Minimal Clinically Important Difference for Safe and Simple Novel Acute Ischemic Stroke Therapies

Jessica S. Cranston, BS; Brett D. Kaplan, BS; Jeffrey L. Saver, MD

**Background and Purpose**—Determining the minimal clinically important difference (MCID) is essential for evaluating novel therapies. For acute ischemic stroke, expert surveys have yielded MCIDs that are substantially higher than the MCIDs observed in actual expert behavior in guideline writing and clinical practice, potentially because of anchoring bias.

**Methods**—We administered a structured, internet-based survey to a cross-section of academic stroke neurologists in the United States. Survey responses assessed demographic and clinical experience, and expert judgment of the MCID of the absolute increase needed in the proportion of patients achieving functional independence at 3 months to consider a novel, safe neuroprotective agent as clinically worthwhile. To mitigate anchoring bias, the survey response framework used a base 1000 rather than base 100 patient framework.

**Results**—Survey responses were received from 122 of 333 academic stroke neurologists, there were 23% women, 72.8% ≥6 years of practice experience, and neurovascular disease accounted for more than half of practice time in >70%. Responder–nonresponder and continuum of resistance tests indicated that responders were representative of the full expert population. Among respondents, the median MCID was 1.3% (interquartile range, 0.8% to >2%).

**Conclusions**—Stroke expert responses to MCID surveys are affected by anchoring and centrality bias. When survey design takes these into account, the expert-derived MCID for a safe acute ischemic stroke treatment is 1.1% to 1.5%, in accord with actual physician behavior in guideline writing and clinical practice. This revised MCID value can guide clinical trial design and grant-funding and regulatory agency decisions. (*Stroke*. 2017;48:2946-2951. DOI: 10.1161/STROKEAHA.117.017496.)

Key Words: device approval ■ sample size ■ stroke ■ treatment outcome

The minimal clinically important difference (MCID) is the smallest change in a treatment outcome that a patient, a care provider, or both would consider worthwhile.

Establishing the MCID for a disease state is an essential prerequisite for clinical trial sample size calculation and informs funding decisions by the National Institutes of Health and other sponsors and drug or device approval decisions by the Food and Drug Administration and other regulatory agencies. For superiority trials, declarations that a novel treatment is clinically superior to standard therapy require that the improved outcome rate exceeds the MCID. For equivalence and noninferiority trials, declarations that 2 treatments are of equal clinical efficacy require that their outcome rates fall within the MCID. The smaller the MCID, the larger the sample size needed for a randomized trial to ensure that the study is adequately powered to detect or exclude a treatment benefit of clinical relevance.

Approaches to establishing the MCID for a particular disease or symptom fall into 3 categories: distribution-based, anchor-based, and Delphi expert-based approaches. Distribution-based approaches statistically derive an MCID from the distribution of outcome data, such as using one half the SD of an end point. They have the advantage of direct calculation from outcome data sets, but the drawback of not clearly correlating with clinically important change.

Anchor-based approaches compare change in end point scores with an external anchor, most commonly a patient global impression of change. Patient judgments that a change has been meaningful are relatively straightforward to elicit for treatments that are applied to patients with a previously stable disease-related health state. When interventions move patients from one to another long-lasting disease–related health state, patients can draw on their personal experience of both the before and the after states to render assessments comparing the 2. However, patient judgments that a change has been meaningful are challenging to derive for treatments that are applied to patients with an abrupt onset new condition, such as acute ischemic stroke. With acute onset conditions, patients can draw on their personal experience of only 1 stable disease-related health state, their own final outcome, and cannot compare this state to any alternative, personally experienced outcome state.
Given the limitations of the distribution- and anchor-based approaches for acute stroke, the survey methods have been the leading technique for determining what is a minimally important change in stroke outcomes. Surveys are administered to physicians, nurses, and other healthcare providers about the worth of different outcome states. Because healthcare providers have direct observational familiarity with a range of stroke outcomes, they are able to knowledgeably make comparative judgments of the value of alternative disease-related health states. But, for simple and safe therapies for acute ischemic stroke, MCIDs derived from expert judgment (5%–10%) have been higher than MCIDs derived from econometric modeling (2%–3%) and higher than MCIDs derived from observations of actual physician behavior and medical guidelines (1%–1.5%). These elevated expert-derived MICD values have been highlighted in European and American consensus statements on acute ischemic stroke clinical trial design.

Recent studies of both nonexperts and experts have shown that human judgment in a wide variety of settings is prone to cognitive biases—systematic deviations from rationality. Notably, the architecture of questions used to elicit expert opinion may bias the resulting responses. Investigations of expert judgment of acute stroke MCID have generally used multiple choice questions, which are subject to anchoring and centrality bias. We sought to determine whether altering the question anchor framework would yield expert-derived MCIDs for acute stroke that better accord with actual expert behavior.

Methods

We designed a survey response framework that would have a base 1000 rather than base 100 framework, to avoid skewing respondent answers toward values in the 2% to 20% range. The survey consisted of 5 questions on a single web page. The first 4 questions elicited information on respondent characteristics, including appointment level, years of clinical practice, board certification, and percent practice devoted to stroke patients. The fifth question addressed the MCID using the below scenario and response options:

In an acute stroke trial, patients are randomized to a novel neuroprotective agent or placebo. The active drug is demonstrated to be safe, without any side effects, and to improve the number of patients who achieve freedom from disability (modified Rankin Score 0–2) at 90 days after stroke. How many patients, per 1000 treated, would need to benefit (by achieving freedom from disability at 90 days) for this neuroprotective drug to be worthwhile to use in clinical practice?

a. 1
b. 2–5
c. 6–10
d. 11–15
e. 16–20
f. >20

The survey was administered to university-affiliated neurologists in the United States who specialize in the care of stroke patients. A systematic effort was made to identify all faculty stroke neurologists at all academic medical centers in the United States. The websites of all medical schools in the United States were searched for stroke neurology physicians, and all individuals identified were sent e-mail invitations to participate in the survey. After the initial survey invitation, 2 subsequent reminders were sent to initial nonresponders. Participants were assured that all survey responses would be kept fully confidential. A total of 335 academically affiliated stroke neurologists were identified, among whom 333 had e-mail addresses that were functional and were included in the target population.

Statistical Analysis

Survey response data were analyzed using univariate statistical methods. Two approaches were taken to evaluate the potential for response bias. For demographic physician variables for which data on both respondents and nonrespondents were publicly available by web search (sex and geographic location), responder and nonresponder groups were directly compared. For demographic physician variables for which data on nonresponders groups were not publicly available, and for expert MCID judgments, the potential for responder bias was evaluated using the continuum of resistance model. The survey recipients were divided into those who responded to the first survey request received (early responders) and those who responded only after repeated survey requests (late responders). The underlying assumption behind this approach is that every subject in the study population has a position on the response continuum that ranges from will always respond to will never respond. Late responders, who would have been categorized as nonresponders if the study had been stopped earlier, are expected to more closely resemble the nonresponders than the early responders. Consequently, comparisons of the characteristics of the early versus late responder groups will provide an estimate of the potential for responder bias arising from the nonresponders. Statistical significance was estimated by χ² tests. A P value ≤0.05 was considered significant.

Table 1. Characteristics of Physician Survey Participants

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Number, % (n=122)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, female</td>
<td>28 (23.0)</td>
</tr>
<tr>
<td>Geographic location</td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>41 (33.6)</td>
</tr>
<tr>
<td>Midwest</td>
<td>26 (21.3)</td>
</tr>
<tr>
<td>South</td>
<td>32 (26.2)</td>
</tr>
<tr>
<td>West</td>
<td>23 (18.9)</td>
</tr>
<tr>
<td>Years of practice</td>
<td></td>
</tr>
<tr>
<td>0–5</td>
<td>36 (29.5)</td>
</tr>
<tr>
<td>6–10</td>
<td>20 (16.4)</td>
</tr>
<tr>
<td>11–15</td>
<td>17 (13.9)</td>
</tr>
<tr>
<td>&gt;15</td>
<td>49 (40.2)</td>
</tr>
<tr>
<td>Board certification, yes</td>
<td>96 (78.7)</td>
</tr>
<tr>
<td>Current appointment level</td>
<td></td>
</tr>
<tr>
<td>Clinical Instructor</td>
<td>6 (5.0)</td>
</tr>
<tr>
<td>Junior faculty</td>
<td>40 (33.1)</td>
</tr>
<tr>
<td>Midcareer</td>
<td>29 (24.0)</td>
</tr>
<tr>
<td>Senior faculty</td>
<td>46 (38.0)</td>
</tr>
<tr>
<td>Proportion of practice devoted to stroke</td>
<td></td>
</tr>
<tr>
<td>&lt;10%</td>
<td>5 (4.1)</td>
</tr>
<tr>
<td>10%–50%</td>
<td>27 (21.3)</td>
</tr>
<tr>
<td>51%–99%</td>
<td>64 (52.9)</td>
</tr>
<tr>
<td>100%</td>
<td>25 (20.7)</td>
</tr>
</tbody>
</table>
Results

Demographics
Survey responses were received from 122 of 333 (37%) academic stroke neurologists. The characteristics of the survey participants are shown in Table 1. One quarter were female, more than three quarters were board certified in vascular neurology, all 4 geographic regions of the country were well represented, early career, midcareer, and senior faculty were each well represented, and the median years in practice was 11 to 15 years. All respondents were engaged in stroke clinical care, with neurovascular disease accounting for more than half of the practice time for >70% of those surveyed.

Representativeness
The respondents appeared representative of the entire sample (Table 2). Responders and nonresponders did not differ in female sex (23.0% versus 28.4%; P=0.30) or geographic region (P=0.54). Among the responders, two thirds were early and one third were late responders. There were no significant differences between early and late responders in years of practice, board certification, current appointment level, and proportion of practice dedicated to stroke (P values ranging from 0.29 to 0.78).

Clinically Important Treatment Effects
The Figure shows the distribution of responses on the MICD. The median MCID per 1000 treated patients was 11 to 15 (interquartile range, 6–10 to >20). Converted to natural base 100 values, the MCID was 1.1% to 1.5% (interquartile range, 0.6%–1.0% to >2%). Early versus late responders did not differ in MCID selections (P=0.88).

Discussion
This study found that, when presented with response options with denominators of 1000 patients, permitting selection of smaller increments of benefit, expert stroke physicians identify 1% to 1.5% as the minimal clinically important increase in the proportion of patients achieving functional independence at 3 months to make a completely safe stroke treatment worth administering in clinical practice. This clinical effect size is equivalent to a number needed to treat for 1 additional patient to be functionally independent of 67 to 91. The responding academic vascular neurologists had extensive practice experience and were well distributed across all regions of the country, and formal responder–nonresponder, and continuum of resistance tests suggested that the respondent experts were representative of US academic vascular neurologists.

The MCID value of 1.1% to 1.5% in this study is substantially lower than the 5% to 10% derived in a prior study that also used an expert survey approach. Compared with that prior derivation, the current study MCID value harmonizes better with estimates of clinically worthwhile treatment effects derived by other techniques. A formal socioeconomic model of stroke impact found that a safe neuroprotective would be cost-effective and clinically worthwhile if it improved outcomes of 2% to 3% of patients. Most importantly, an MCID value of 1% to 1.5% is also indicated by actual physician behaviors in relation to the existing therapy of antiplatelet treatment with aspirin for acute ischemic stroke patients not being treated with reperfusion interventions. In the most recent Cochrane meta-analysis, analyzing 41,291 patients randomized in 4 trials, acute aspirin increases the rate of functional independence at long-term follow-up from 53.8% to 55.0%, an absolute increase of 1.2% (P=0.01). On the basis of this demonstrated treatment effect, acute antiplatelet therapy with aspirin is endorsed in stroke treatment guidelines worldwide as a performance measure used by hospital accrediting agencies in the United States to assess hospital care quality, and is essentially universally used in actual clinical practice.7 The risk and administration mode of acute aspirin is similar in key respects to the theoretical neuroprotective agent described in the current survey vignette: simple to administer and a low incidence of adverse events. Accordingly, the worldwide guideline

Table 2. Evaluation for Responder Bias

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Early responders (n=81)</th>
<th>Late responders (n=41)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex, female</td>
<td>28 (23.0)</td>
<td>60 (18.1)</td>
<td>0.30</td>
</tr>
<tr>
<td>Geographic location</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northeast</td>
<td>41 (33.6)</td>
<td>72 (34.1)</td>
<td></td>
</tr>
<tr>
<td>Midwest</td>
<td>26 (21.3)</td>
<td>46 (21.8)</td>
<td>0.54</td>
</tr>
<tr>
<td>South</td>
<td>32 (26.2)</td>
<td>65 (30.8)</td>
<td></td>
</tr>
<tr>
<td>West</td>
<td>23 (18.9)</td>
<td>28 (13.3)</td>
<td></td>
</tr>
<tr>
<td>Difference between early and late responders did not differ in MCID selections (P=0.88).</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
MCID estimates.\textsuperscript{14} Presentation formats can yield substantially different expert judgments accords with prior studies indicating that varying when multiple choice questions are used to elicit physician anchoring and centrality bias may influence MCID estimates.\textsuperscript{11} This approach is subject to both anchoring and centrality bias. The options presented anchored, indeed constrained, the permitted responses to the 1.7% to 20% range, and the middle range of choices implicitly suggested that the correct answer would lie in the 3% to 10% range. The finding that anchoring and centrality bias may influence MCID estimates when multiple choice questions are used to elicit physician judgments accords with prior studies indicating that varying presentation formats can yield substantially different expert MCID estimates.\textsuperscript{14}

Additional potential, but less likely, contributors to the MCIDs between the studies are treatment effect magnitude framework and end point type. The current study asked physicians to consider the benefit per 100 (BPH)—the proportion of patients who benefit among 100 treated. The BPH is a version of the absolute risk reduction. In contrast, the prior survey study asked physicians to consider the number needed to treat (NNT) for 1 patient to benefit. The NNT is the inverse of the absolute risk reduction. We are not aware of any studies showing that MCID values differ when queries are framed using BPH versus NNT perspectives. The BPH aligns better with the decision-making perspective of physicians, as they participate in the same treatment decisions multiple times over the course of their practice. A BPH framework may therefore be more appropriate when MCID derivations use physician and healthcare provider informants. The NNT, while also useful for physicians, aligns better with the decision-making perspective of patients, as they typically participate in the examined treatment decision only once. An NNT framework may therefore be more appropriate when MCID derivations use patient and family informants. The end point in the prior study was improvement by 1 step anywhere along the 7-level modified Rankin Scale (mRS) of disability, whereas the current study used improvement to an mRS score of the 0 to 2 range. Some 1-step health state transitions on the mRS scale are valued more highly and others less highly than the transition from mRS score of 3 to mRS score of 2.\textsuperscript{15}

It is important to note that the findings of this study are specific to treatments for acute ischemic stroke that are both safe and easy to administer, like many putative neuroprotective agents and acute aspirin. For such agents, physician informants, when contemplating the MCID, may focus solely on the degree of benefit an agent confers. In contrast, some acute stroke therapies have substantial risks of causing harm (eg, symptomatic intracerebral hemorrhage fibrinolytic therapy) or substantial personnel effort and expense to administer (eg, endovascular thrombectomy by specialized teams in neurocatheterization laboratories). To offset these important drawbacks of some treatments, physician informants will likely specify that higher MCIDs are needed for therapies with safety, physician effort, or expense limitations.

The findings of this study have direct implications for stroke trial funding and regulatory oversight bodies. For example, in the past, the National Institute of Neurological Disorders and Stroke encouraged that the proposals for phase 3 acute ischemic stroke trials choose sample sizes based primarily on the MCID, and not on larger effect magnitudes that may have been observed in preclinical and phase 2 studies. This approach had the merit of ensuring that trial results would definitively determine whether the studied intervention was of any clinical benefit. However, given the small MCID for acute ischemic stroke, trial sample sizes must be large enough to fully exclude the MCID. For a population with a control functional independence outcome rate by at least the MCID of 1.3%, with a 2-tailed $\alpha$ of 0.05, would need to enroll 45 744 patients (22 872 in each arm). Such a large trial is financially infeasible, lending support to the more recent approach of National Institute of Neurological Disorders and Stroke, permitting trial designs to detect large clinical effects when well-supported by preliminary data (Scott Janis, PhD, Oral Personal Communication, April 17, 2014).

From a regulatory perspective, the US Food and Drug Administration device branch has at times required that sponsors of pivotal acute ischemic stroke trials specify the MCID in advance and established a framework to assess trial success

**Figure**. Physician responses of minimal clinically important difference (MCID). Number of additional patients, per 1000 treated, that expert stroke neurologists indicated would need to benefit, by achieving freedom from disability at 90 days, for a safe neuroprotective drug to be worthwhile for use in clinical practice ($n = 122$). Left. Sixty-eight percent of experts indicated $\leq 20$ per 1000 and 32% indicated $>20$ per 1000. Right. More granular categories, with median response in the 11 to 15 per 1000 category.
using 2 criteria: not only the usual achievement of statistically significant differences between the treatment arms but also a requirement that the point estimates for superiority exceed the prespecified MCID, thereby indicating probable rather than possible clinical importance.10,11 Sponsors and trialists should not simply state that the MCID is the exact treatment effect magnitude that the trial is powered to detect, unless that is truly the case. Often trials are powered to detect effect magnitudes that substantially exceed the MCID, given the infeasible large size needed for a trial to detect the minimum important difference. As a result, trials finding an actual clinically substantial benefit of therapy could be incorrectly perceived as having ambiguous results. For example, a stroke trial may be powered to detect an absolute 8% improvement in disability outcomes between groups, even though the MCID is 2%, simply because the governmental or industry sponsor cannot afford to support a large trial adequately powered to detect a 2% difference. If the trial results in a statistically significant group difference in a disability of 6%, the results are appropriately interpreted as not only statistically but also clinically significant. But failure to prespecify the MCID, and the fact that the MCID is lower than the difference the trial is powered to detect, can obscure this reality.

To avoid this uncertainty, for acute ischemic stroke therapies, it is important that sponsors, trialists, and regulators be aware that it often will be appropriate to prespecify an MCID value that is much lower than the effect size a registration study is powered to detect. As a result, if the study shows statistical superiority for the novel treatment, even if the effect size point estimate is mildly below the effect the study was powered to identify, the therapy will still be recognized as exceeding the MCID to a substantial degree and the study can straightforwardly be interpreted as having a positive result.

This study has limitations. The MCID survey question in the current survey was itself subject to anchoring and centrality bias. This aspect was planned, as it permitted the current study to meet its 2 aims of (1) demonstrating that expert-derived MCID values are subject to these cognitive biases and the values will alter depending on the anchoring used in the question architecture, and (2) with appropriate anchoring, expert surveys yield MCID estimates that accord, rather than conflict, with the more accurate MCIDs found by analysis of actual expert clinical practice. However, future studies are desirable that use question architectures that may be less prone to cognitive biases, such as offering response options covering a range of effect sizes that spans 3, rather than 2, orders of magnitude. The response rate to the survey invitation was moderate. But there was no evidence of nonresponse bias from analysis of demographic features of responders and nonresponders. In addition, continuum of resistance analysis, showing no differences in demographic features and substantive answers of early versus late waver responders, supported that the responding physicians were not systematically different from all physician experts invited to participate.10,11,17

**Conclusions**

In conclusion, this study provides evidence that physician expert responses to MCID surveys may be affected by anchoring and centrality bias. When survey design takes these into account, the expert-derived MCID for a safe acute ischemic stroke treatment is 1.1% to 1.5%, in accord with actual physician behavior in guideline writing and clinical practice. This revised MCID value, supported by both formal expert surveys and by clinician behavior in guideline writing and in clinical practice, can guide clinical trial design and grant-funding and regulatory agency decisions.

**Sources of Funding**

This work was supported in part by Award National Institutes of Health-National Institute of Neurological Disorders and Stroke U10NS086497 (Dr Saver).

**Disclosures**

Dr Saver is an employee of the University of California. The University of California has patent rights in retrieval devices for stroke. Dr Saver has served as an unpaid site investigator in multicenter trials run by Medtronic and Stryker for which the UC Regents received payments on the basis of clinical trial contracts for the number of subjects enrolled. Dr Saver has received funding for services as a scientific consultant on trial design and conduct to Medtronic/Coviden, Stryker, Neuravi, BrainsGate, Pfizer, Squibb, Boehringer Ingelheim (prevention only), ZZ Biotech, and St. Jude Medical. The University of California has released the Rankin Focused Assessment for free use under a Creative Commons license and has copyright for Rankin Scale training vignettes. The other authors report no conflict.

**References**


Minimal Clinically Important Difference for Safe and Simple Novel Acute Ischemic Stroke Therapies
Jessica S. Cranston, Brett D. Kaplan and Jeffrey L. Saver

*Stroke*. 2017;48:2946-2951; originally published online September 20, 2017;
doi: 10.1161/STROKEAHA.117.017496

*Stroke* is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2017 American Heart Association, Inc. All rights reserved.
Print ISSN: 0039-2499. Online ISSN: 1524-4628

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://stroke.ahajournals.org/content/48/11/2946

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Stroke* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at: http://www.lww.com/reprints

Subscriptions: Information about subscribing to *Stroke* is online at: http://stroke.ahajournals.org/subscriptions/