Original research

Early neurological improvement as a predictor of outcomes after endovascular thrombectomy for stroke: a systematic review and meta-analysis

Hassan Kobeissi 1, Sherief Ghozy 1, Cem Bilgin, 2 Ramanathan Kadirvel 1, 2 David F Kallmes 2

ABSTRACT

Background Early neurological improvement (ENI) is a potential predictor for 90-day outcomes following mechanical thrombectomy for acute ischemic stroke (AIS). We performed a systematic review and meta-analysis to better understand whether ENI can be used as a surrogate for long-term outcomes following mechanical thrombectomy for AIS.

Methods Following the PRISMA guidelines, a systematic literature review of the English language literature was conducted using PubMed, MEDLINE, and Embase. ENI definition, including timing and degree of improvement on the National Institutes of Health Stroke Scale (NIHSS), was catalogued for each included study. Outcomes of interest included 90-day modified Rankin Scale (mRS) 0–2, symptomatic intracranial hemorrhage (sICH), and mortality. We calculated pooled ORs and their corresponding 95% confidence intervals (CI) for all definitions of ENI.

Results We included nine studies with 2355 patients in our analysis. ENI definitions included improvement in NIHSS of 8 points, 4 points, 12%, and 30% or greater. There was a significant association between ENI and mRS 0–2 rates (OR 8.62, 95% CI 4.86 to 15.29; p<0.001). Significance of the association was maintained across all definitions (p<0.001). Moreover, achieving ENI was a significant predictor of reduced odds for reported sICH rates (OR 0.11, 95% CI 0.06 to 0.21; p<0.001). There was a significant association between ENI and reduction in mortality rates (OR 0.09, 95% CI 0.05 to 0.15; p<0.001).

Conclusions Broadly defined, ENI is a promising predictor of good functional outcome at 90 days and is associated with lower rates of mortality and sICH.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Early neurological improvement has thus far been used to assess the clinical improvement in patients 24 hours after stroke treatment.

WHAT THIS STUDY ADDS

⇒ This study examines the literature on early neurological improvement and its predictive value for long-term outcomes.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE AND/OR POLICY

⇒ Our systematic review and meta-analysis indicates that early neurological improvement is a promising predictor for 90-day outcomes following mechanical thrombectomy for acute ischemic stroke.
⇒ Our study also compares different definitions of early neurological improvement and highlights the need for a consistent definition of the term.

INTRODUCTION

Stroke is a widespread medical event that impacts patients’ lives, often resulting in reduced independence. Assessing outcomes following treatment for acute ischemic stroke (AIS) traditionally includes a multitude of measurements. A good outcome is commonly defined as a modified Rankin Scale (mRS) score of 0–2 at 90 days post-treatment. The mRS is typically measured at 90 days after treatment and is used to assess the degree of disability in patients who have suffered a stroke. The National Institutes of Health Stroke Scale (NIHSS) is a parameter which can be measured sooner than 90 days post-treatment, most commonly at admission and after 24 hours as a deficit rating scale, and can be used as a predictor of functional outcome and mortality.

Early neurological improvement (ENI) has been used as a 24-hour measurement to assess a change in NIHSS score. ENI has been defined in several different ways, most commonly as an integer improvement in the NIHSS score after 24 hours of 8, but also as a percentage change in the NIHSS score and as improvement of 4, 6, and 10 in the NIHSS score after 24 hours. If it can be validated as a reliable early predictor of long-term outcome, ENI will be a valuable tool for informing patients and their families of their prognosis in a timely manner. To our knowledge, a meta-analysis and systematic review of the literature addressing ENI as a predictor of 90-day outcomes has not yet been performed.

To assess whether ENI can be used as a reliable predictor of good functional outcome following mechanical thrombectomy treatment for AIS, we performed a systematic review and meta-analysis of studies that reported on ENI and 90-day functional outcome, including a comparison of different definitions of ENI.
Ischemic stroke

METHODS

Search strategy

On 28 January 2022, following the PRISMA guidelines for performing systematic reviews, a systematic literature review of the English language literature was conducted within the Nested Knowledge Autotlit software per the drafted protocol from inception, using PubMed, Embase, Web of Science, and Scopus. Based on each database, different combinations of possible keywords and/or MeSH terms were used for that purpose. Keywords and MeSH terms included: "stroke", "cerebral infarction", "thrombectomy", "NIHSS", "National Institutes of Health Stroke Scale", "early", "neurological improvement", "neurological outcome", "functional outcome", "functional independence", "clinical outcome", "clinical improvement". Moreover, we did an extensive manual search through the references of the included articles to retrieve any missed papers.

Screening process

We included all original studies fulfilling our pre-determined PICO: Population was patients with AIS with reported ENI. Intervention was thrombectomy, there was no Control group, the Outcomes of interest were the modified Rankin Scale (mRS) 0–2, mortality, and symptomatic intracerebral hemorrhage (sICH). We excluded papers where the ENI was not described, review articles, duplicate studies including the same patients presented in another included paper, case reports, case series with fewer than five patients, and conference abstracts. We did not pose any limitations regarding sample size, study design, or patients' characteristics.

Two authors carried out the title and abstract screening against the pre-defined criteria. This was followed by a full text screening of any retained studies of the first screening step. In both stages, the senior author was consulted to resolve any conflicts in the decisions.

Data extraction

Following a pilot extraction, an extraction sheet was built and the extraction was performed by at least two authors. The extracted data included study characteristics, baseline data of the included patients, and the aforementioned outcomes of interest. After performing the extraction, a third author carried out an extensive revision of the extracted data to avoid any prior mistakes.

Risk of bias

We assessed the risk of bias using ROBINS-I, a tool for assessing risk of bias in non-randomised studies of interventions. Two authors evaluated the quality of each study, which was adjudicated by a third author when needed.

Statistical analysis

Using R software version 4.1.2, we calculated pooled ORs and their corresponding 95% CIs. A random model was adopted to pool all data due to heterogeneity among the included studies in defining ENI. Heterogeneity was assessed using Q statistics and the I² test, where I² > 50% or p < 0.05 were considered significant. Among the performed analysis, the number of studies was < 10 so publication bias (Egger’s regression test) and the impact of sample size (meta-regression) were not tested. Since the included studies used different definitions of ENI, we used subgroup analyses to investigate different definitions and test for statistical differences among them.

RESULTS

Search and screening results

Following the removal of 1586 duplicate records, we retrieved 800 papers for further screening. We excluded 776 records at the title and abstract screening stage, thus retaining 24 records for full-text screening. Finally, nine papers were determined to satisfy our inclusion criteria with the appropriate report of outcomes of interest (see online supplemental figure S1).

Study characteristics and risk of bias

Of the nine included studies, six used a retrospective design and five were multicenter studies. The sample size of the included studies ranged from 91 individuals to 568, with reported outcomes at 3 months of follow-up for all studies. The median NIHSS score at baseline and patients’ characteristics/comorbidities are shown in table 1. Five papers included in the meta-analysis defined ENI as a reduction in NIHSS score of 8 or more, Weyland et al defined ENI as any improvement in the NIHSS score in 24 hours, Soize et al defined ENI as a reduction in NIHSS score of 4 or more, Pu et al defined ENI as a 12% or greater reduction in NIHSS score, and Cao et al defined ENI as a 30% or greater improvement in NIHSS score. Definitions of both ENI and sICH are shown in online supplemental table S1. All definitions of ENI focused on improvement of score from baseline to 24 hours after baseline.

For all included studies there was no high risk of bias among all assessed domains. However, certain specific bias risk was identified in specific study aspects. A detailed assessment of risk of bias is shown in online supplemental table S2.

Functional independence (mRS 0–2)

Nine studies with 2355 patients reported the functional independence rates in relation to ENI. Overall, there was a significant association between ENI and mRS 0–2 rates (OR 8.62, 95% CI 4.86 to 15.29; p < 0.001); however, a significant heterogeneity was observed among the included studies (I² = 84%, p < 0.001). On further stratification of studies based on their definition of ENI, the significance of the association was maintained across all definitions (p < 0.001; figure 1). Moreover, the residual heterogeneity dropped to an insignificant level between studies (I² = 37%, p = 0.171), confirming the different definition as the main source of the overall heterogeneity. Nevertheless, the odds of achieving functional independence were significantly variable among different ENI subgroups/definitions, as shown by the test for subgroup differences (p < 0.001).

Incidence of sICH

Three studies of 1208 patients reported the incidence rates of sICH as associated with ENI. Overall, having ENI was a significant predictor of reduced odds for the reported sICH rates (OR 0.11, 95% CI 0.06 to 0.21; p < 0.001), with no heterogeneity among the included studies (I² = 0%, p = 0.694). This significance was present across different ENI definitions, except for ≥ 8 NIHSS points (OR 0.16, 95% CI 0.01 to 2.99; p = 0.220); however, there was no significant variability across different ENI subgroups (p = 0.694; figure 2).

Mortality rates

Five studies with 1161 patients reported the association between ENI and subsequent mortality. There was a significant association between ENI and the reduction in mortality rates (OR 0.09, 95% CI 0.05 to 0.15; p < 0.001), with no significant heterogeneity among the included studies (I² = 18%, p = 0.302). This
In this meta-analysis and systematic review we have shown that ENI of various definitions is correlated with 90-day outcomes following thrombectomy for AIS. These findings are important because it will allow clinicians to inform patients and their families about the likelihood of long-term outcomes without the need to wait the full 90 days. Additionally, we found that patients who achieve ENI had a significantly lower rate of mortality at 90 days and lower rates of sICH compared with those who did not achieve ENI. This association between ENI and good outcomes is likely due to the fact that patients who demonstrate a significant improvement in the NIHSS score at 24 hours are poised to improve further until the clinical course for recovery from their stroke stabilizes at day 4.4-22

We found that, although there were multiple definitions of ENI, all definitions were associated with good functional outcomes at 90 days. This holds true for safety outcomes (mortality, sICH), except in the study by Heit et al in which no significant association between sICH and ENI was found. Using retrospective analyses, previous authors have found that ENI is significantly correlated with 90-day outcomes.6,23-24 These previous studies, however, had small patient cohorts, making it difficult to extrapolate the results to broader patient populations. However, even though previous studies have had small cohorts, the association between ENI and good outcomes has been consistent with the present study.8 Our study adds to the current literature by increasing the power of these studies with a larger cohort of 2355 patients in the form of a systematic review and meta-analysis. With this larger patient population, we achieved results consistent with the smaller studies that were analyzed. Additionally, our study examined the value of using different definitions of ENI, with results showing that all definitions of ENI were significantly correlated with good functional and safety outcomes.

Our meta-analysis has a number of limitations. The major limitation that must be noted is the heterogeneity in the definitions of ENI. Most of the studies analyzed did not agree on a definition of ENI, and definitions ranged from any improvement in NIHSS score at 24 hours to an 8-point improvement in NIHSS score at 24 hours. Other studies used percentage thresholds ranging from 12% to 30% improvement in NIHSS at 24 hours. This means that we were not able to pool together these studies under a single definition of ENI, thereby decreasing the power of our study. Similarly, NIHSS scores trend towards being higher for left hemisphere strokes than for right hemisphere strokes, and this bias must be considered when using the NIHSS score.35

Another limitation which must be noted is that we did not have access to patient-level data. It has been shown that underlying pathologies such as increased blood pressure and increased mean glucose levels, and varying treatment modalities such as pretreatment with intravenous thrombolysis are associated with outcomes.26-28 We were unable to control for these factors since we did not have access to data that detailed individual patient characteristics and treatment. Furthermore, there were several studies which would have otherwise been eligible for our analysis but did not include patient data. Instead, these studies only reported on the analysis of their data without providing the raw data for our analysis. Despite these limitations, we were able through pooling the data for our analysis. Despite these limitations, we were able to achieve results consistent with the smaller studies that were analyzed.

We acknowledge that our study did not stratify characteristics by ENI status; relevant characteristics for the entire study are listed in the “ENI” section. We found that, although there were multiple definitions of ENI, all definitions were associated with good functional outcomes at 90 days. This holds true for safety outcomes (mortality, sICH), except in the study by Heit et al in which no significant association between sICH and ENI was found. Using retrospective analyses, previous authors have found that ENI is significantly correlated with 90-day outcomes.6,23-24 These previous studies, however, had small patient cohorts, making it difficult to extrapolate the results to broader patient populations. However, even though previous studies have had small cohorts, the association between ENI and good outcomes has been consistent with the present study.8 Our study adds to the current literature by increasing the power of these studies with a larger cohort of 2355 patients in the form of a systematic review and meta-analysis. With this larger patient population, we achieved results consistent with the smaller studies that were analyzed. Additionally, our study examined the value of using different definitions of ENI, with results showing that all definitions of ENI were significantly correlated with good functional and safety outcomes.
Ischemic stroke

to obtain statistically significant results with low heterogeneity between studies.

It is important to note that integer changes in NIHSS scores at 24 hours have their own limitations. For example, a patient who has a 5-point improvement in the NIHSS score from 15 to 10 (33%) may fare worse than a patient who has a 4-point improvement in the NIHSS score from 7 to 3 (57%). Though the former 5-point improvement had a greater absolute change

<table>
<thead>
<tr>
<th>Study</th>
<th>Events</th>
<th>ENI Total</th>
<th>Events</th>
<th>No ENI Total</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95% CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 or more NIHSS points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guenego et al., 2021</td>
<td>51</td>
<td>70</td>
<td>40</td>
<td>167</td>
<td>8.52</td>
<td>4.51</td>
<td>16.09</td>
<td>11.6%</td>
</tr>
<tr>
<td>Sun et al., 2021</td>
<td>34</td>
<td>62</td>
<td>34</td>
<td>125</td>
<td>3.25</td>
<td>1.72</td>
<td>6.14</td>
<td>11.6%</td>
</tr>
<tr>
<td>Wirtz et al., 2019</td>
<td>23</td>
<td>52</td>
<td>14</td>
<td>103</td>
<td>5.04</td>
<td>2.30</td>
<td>11.06</td>
<td>10.8%</td>
</tr>
<tr>
<td>Heit et al., 2019</td>
<td>19</td>
<td>31</td>
<td>22</td>
<td>60</td>
<td>2.73</td>
<td>1.12</td>
<td>6.68</td>
<td>10.2%</td>
</tr>
<tr>
<td>Cai et al., 2021</td>
<td>38</td>
<td>50</td>
<td>62</td>
<td>177</td>
<td>5.87</td>
<td>2.86</td>
<td>12.06</td>
<td>11.2%</td>
</tr>
<tr>
<td>Random effects model</td>
<td>165</td>
<td>265</td>
<td>172</td>
<td>632</td>
<td>4.84</td>
<td>3.21</td>
<td>7.39</td>
<td>55.5%</td>
</tr>
</tbody>
</table>

Heterogeneity: $I^2 = 37\%$, $r^2 = 0.081$, $p = 0.171$
Test for effect in subgroup: $z = 7.554$ ($p < 0.001$)

Any ENI

<table>
<thead>
<tr>
<th>Study</th>
<th>Events</th>
<th>ENI Total</th>
<th>Events</th>
<th>No ENI Total</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95% CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weyland et al., 2021</td>
<td>229</td>
<td>434</td>
<td>13</td>
<td>115</td>
<td>8.76</td>
<td>4.78</td>
<td>16.09</td>
<td>11.8%</td>
</tr>
<tr>
<td>Random effects model</td>
<td>229</td>
<td>434</td>
<td>13</td>
<td>115</td>
<td>8.76</td>
<td>4.78</td>
<td>16.09</td>
<td>11.8%</td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable
Test for effect in subgroup: $z = 7.007$ ($p < 0.001$)

4 or more NIHSS points

<table>
<thead>
<tr>
<th>Study</th>
<th>Events</th>
<th>ENI Total</th>
<th>Events</th>
<th>No ENI Total</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95% CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Soize et al., 2019</td>
<td>106</td>
<td>128</td>
<td>7</td>
<td>116</td>
<td>75.03</td>
<td>30.76</td>
<td>182.98</td>
<td>10.2%</td>
</tr>
<tr>
<td>Random effects model</td>
<td>106</td>
<td>128</td>
<td>7</td>
<td>116</td>
<td>75.03</td>
<td>30.76</td>
<td>182.98</td>
<td>10.2%</td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable
Test for effect in subgroup: $z = 9.492$ ($p < 0.001$)

12% improvement in NIHSS

<table>
<thead>
<tr>
<th>Study</th>
<th>Events</th>
<th>ENI Total</th>
<th>Events</th>
<th>No ENI Total</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95% CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pu et al., 2018</td>
<td>231</td>
<td>284</td>
<td>59</td>
<td>264</td>
<td>16.62</td>
<td>10.99</td>
<td>25.14</td>
<td>12.7%</td>
</tr>
<tr>
<td>Random effects model</td>
<td>231</td>
<td>284</td>
<td>59</td>
<td>264</td>
<td>16.62</td>
<td>10.99</td>
<td>25.14</td>
<td>12.7%</td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable
Test for effect in subgroup: $z = 13.310$ ($p < 0.001$)

30% improvement in NIHSS

<table>
<thead>
<tr>
<th>Study</th>
<th>Events</th>
<th>ENI Total</th>
<th>Events</th>
<th>No ENI Total</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95% CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cao et al., 2017</td>
<td>30</td>
<td>42</td>
<td>10</td>
<td>55</td>
<td>11.25</td>
<td>4.32</td>
<td>29.32</td>
<td>9.8%</td>
</tr>
<tr>
<td>Random effects model</td>
<td>30</td>
<td>42</td>
<td>10</td>
<td>55</td>
<td>11.25</td>
<td>4.32</td>
<td>29.32</td>
<td>9.8%</td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable
Test for effect in subgroup: $z = 4.952$ ($p < 0.001$)

12% improvement in NIHSS

<table>
<thead>
<tr>
<th>Study</th>
<th>Events</th>
<th>ENI Total</th>
<th>Events</th>
<th>No ENI Total</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95% CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pu et al., 2018</td>
<td>8</td>
<td>284</td>
<td>52</td>
<td>284</td>
<td>0.13</td>
<td>0.06</td>
<td>0.28</td>
<td>65.6%</td>
</tr>
<tr>
<td>Random effects model</td>
<td>8</td>
<td>284</td>
<td>52</td>
<td>284</td>
<td>0.13</td>
<td>0.06</td>
<td>0.28</td>
<td>65.6%</td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable
Test for effect in subgroup: $z = -5.244$ ($p < 0.001$)

8 or more NIHSS points

<table>
<thead>
<tr>
<th>Study</th>
<th>Events</th>
<th>ENI Total</th>
<th>Events</th>
<th>No ENI Total</th>
<th>Odds Ratio</th>
<th>OR</th>
<th>95% CI</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heit et al., 2019</td>
<td>0</td>
<td>31</td>
<td>5</td>
<td>60</td>
<td>0.16</td>
<td>0.01</td>
<td>2.99</td>
<td>4.5%</td>
</tr>
<tr>
<td>Random effects model</td>
<td>0</td>
<td>31</td>
<td>5</td>
<td>60</td>
<td>0.16</td>
<td>0.01</td>
<td>2.99</td>
<td>4.5%</td>
</tr>
</tbody>
</table>

Heterogeneity: not applicable
Test for effect in subgroup: $z = -1.226$ ($p = 0.220$)

Random effects model

Figure 1  Forest plot of modified Rankin Scale score (mRS) of 0–2 at 90 days. Results were stratified by definition of early neurological improvement (ENI).

Figure 2  Forest plot of symptomatic intracranial hemorrhage (sICH). Results were stratified by definition of early neurological improvement (ENI).

NIHSS, National Institutes of Health Stroke Scale.
in the NIHSS score, the latter may have better outcomes due to both their final NIHSS score and greater percentage change.\textsuperscript{29} Future papers could assess early NIHSS as a percentage versus an integer in order to determine the optimal threshold for ENI as it correlates with functional outcomes, as current literature suggests the use of an absolute 24-hour NIHSS at presentation and final NIHSS at 24 hours likely correlate with good outcomes following treatment. Based on existing literature, it is difficult to ascertain whether these NIHSS values hold a better predictive value than ENI.

Our study, while important, demonstrates the need for a unified definition of ENI. This would allow a higher-powered analysis of studies to be carried out to better assess the predictive value of ENI as a surrogate for long-term outcomes. As more data regarding ENI and long-term outcomes are published, clinicians will be able to better inform patients and their families about their prognosis following thrombectomy treatment for AIS.

CONCLUSIONS

In this meta-analysis of a systematic review of nine studies, ENI was shown to be a promising predictor of long-term functional outcomes in patients treated with mechanical thrombectomy for AIS. Additionally, ENI was found to be significantly associated with lower rates of sICH and mortality. Future prospective studies are needed to further validate the association and to investigate if it is clinically meaningful.

Acknowledgements The authors acknowledge Karl Holub, Stephen Mead, Jeffrey Johnson, and Darian Lehmann-Plantenberg for their design, development, and support of the Nested Knowledge meta-analytical software.

Contributors All authors contributed to the manuscript. HK, SG, RK, and DFK were responsible for the conception and design of the work. HK, SG, and CB were responsible for collection of the data. All authors were involved in the drafting of the article, critical revision of the article, and final approval of the version to be published. HK is responsible for the overall content as guarantor.

Funding Research reported in this publication was supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under Award Number R01NS076491.

Disclaimer The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests DFK holds equity in Nested Knowledge, Superior Medical Editors, and Convex Medical, Marblehead Medical; is a consultant for MicroVention, Medtronic, Balt, and Insera Therapeutics; Data Safety Monitoring Board for Vesalio; and receives royalties from Medtronic.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The data that support the findings of this study are available from the corresponding author upon reasonable request.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not be have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

ORCID iDs
Hassan Kobeissi http://orcid.org/0000-0001-7446-624X
Sherief Ghozy http://orcid.org/0000-0001-5629-3023
Ramanathan Kadirvel http://orcid.org/0000-0002-6786-9953

REFERENCES

2 Weischer N, Vermeulen M, Roos YB, et al. What should be defined as good outcome in stroke trials; a modified Rankin score of 0-1 or 0-2? Neurology 2008;255:867–74.