branching variability of the MCA for M2 occlusion assessment, general variability of the posterior circulation, and difficulty in incorporating the anterior cerebral arteries into the scoring for ICA occlusion assessment (ie, an ICA terminus occlusion may still perfuse the ipsilateral anterior cerebral artery through the anterior communicating artery). Any of these may affect both intra- and interrater agreement. Nonocclusive thrombus may also be a challenge and account for some disagreement. Additionally, the evolution of grading scales and nomenclature may have had an effect on reliability because raters may have developed inherent biases based on timing and the institution of training. We chose to study the standard mTICI, given its widespread use throughout the large mechanical thrombectomy trials.

Our study is the first in the literature, to our knowledge, to assess intrarater reliability of the mTICI score. When raters were compared against themselves across the 2 time points, they had substantial agreement, higher than when raters were compared against each other, demonstrating a personal consistency in assessment. When answers were dichotomized to either $\leq 2a$ or $\geq 2b$, intrarater reliability rose to substantial-to-almost perfect agreement. The interrater reliability also improved on the second survey for both the ordinal and dichotomized analyses. This outcome may be due to a learning curve for the survey platform or may be evidence of a further familiarity with the mTICI system. This possibility could suggest that standardized training for mTICI, similar to what is done for clinical stroke assessment, will improve reporting consistency. However, the importance of angiogram interpretation standardization is less than that of clinical stroke assessment because an angiogram is a static record that can always be adjudicated at a later date and a patient’s clinical state is fluid with the only record being that of the clinician’s assessment on the day of examination (assuming these are not video-recorded).

Limitations from our study include heterogeneous occlusion locations with unequal representation of certain occlusion locations. Although the locations are heterogeneous and unequal, this situation occurs in practice. More education for raters by way of

case examples, explanation of rating scales from the literature and trials, and tutorials on how to deal with nuances before their formal scoring on the Web-based platform would have set the stage for more standard results. However, a simple 1-page definition of scores does allow a better real-world assessment of interventionalists' ratings, whereas heavy training would bias results away from the current state of practice and would be more of an assessment of how our training package standardizes results. We did not want to influence the scoring with our hypothesized biases.

Of the 67 individuals reviewed, reperfusion in 70.15%–83% of them was rated as 2b or 3, depending on the rater. The lack of heterogeneity in categories produces a high estimate of chance agreement, which produces a lower κ score. This is concordant with the HERMES collaboration reporting of the 5 thrombectomy trials published in 2015, which reported mTICI scores of 2b or 3 in 71% of cases. κ statistics could be artificially low, given that there was not enough representation of 0, 1, or 2a mTICI cases. Last, we compared the proceduralists' mTICI assessment with that of our raters, but we have no knowledge of what the original proceduralists' cutoffs for TICI grading were due to the retrospective nature of our study.

CONCLUSIONS

The mTICI score is a practical metric for assessing reperfusion after mechanical thrombectomy, though not without limitations. Agreement improved when scores were dichotomized around the clinically relevant threshold of successful revascularization. Interrater reliability improved with time; this feature suggests that formal training of interventionalists, similar to the design of our study, may improve reporting reliability. Agreement of the mTICI scale is highest with M1 and ICA occlusion and becomes less reliable with more distal or posterior circulation occlusions. These findings should be accounted for when developing research trials or with future modifications to stroke revascularization scores.

ACKNOWLEDGMENTS

We would like to thank Ron Price, Greg Klitz, Virgilio Conanan, and Rahul Williamson for their technical assistance with image acquisition and creation of the secure survey Web site.

Disclosures: Joseph C. Serrone—UNRELATED: Payment for Lectures Including Service on Speakers Bureaus: Chicago Review Course, Comments: I speak annually at the Chicago Review Course in Neurosurgery and receive an honorarium of <\$1000. It does not conflict with this article. Koji C. Ebersole—UNRELATED: Consultancy: Stryker Neurovascular, Comments: physician proctor and medical consultant. Andrew J. Ringer—UNRELATED: Consultancy: MicroVention, Comments: no payments in >2 years; Expert Testimony: Mayfield Brain & Spine, Comments: none relevant to this work.

REFERENCES

- Goyal M, Menon BK, van Zwam WH, et al; HERMES Collaborators. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet* 2016;387:1723–31 [CrossRef Medline](#)
- Nogueira RG, Jadhav AP, Haussen DC, et al; DAWN Trial Investigators. Thrombectomy 6 to 24 hours after stroke with a mismatch between deficit and infarct. *N Engl J Med* 2018;378:11–21 [CrossRef Medline](#)
- Albers GW, Marks MP, Kemp S, et al; DEFUSE 3 Investigators. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. *N Engl J Med* 2018;378:708–18 [CrossRef Medline](#)
- Berkhemer OA, Fransen PS, Beumer D, et al; MR CLEAN Investigators. A randomized trial of intraarterial treatment for acute ischemic stroke. *N Engl J Med* 2015;372:11–20 [CrossRef Medline](#)
- Goyal M, Demchuk AM, Menon BK, et al; ESCAPE Trial Investigators. Randomized assessment of rapid endovascular treatment of ischemic stroke. *N Engl J Med* 2015;372:1019–30 [CrossRef Medline](#)
- Campbell BC, Mitchell PJ, Kleinig TJ, et al; EXTEND-IA Investigators. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *N Engl J Med* 2015;372:1009–18 [CrossRef Medline](#)
- Saver JL, Goyal M, Bonafe A, et al; SWIFT PRIME Investigators. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. *N Engl J Med* 2015;372:2285–95 [CrossRef Medline](#)
- Jovin TG, Chamorro A, Cobo E, et al; REVASCAT Trial investigators. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *N Engl J Med* 2015;372:2296–2306 [CrossRef Medline](#)
- Bracard S, Ducrocq X, Mas JL, et al; THRACE Investigators. Mechanical thrombectomy after intravenous alteplase versus alteplase alone after stroke (THRACE): a randomised controlled trial. *Lancet Neurol* 2016;15:1138–47 [CrossRef Medline](#)
- Tomsick T, Broderick J, Carrozella J, et al; Interventional Management of Stroke II Investigators. Revascularization results in the Interventional Management of Stroke II trial. *AJNR Am J Neuroradiol* 2008;29:582–87 [CrossRef Medline](#)
- Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions*. Wiley; 2003
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74 [Medline](#)
- Mori E, Yoneda Y, Tabuchi M, et al. Intravenous recombinant tissue plasminogen activator in acute carotid artery territory stroke. *Neurology* 1992;42:976–82 [CrossRef Medline](#)
- Higashida RT, Furlan AJ, Roberts H, et al; Technology Assessment Committee of the American Society of Interventional and Therapeutic Neuroradiology; Technology Assessment Committee of the Society of Interventional Radiology. Trial design and reporting standards for intra-arterial cerebral thrombolysis for acute ischemic stroke. *Stroke* 2003;34:e109–37 [CrossRef Medline](#)
- Khatri P, Neff J, Broderick JP, et al; IMS-I Investigators. Revascularization end points in stroke interventional trials: recanalization versus reperfusion in IMS-I. *Stroke* 2005;36:2400–03 [CrossRef Medline](#)
- Noser EA, Shaltoni HM, Hall CE, et al. Aggressive mechanical clot disruption: a safe adjunct to thrombolytic therapy in acute stroke? *Stroke* 2005;36:292–96 [CrossRef Medline](#)
- Demchuk AM, Burgin WS, Christou I, et al. Thrombolysis in brain ischemia (TIBI) transcranial Doppler flow grades predict clinical severity, early recovery, and mortality in patients treated with intravenous tissue plasminogen activator. *Stroke* 2001;32:89–93 [CrossRef Medline](#)
- Qureshi AI. New grading system for angiographic evaluation of arterial occlusions and recanalization response to intra-arterial thrombolysis in acute ischemic stroke. *Neurosurgery* 2002;50:1405–14; discussion 1414–15 [Medline](#)
- Bar M, Mikulik R, Jonszta T, et al. Diagnosis of recanalization of the intracranial artery has poor inter-rater reliability. *AJNR Am J Neuroradiol* 2012;33:972–74 [CrossRef Medline](#)
- Gaha M, Roy C, Estrade L, et al. Inter- and intraobserver agreement in scoring angiographic results of intra-arterial stroke therapy. *AJNR Am J Neuroradiol* 2014;35:1163–69 [CrossRef Medline](#)
- Suh SH, Cloft HJ, Fugate JE, et al. Clarifying differences among thrombolysis in cerebral infarction scale variants: is the artery half open or half closed? *Stroke* 2013;44:1166–68 [CrossRef Medline](#)
- Volny O, Cimflova P, Szeder V. Inter-rater reliability for thrombolysis in cerebral infarction with TICI 2c category. *J Stroke Cerebrovasc Dis* 2017;26:992–94 [CrossRef Medline](#)

23. Fahed R, Ben Maacha M, Ducroux C, et al; ASTER Trial Investigators. **Agreement between core laboratory and study investigators for imaging scores in a thrombectomy trial.** *J Neurointerv Surg* 2018;10:e30 [CrossRef Medline](#)
24. Saver JL, Jahan R, Levy EI, et al; SWIFT Trialists. **Solitaire flow restoration device versus the Merci Retriever in patients with acute ischaemic stroke (SWIFT): a randomised, parallel-group, non-inferiority trial.** *Lancet* 2012;380:1241–49 [CrossRef Medline](#)
25. Nogueira RG, Lutsep HL, Gupta R, et al; TREVO 2 Trialists. **Trevo versus Merci retrievers for thrombectomy revascularisation of large vessel occlusions in acute ischaemic stroke (TREVO 2): a randomised trial.** *Lancet* 2012;380:1231–40 [CrossRef Medline](#)
26. Zhang G, Treurniet KM, Jansen IG, et al; for the MR CLEAN Registry Investigators. **Operator versus core lab adjudication of reperfusion after endovascular treatment of acute ischemic stroke.** *Stroke* 2018;49:2376–82 [CrossRef Medline](#)