Journal of the American Heart Association

ORIGINAL RESEARCH

Validity of Early Outcomes as Indicators for Comparing Hospitals on Quality of Stroke Care

Marzyeh Amini, PhD; Frank Eijkenaar, PhD; Hester F. Lingsma, PhD; Sanne J. den Hartog, MD, PhD; Susanne G. H. Olthuis, MD; Jasper Martens, MD; Bart van der Worp, MD, PhD; Wim van Zwam, MD, PhD; Anouk van der Hoorn, MD, PhD; Stefan D. Roosendaal, MD, PhD; Bob Roozenbeek, MD, PhD; Diederik Dippel, MD, PhD; Nikki van Leeuwen, PhD; on behalf of the MR CLEAN Registry Investigators*

BACKGROUND: Insight into outcome variation between hospitals could help to improve quality of care. We aimed to assess the validity of early outcomes as quality indicators for acute ischemic stroke care for patients treated with endovascular therapy (EVT).

METHODS AND RESULTS: We used data from the MR CLEAN (Multicenter Randomized Controlled Trial of Endovascular Treatment for Acute Ischemic Stroke in the Netherlands) Registry, a large multicenter prospective cohort study including 3279 patients with acute ischemic stroke undergoing EVT. Random effect linear and proportional odds regression were used to analyze the effect of case mix on between-hospital differences in 2 early outcomes: the National Institutes of Health Stroke Scale (NIHSS) score at 24 to 48 hours and the expanded thrombolysis in cerebral infarction score. Between-hospital variation in outcomes was assessed using the variance of random hospital effects (tau^2). In addition, we estimated the correlation between hospitals' EVT-patient volume and (case-mix–adjusted) outcomes. Both early outcomes and case-mix characteristics varied significantly across hospitals. Between-hospital variation in the expanded thrombolysis in cerebral infarction score was not influenced by case-mix adjustment (tau ^{2}=0.17 in both models). In contrast, for the NIHSS score at 24 to 48 hours, case-mix adjustment led to a decrease in variation between hospitals (tau ^{2} decreases from 0.19 to 0.17). Hospitals' EVT-patient volume was strongly correlated with higher expanded thrombolysis in cerebral infarction scores (r=0.48) and weakly with lower NIHSS score at 24 to 48 hours (r=0.15).

CONCLUSIONS: Between-hospital variation in NIHSS score at 24 to 48 hours is significantly influenced by case-mix but not by patient volume. In contrast, between-hospital variation in expanded thrombolysis in cerebral infarction score is strongly influenced by EVT-patient volume but not by case-mix. Both outcomes may be suitable for comparing hospitals on quality of care, provided that adequate adjustment for case-mix is applied for NIHSS score.

Key Words: acute ischemic stroke ■ case-mix ■ early outcome ■ expanded thrombolysis in cerebral infarction ■ hospitals' patient volume ■ National Institutes of Health Stroke Scale ■ quality of care

Insight into between-hospital differences in outcome might help to improve the quality of care. One of the most important considerations in this regard is the selection of valid and reliable outcome indicators that are used for benchmarking hospitals. Outcome measures reflect the impact of health care services and interventions on the health status of patients,[1–3] and differences in outcome may represent real differences in quality of care. For example, the differences may be attributable to variation in the use of treatments, the process of care,

Correspondence to: Marzyeh Amini, PhD, Department of Public Health, Erasmus Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands. Email: m.amini@erasmusmc.nl

Supplemental Material is available at https://www.ahajournals.org/doi/suppl/10.1161/JAHA.122.027647

*A complete list of the MR CLEAN Registry Investigators can be found in the Appendix at the end of the article.

For Sources of Funding and Disclosures, see page 11.

JAHA is available at: www.ahajournals.org/journal/jaha

CLINICAL PERSPECTIVE

What Is New?

- Using data from a large nationwide registry, this study assessed the effect of hospitals' case-mix and patient volume on between-hospital variation in the early outcomes "reperfusion rate" and "neurological deficit" after endovascular therapy of ischemic stroke.
- Rates of successful reperfusion were significantly higher in high-volume hospitals than in low-volume hospitals, and case-mix adjustment did not influence between-hospital variation in this outcome.
- Variation in neurological deficit after endovascular therapy was not related to patient volume, whereas case-mix explained a significant portion of the variation between hospitals.

What Are the Clinical Implications?

- For benchmarking endovascular therapy hospitals on the quality of acute ischemic stroke care, both early outcomes might be suitable provided adequate adjustment for case-mix is performed for the outcome neurological deficit.

Nonstandard Abbreviations and Acronyms

AIC	Akaike Information Criterion
eTICI	expanded thrombolysis in cerebral infarction
EVT	endovascular therapy
mRS	modified Rankin Scale
NIHSS	National Institutes of Health Stroke Scale

and/or less measurable aspects, such as care providers' experience in performing certain treatments.[4–7] However, a meaningful comparison of hospitals on outcome requires adequate accounting for important methodological issues, specifically adjustments for case-mix and random variation.[1,8,9]

Selecting appropriate outcome measures for comparing hospitals on quality of care for acute ischemic stroke poses a major challenge for investigators. In stroke care, the modified Rankin Scale (mRS) score at 90 days is a functional outcome that describes the degree of overall disability or dependence in daily life after stroke care, and it has been used as a measure of treatment outcome in trials and benchmarking exercises.[10] However, the long time span between treatment and mRS outcome assessment may experience loss to follow-up, leading to incomplete outcome data, which may threaten the validity and reliability of the mRS score as a quality indicator.[11] More important, previous studies have shown that significant differences between endovascular therapy (EVT) hospitals in mRS score at 90 days are the result of differences in case-mix rather than the quality of care.[8,9] Therefore, other outcomes should be considered that may be more valid representations of the quality of acute stroke care and, thus, are more useful for between-hospital comparisons. A potentially suitable early outcome measure in this respect is the reperfusion rate, as measured by the expanded thrombolysis in cerebral infarction (eTICI) score, because a high rate of reperfusion is a major contributor to better outcome after EVT.[12] Another potentially suitable early outcome measure is the neurological deficit at 24 to 48 hours, as measured with the National Institutes of Health Stroke Scale (NIHSS), which is easily assessed during hospital stay and is a reliable measure of stroke neurological deficit.[13]

In the Netherlands and other countries, specialized treatment of patients with ischemic stroke is centralized in EVT hospitals.[14,15] This centralization leads to higher volume of EVT in these hospitals. Previous research has shown that patient volume has a minor influence on functional outcome (mRS score) after stroke, possibly because this outcome poorly reflects the increased quality of care that may result from higher volumes.[9] Analysis of the impact of hospital's EVT-patient volume and case-mix on between-hospital differences in different types of early outcome measures may provide more insight into these measures' validity as indicators of quality of stroke care.[16] Therefore, using data from a large nationwide registry, we aimed to assess the effect of case-mix and hospitals' patient volume on between-hospital variation in reperfusion rate (eTICI score) and neurological deficit (NIHSS score) after EVT for ischemic stroke, to evaluate their validity as quality indicators.

METHODS

Data cannot be made available, as no patient approval has been obtained for sharing coded data. However, syntax files and output of statistical analyses will be made available on reasonable request to the corresponding author.

Study Design and Patients

We used data collected between March 2014 and November 2017 from the MR CLEAN (Multicenter Randomized Controlled Trial of Endovascular Treatment for Acute Ischemic Stroke in the Netherlands) Registry, a prospective, observational study in all 17 hospitals that perform EVT in the Netherlands (Figure S1).[14] We

applied the following inclusion criteria: treatment at age ≥18 years, treatment in a hospital that participated in the MR CLEAN trial, and proximal intracranial vessel occlusion in the anterior circulation (internal carotid artery, internal carotid artery terminus, middle [M1/M2] cerebral artery, or anterior [A1/A2] cerebral artery), as shown by computed tomography angiography. Details on the study design and objectives of the MR CLEAN Registry have been described previously.[14]

The MR CLEAN Registry was approved by the ethics committee of the Erasmus University Medical Center, Rotterdam, the Netherlands (MEC-2014-235). With this approval, the registry was approved by the research board of each participating center. At UMC Utrecht, approval to participate in the study has been obtained from that university's own research board and ethics committee. All subjects gave informed consent.

Variables

Outcome Measures

We analyzed variation in 2 early outcome measures: the ordinal eTICI score and the continuous NIHSS score at 24 to 48 hours after EVT. The eTICI score ranges from 0 (no antegrade reperfusion of the occluded vascular territory) to 3 (successful reperfusion).[14,17] Reaching a score of 2B or higher required the completion of digital subtraction angiography runs, including both anteroposterior and lateral views after EVT. If a lateral view was missing, 2A was the highest possible score. An imaging core laboratory adjudicated all patient imaging, but we used the self-reported, locally assessed eTICI for the current analysis. The members of the core laboratory were blinded to all clinical data, except for the symptom side. The NIHSS score measures neurological deficit at 24 to 48 hours after EVT. NIHSS scores were assessed by the treating neurologist.[18]

Case-Mix Variables

We considered the possibility that prognostic factors for reperfusion and neurological deficit are not equally distributed across hospitals, thus possibly confounding the estimation of hospitals' effect on early outcomes. Patients' age, sex, relevant medical history (ie, previous stroke, atrial fibrillation, hypertension, and diabetes), prestroke mRS score, baseline score on the NIHSS, occlusion location, systolic and diastolic blood pressure at admission, and the time between stroke onset and arrival at the emergency department of the EVT hospital were considered as potential confounders in case-mix–adjusted models. These patient and neuroimaging characteristics were selected on the basis of clinical knowledge and previous studies.[1,8] Specifically, each characteristic must be associated with outcome and should not be influenceable by hospitals.

Hospital Patient Volume

We assessed hospital patient volume, which reflects the experience of a hospital with EVT and, thus, represents to some extent quality of care. Volume was defined as the number of patients treated in each hospital each year during the study period. For illustrative purposes and visualization of crude data, volume was divided into quartiles (≤29, 30–32, 33–36, and >36 patients), but in the statistical analyses, we used the continuous measure of absolute patient volume.

Statistical Analysis

We used Pearson χ^2 statistics and the nonparametric Kruskal-Wallis tests for univariable comparisons of hospitals on the eTICI score, NIHSS score at 24 to 48 hours, and case-mix variables. Between-hospital differences in early outcome were analyzed using both random effect proportional odds regression (for the ordinal eTICI score) and random effect linear regression (for the continuous NIHSS score at 24–48 hours) models. Patients who had died before NIHSS assessment received the maximum NIHSS score of 42. NIHSS scores at 24 to 48 hours were then log10 transformed to meet the assumption of normally distributed residuals in the regression model, after adding 1 point to all NIHSS scores, so that the log10-transformed NIHSS score of 0 would remain 0.[19]

Separately for the 2 outcomes, we first fitted an unadjusted model including only a random hospital intercept, providing insight into between-hospital variation in outcome accounting only for random variation. Next, we ran a second model in which we, in addition to the random hospital effect, also adjusted for the individual-level fixed effects of the case-mix variables on the outcomes. These regression models were used to estimate each hospital's effect on the 2 early outcomes, without and with adjustment for case-mix.

We calculated the correlation between hospitals' effect estimates from the unadjusted model with those from the adjusted model to provide insight into the influence of case-mix on the hospital comparisons. In addition, we compared the between-hospital variation in outcome (measured by tau^2, the variance of the random hospital effects) between the 2 models. The model fit was assessed using Akaike Information Criterion (AIC), with a lower AIC value indicating a better fit.[20]

To assess the influence of EVT-patient volume on variation in outcome, we calculated the correlation between hospitals' absolute EVT-patient volume and the hospitals' effect estimates on outcomes, with and without case-mix adjustment.

All analyses were conducted with R statistical software version 3.4.3 (R Foundation for Statistical Computation, Vienna, Austria), using the *clmm* module

in the *ordinalimputation* package and the *lmer* module in the *lme4* package. Statistical significance was assessed at $P<0.05$ in all analyses.

Multiple imputation was used to deal with missing case-mix values, which ranged from 0.7% (previous diabetes) to 56% (diastolic blood pressure). To handle the missing data, we first evaluated the pattern of missing data using the *md.pattern* function of the mice package in R. We also used *mcar_test* to assess whether data are missing completely at random. The high portion of missingness in systolic and diastolic blood pressures was related to age, NIHSS score baseline, location of occlusion, and prestroke mRS score. The Little test of missing completely at random showed that the missing data are not missing completely at random (ie, data are missing at random).[21] We conducted a sensitivity analysis in which we only used data from cases with complete data (ie, complete case analysis). Results were highly similar compared with those based on multiple imputation, but because complete case analysis significantly decreases the sample size and statistical power, we proceeded with the multiple imputed data. Therefore, we fitted imputation models[22] and imputed data 5 times using both the case-mix and outcome variables. Each imputed data set was analyzed separately, after which the results were pooled.

RESULTS

Descriptive Analyses

A total of 3279 patients were included in the study (Figure S2). At the hospital level, the median patient age ranged from 68 to 77 years (Table). Case-mix differences between hospitals were statistically significant for previous stroke (range, 0%–26%), atrial fibrillation (range, 13%–37%), hypertension (range, 41%–67%), prestroke mRS score, and location of occlusion. Also, the median significantly differed across hospitals for baseline NIHSS score (range, 13–17), time from stroke onset to arrival at the emergency department of the EVT hospital (range, 52–160 minutes), systolic blood pressure (range, 142–151 mm Hg), and diastolic blood pressure (range, 78.5–88 mm Hg). The annual number of patients receiving EVT also varied considerably between hospitals across the 4 years (Table).

No reperfusion (eTICI score, 0) was achieved for 8% to 30% of patients across hospitals. A total of 0% to 7% of patients achieved an eTICI 1 score, and 8% to 31% received an eTICI 2A score. Successful reperfusion (eTICI score, >2B) was achieved in 30% to 80% of patients across hospitals ($P<0.001$). The median NIHSS score at 24 to 48 hours after EVT varied significantly ($P=0.001$) between hospitals from 8 to 15 (Table).

Effect of Case-Mix Adjustment on Variation in Outcome

For reperfusion rate (eTICI score), the between-hospital variation in outcome (tau^2) in the unadjusted model was 0.17, with an AIC of 6219. The case-mix–adjusted model yielded similar figures ($tau^2=0.17$; AIC=6186), suggesting no significant influence of case-mix adjustment on between-hospital variation in eTICI score (Figure 1A). This finding was underscored by a strong positive correlation ($r=0.99$) between the hospital effect estimates from the unadjusted model and those from the case-mix–adjusted model. In contrast, for neurological deficit (NIHSS score), this correlation between unadjusted and case-mix–adjusted hospital estimates was much lower ($r=0.51$). For this early outcome, tau^2 reduced from 0.19 in the unadjusted model (AIC=3541) to 0.17 in the case-mix adjustment model (AIC=2348), suggesting that some of the between-hospital variation in outcome is driven by differences in case-mix (Figure 1B).

Effect of Hospital EVT-Patient Volume

Reperfusion rates differed across quartiles of hospitals' patient volume; higher-volume hospitals had significantly larger proportions of patients with successful reperfusion (eTICI score, >2B; $P<0.001$; Figure 2A). In contrast, there were no significant differences in neurological deficit (NIHSS score) across quartiles of patient volume ($P=0.71$; Figure 2B).

When the effect of volume was analyzed as a continuous variable, unadjusted successful reperfusion rates (eTICI score, >2B) were higher in higher-volume hospitals ($r=0.49$; $P<0.05$). After adjusting for case-mix, the correlation decreased slightly but still was statistically significant ($r=0.48$; $P<0.05$; Figure 3A). In contrast, the unadjusted ($r=0.09$; $P=0.68$) and adjusted ($r=0.15$; $P=0.73$) neurological deficit was independent of EVT-patient volume (Figure 3B).

DISCUSSION

In this observational study, we found significant variation in the early outcome measures eTICI score (reperfusion rate) and NIHSS score (neurological deficit) between hospitals providing EVT for acute ischemic stroke in the Netherlands. Rates of successful reperfusion were significantly higher in high-volume hospitals than in low-volume hospitals, and case-mix adjustment did not influence between-hospital variation in this outcome. In contrast, variation in neurological deficit after EVT was not related to patient volume, whereas case-mix did explain a significant portion of the variation between hospitals.

Table . Characteristics of Patients Treated in Intervention Hospitals in the MR CLEAN Registry (N=3279)

Characteristic	Total population, n (%)/ median (IQR)	Hospital-level range, median/%	*P* value	Missing N (%)
Case-mix				
Age, y	72 (61–80)	68–77	0.001	0
Men	1696 (52)	39–55	0.79	0
Medical history				
Previous stroke	546 (17)	0–26	<0.001	27 (0.8)
Atrial fibrillation	772 (24)	13–37	<0.001	43 (1.3)
Hypertension	1688 (53)	41–67	<0.001	66 (2.0)
Diabetes	532 (16)	12–25	0.09	24 (0.7)
Prestroke modified Rankin Scale score				72 (2.2)
0	2170 (68)	45–87	<0.001	
1	424 (13)	3–19		
2	241 (8)	0–16		
≥3	372 (12)	5–25		
NIHSS score at baseline	16 (11–20)	13–17	<0.001	55 (1.7)
Location of occlusion			<0.001	160 (4.99)
M1	1815 (58)	45–70		
M2	455 (15)	4–31		
Intracranial ICA	161 (5)	0–10		
ICA-T	663 (21)	9–27		
Other (M3/anterior)	25 (1)	0–4		
Time from onset to arrival at ED, min	135 (65–195)	52–160	<0.001	124 (3.8)
Systolic blood pressure at admission, mm Hg	150 (131–165)	142–151	<0.001	1833 (55.9)
Diastolic blood pressure at admission, mm Hg	80 (70–91)	78.5–88	<0.001	1837 (56.0)
Hospital's EVT-patient volume				
2014	187 (6*)	0–18†		
2015	822 (25*)	1–13†		
2016	1131 (34*)	1–14†		
2017	1139 (35*)	0–13†		
Outcomes				
eTICI score			<0.001	88 (2.7)
0	535 (16)	8–30		
1	95 (3)	0–7		
2A	596 (18)	8–31		
2B, 2C, or 3	1965 (60)	30–80		
NIHSS score 24–48 h after EVT	10 (4–17)	8–15	<0.001	253 (7.7)

ED indicates emergency department; eTICI, expanded thrombolysis in cerebral infarction; EVT, endovascular therapy; ICA, internal carotid artery; ICA-T, ICA terminus; IQR, interquartile range; MR CLEAN, Multicenter Randomized Controlled Trial of Endovascular Treatment for Acute Ischemic Stroke in the Netherlands; M1/M2, middle cerebral artery; and NIHSS, National Institutes of Health Stroke Scale.

*Hospital EVT-patient volume as a percentage of all patients receiving EVT and treated in the Netherlands in the relevant year.

†The hospital-level range in EVT-patient volume is defined as the range in the percentage of all patients receiving EVT and treated in each hospital relative to all patients receiving EVT and treated in the Netherlands in the relevant year.

Validity of and Actionability of Early Outcomes as Quality Indicator

Benchmarking hospitals on outcomes is useful for improving quality of care only if the measured outcomes are affected by the provided care and reflect the quality performance of hospitals.[23] In other words, selecting outcomes for benchmarking should involve considering the interpretability, actionability, and feasibility of each outcome; and the outcomes must be valid indicators of the quality of care.[16] The validity is often described as sensitivity to adjustment for case-mix.[24] In that sense, our findings suggest that the outcomes reperfusion rate and neurological deficit, measured early

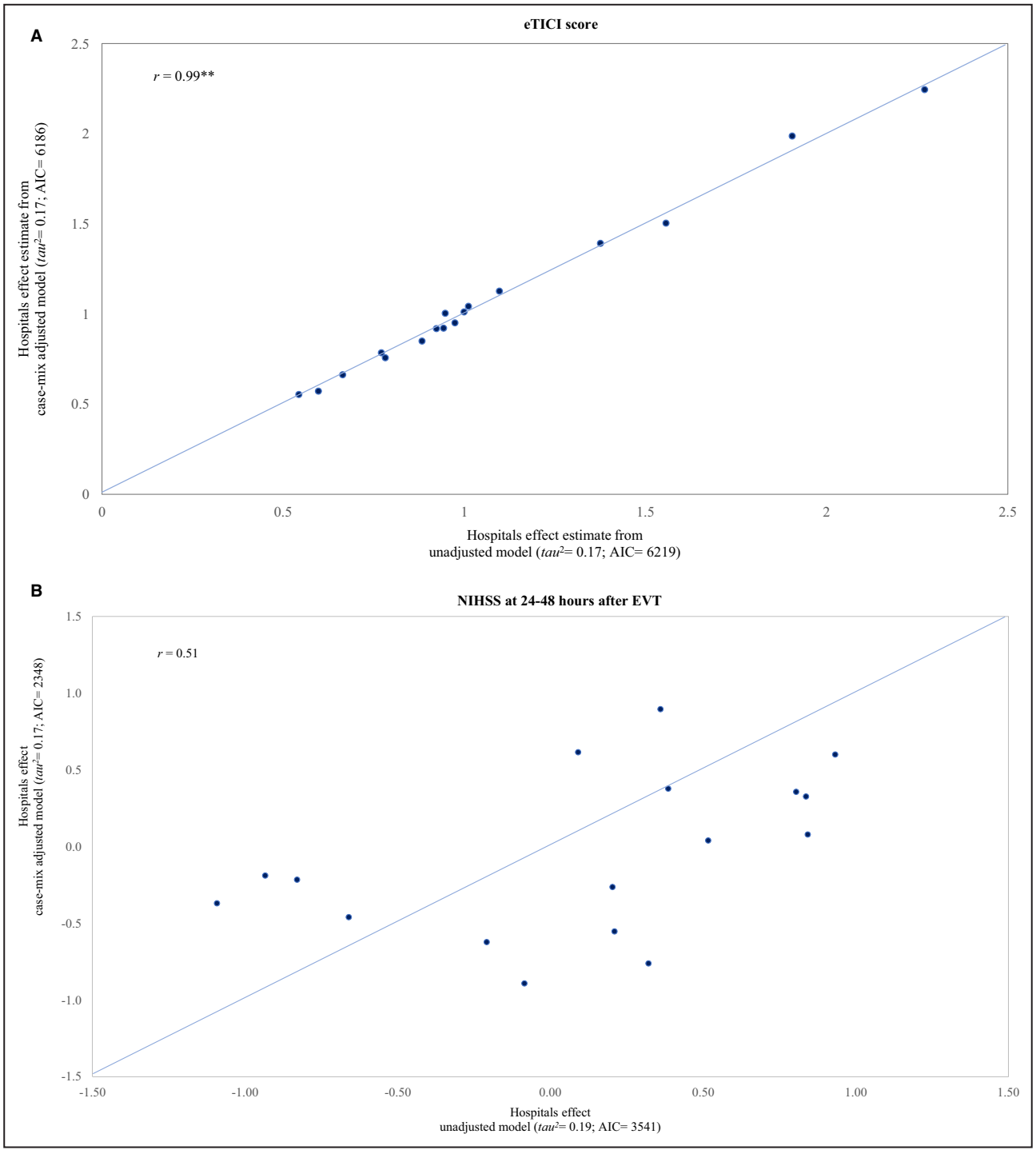

Figure 1. Plots of hospital effects from unadjusted and case-mix–adjusted models for expanded thrombolysis in cerebral infarction (eTICI) score (A) and National Institutes of Health Stroke Scale (NIHSS) score 24 to 48 hours after endovascular therapy (EVT; B). Results are from the random effect proportional odds regression analysis for ordinal eTICI score and random effect linear regression analysis for NIHSS score 24 to 48 hours after EVT. Each dot represents a hospital effect estimate related to the relevant outcome. A lower Akaike Information Criterion (AIC) value indicates a better model fit. The tau^2 is the variance of the random hospital effects. **Correlation is significant at the 0.01 level.

after treatment, might be more valid than the functional outcome measure mRS score at 90 days, because clinical status later can be influenced by other factors than quality of care, such as depression and recurrent ischemic events.[25] In our previous study in which we analyzed the effect of differences in performance

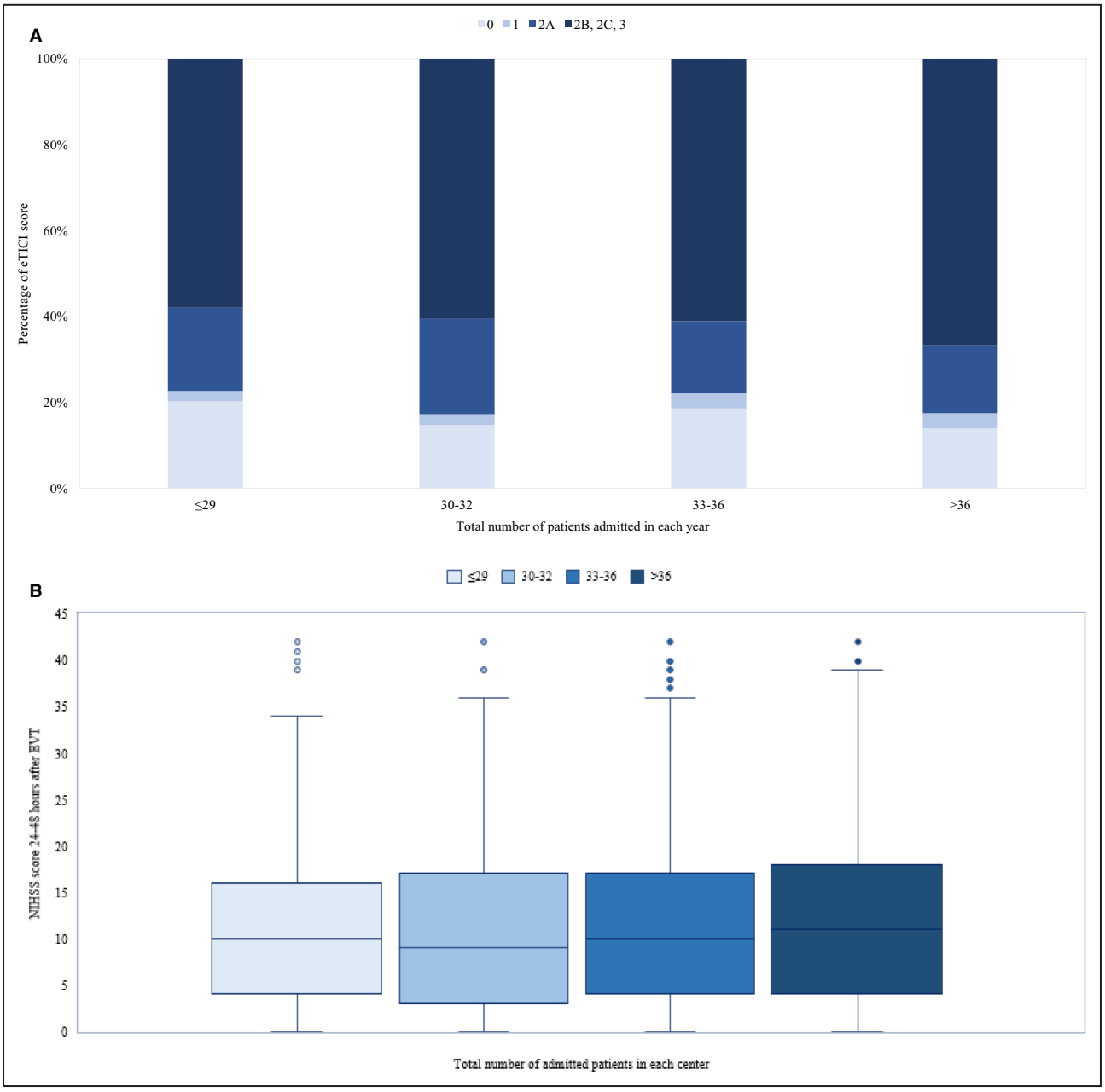

Figure 2. Distribution of expanded thrombolysis in cerebral infarction (eTICI) score (A) and National Institutes of Health Stroke Scale (NIHSS) score 24 to 48 hours after endovascular therapy (EVT; B) across quartiles of annual patient volume in EVT hospitals.

on structure and processes of care, and case-mix on between-hospital differences in mRS score at 90 days, we concluded that between-hospital variation in the mRS outcome of patients with ischemic stroke mostly reflects differences in case-mix, rather than differences in structure or process of care.[9] In this study, we found that variation in neurological deficit after EVT, similar to mRS score at 90 days,[9] is substantially influenced by differences in case-mix. Many of the case-mix variables (patient characteristics, like age, and history of another disease) are not modifiable by the care provided. This emphasizes that case-mix adjustment should be done thoroughly and accurately when comparing hospitals on these 2 outcomes for it to be a valid measure for comparing providers on quality of care.[25] But the use of any of these outcomes for benchmarking depends on the time period during which each performance measure needs to be evaluated. For example, some processes of care are required to be performed within 24 hours after treatment, others before discharge, and

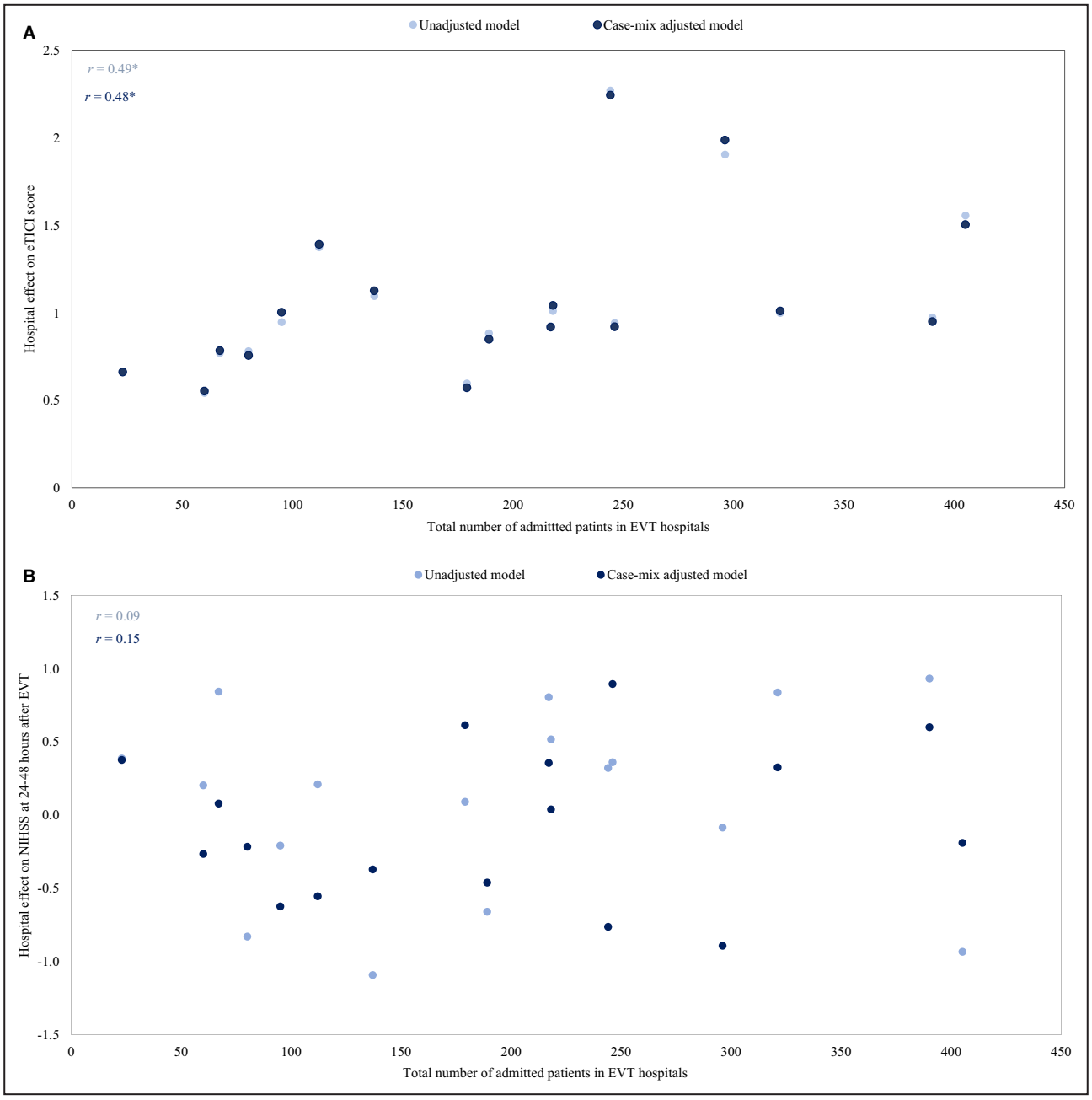

Figure 3. Correlation between hospitals effect on expanded thrombolysis in cerebral infarction (eTICI) score (A) and National Institutes of Health Stroke Scale (NIHSS) score 24 to 48 hours after endovascular therapy (EVT; B), and absolute hospital EVT-patient volume.
Each dot represents a hospital effect estimate related to the relevant outcome. *Correlation is significant at the 0.05 level.

still others within 3 months of discharge. This requires giving due consideration to the circumstances of clinical care when specifying the period of care.[16]

For improvement purposes, indicators should be actionable for hospitals. This is questionable for both functional and neurological deficit outcomes because it is not obvious what improving quality of care would require. This may be different for the reperfusion rate (eTICI score), for which we found that between-hospital variation does not reflect differences in case-mix. However, using this outcome measure for comparing hospitals on quality still requires caution. The eTICI score is a "technical" outcome measure that is mainly influenced by the skills and experience of the intervention radiologist in the angiography unit.[7] Even with adjustment for various case-mix variables, like in this study, between-hospital differences might still reflect unobserved confounding attributable to patient

selection. It is possible that more experienced interventionalists are better able to select patients with occlusions (M1) that are easy to handle. As a result, experience may translate into ability to choose the "right" patients for intervention. Case-mix adjustment is only possible for measured confounders, but the above-mentioned mechanism would lead to unmeasured confounding (selection bias), which is difficult to account for in between-hospital comparisons. We do not oppose that, on large-vessel occlusion, patients who did not undergo EVT in the participating centers. However, our study concerns patients treated within 6 hours from onset. The Dutch intervention hospitals providing EVT generally adhere to the Dutch guidelines, which use a liberal indication, as was done in the MR CLEAN trial. No restrictions were found with regard to the Alberta Stroke Program Early CT Score (ASPECTS), anticoagulants, or previous disability. No selection was performed with CT perfusion or collaterals in this time window.

In addition, eTICI only reflects the quality of a specific (short-term) care process and not the overall picture of quality of care. For example, patients with an eTICI score of 3 might still die within the hospital because of complications, like pneumonia and groin bleeds. Therefore, obtaining an overall picture of the quality of the full care path for patients with stroke requires combining comparisons on eTICI score with other relevant clinical and radiological measures.

Hospital's EVT-Patient Volume and Outcome Relationship

We found no significant correlation between neurological deficit (NIHSS score) after EVT and hospitals' patient volume. In contrast, we did find that hospitals that treated more patients had significantly higher rates of successful reperfusion (eTICI score, >2b). High-volume hospitals might have more focus on preoperative protocols with detailed specification of care processes, including collaboration between different professionals, like neurologists, radiologists, anesthesiologists, intensive care unit staff, and nursing staff, for rapid imaging and shorter procedural duration.[26,27] Therefore, when it comes to improving the quality of stroke care, the eTICI score is more actionable for hospitals because it is more sensitive to, for example, logistical improvements that reduce, for example, periprocedural or procedural times to treatment.[27] Further research is needed to understand what processes or other factors underlie the variation in this outcome as this information would be of great importance to providers and patients.

One of the concerns has been that the more severe and/or complex cases are preferentially transferred to larger centers, thus causing a negative impact on hospital-level clinical outcomes. Our results suggest that for eTICI score, this negative impact is not observed, at least when the sample of EVT only cases is considered. Whether this relationship is similar across all stroke cases seen at a hospital that performs EVT is not known and remains an important question.

Strengths and Limitations of the Study

This is the first study that uses nationwide registry data to analyze between-hospital differences in early outcomes in stroke care. Using sophisticated statistical approaches, we examined the influence of case-mix and patient volume on these differences. Specifically, the use of random effect regression modeling allowed us to estimate (the variance of) hospital effects on 2 early outcomes adjusted for chance variation and potential confounding factors.

A limitation of this study is the unavailability of information on other factors that might impact between-hospital differences in outcome. This includes information on unmeasured patient characteristics, care processes, hospital characteristics, and patient volume of interventionists. Another potential limitation is that missing values may have introduced some bias, although we believe to have mitigated this issue considerably using multiple imputation, which is the preferred method over complete case analysis.[28,29] Furthermore, eTICI score is usually self-reported immediately after the treatment, which may lead to response bias.[30] A final limitation is that we only assessed the validity of 2 early outcomes for outcome variation between hospitals. Conclusions might be different for other outcomes that are relevant to patients with stroke treated with EVT, like bleeding complications or NIHSS score at discharge. But an outcome measured at discharge is difficult to interpret, as the timing of discharge is also dependent on the outcome.

CONCLUSIONS

In patients with ischemic stroke, between-hospital variation in neurological deficit after EVT is influenced by case-mix but not by patient volume, whereas variation in reperfusion rate is heavily influenced by patient volume and not by (observed) case-mix differences. For benchmarking EVT hospitals on quality of short-term care for patients receiving EVT, both early outcomes might be suitable provided adequate adjustment for case-mix is performed for neurological deficit. The reperfusion rate might be a particularly valid and actionable outcome indicator for improving quality of care because it appears to reflect true differences in quality of care. Although it reflects only a small part of the care process of patients with stroke, it may be useful for systematically monitoring and subsequently improving the quality of acute stroke care.

APPENDIX

MR CLEAN Registry Investigators

Executive committee

Diederik W. J. Dippel[1]; Aad van der Lugt[2]; Charles B. L. M. Majoie[3]; Yvo B. W. E. M. Roos[4]; Robert J. van Oostenbrugge[5]; Wim H. van Zwam[6]; Jelis Boiten[14]; Jan Albert Vos[8]

Study coordinators

Josje Brouwer [4]; Sanne J. den Hartog[1,2,40]; Wouter H. Hinsenveld [5,6]; Manon Kappelhof[3]; Kars C. J. Compagne[2]; Robert- Jan B. Goldhoorn[5,6]; Maxim J. H. L. Mulder[1,2]; Ivo G. H. Jansen[3]

Local principal investigators

Diederik W. J. Dippel[1]; Bob Roozenbeek[1]; Aad van der Lugt[2]; Adriaan C. G. M. van Es[2]; Charles B. L. M. Majoie[3]; Yvo B. W. E. M. Roos[4]; Bart J. Emmer[3]; Jonathan M. Coutinho[4]; Wouter J. Schonewille[7]; Jan Albert Vos[8]; Marieke J. H. Wermer[9]; Marianne A. A. van Walderveen[10]; Julie Staals[5]; Robert J. van Oostenbrugge[5]; Wim H. van Zwam[6]; Jeannette Hofmeijer[11]; Jasper M. Martens[12]; Geert J. Lycklama à Nijeholt[13]; Jelis Boiten[14]; Sebastiaan F. de Bruijn[15]; Lukas C. van Dijk[16]; H. Bart van der Worp[17]; Rob H. Lo[18]; Ewoud J. van Dijk[19]; Hieronymus D. Boogaarts[20]; J. de Vries[22]; Paul L. M. de Kort[21]; Julia van Tuijl[21]; Jo Jo P. Peluso[26]; Puck Fransen[22]; Jan S. P. van den Berg[22]; Boudewijn A. A. M. van Hasselt[23]; Leo A. M. Aerden[24]; René J. Dallinga[25]; Maarten Uyttenboogaart[28]; Omid Eschgi[29]; Reinoud P. H. Bokkers[29]; Tobien H. C. M. L. Schreuder[30]; Roel J. J. Heijboer[31]; Koos Keizer[32]; Lonneke S. F. Yo[33]; Heleen M. den Hertog[22]; Emiel J. C. Sturm[35]; Paul Brouwers[34]

Imaging assessment committee

Charles B. L. M. Majoie[3] (chair); Wim H. van Zwam[6]; Aad van der Lugt[2]; Geert J. Lycklama à Nijeholt[13]; Marianne A. A. van Walderveen[10]; Marieke E. S. Sprengers[3]; Sjoerd F. M. Jenniskens[27]; René van den Berg[3]; Albert J. Yoo[38]; Ludo F. M. Beenen[3]; Alida A. Postma[6]; Stefan D. Roosendaal[3]; Bas F. W. van der Kallen[13]; Ido R. van den Wijngaard[13]; Adriaan C. G. M. van Es[2]; Bart J. Emmer[3]; Jasper M. Martens[12]; Lonneke S. F. Yo[33]; Jan Albert Vos[8]; Joost Bot[36], Pieter-Jan van Doormaal[2]; Anton Meijer[27]; Elyas Ghariq[13]; Reinoud P. H. Bokkers[29]; Marc P. van Proosdij[37]; G. Menno Krietemeijer[33]; Jo P. Peluso[26]; Hieronymus D. Boogaarts[20]; Rob Lo[18]; Wouter Dinkelaar[2]; Auke P. A. Appelman[29]; Bas Hammer[16]; Sjoert Pegge[27]; Anouk van der Hoorn[29]; Saman Vinke[20]

Writing committee

Diederik W. J. Dippel[1] (chair); Aad van der Lugt[2]; Charles B. L. M. Majoie[3]; Yvo B. W. E. M. Roos[4]; Robert J. van Oostenbrugge[5]; Wim H. van Zwam[6]; Geert J. Lycklama à Nijeholt[13]; Jelis Boiten[14]; Jan Albert Vos[8]; Wouter J. Schonewille[7]; Jeannette Hofmeijer[11]; Jasper M. Martens[12]; H. Bart van der Worp[17]; Rob H. Lo[18]

Adverse event committee

Robert J. van Oostenbrugge[5] (chair); Jeannette Hofmeijer[11]; H. Zwenneke Flach[23]

Trial methodologist

Hester F. Lingsma[40]

Research nurses/local trial coordinators

Naziha el Ghannouti[1]; Martin Sterrenberg[1]; Corina Puppels[7]; Wilma Pellikaan[7]; Rita Sprengers[4]; Marjan Elfrink[11]; Michelle Simons[11]; Marjolein Vossers[12]; Joke de Meris[14]; Tamara Vermeulen[14]; Annet Geerlings[19]; Gina van Vemde[22]; Tiny Simons[30]; Cathelijn van Rijswijk[21]; Gert Messchendorp[28]; Nynke Nicolaij[28]; Hester Bongenaar[32]; Karin Bodde[24]; Sandra Kleijn[34]; Jasmijn Lodico[34]; Hanneke Droste[34]; Maureen Wollaert[5]; Sabrina Verheesen[5]; D. Jeurrissen[5]; Erna Bos[9]; Yvonne Drabbe[15]; Michelle Sandiman[15]; Marjan Elfrink[11]; Nicoline Aaldering[11]; Berber Zweedijk[17]; Mostafa Khalilzada[15]; Jocova Vervoort[21]; Hanneke Droste[34]; Nynke Nicolaij[2]; Michelle Simons[11]; Eva Ponjee[22]; Sharon Romviel[19]; Karin Kanselaar[19]; Erna Bos[9]; Denn Barning[10].

PhD/Medical students

Esmee Venema[40]; Vicky Chalos[1,40]; Ralph R. Geuskens[3]; Tim van Straaten[19]; Saliha Ergezen[1]; Roger R. M. Harmsma[1]; Daan Muijres[1]; Anouk de Jong[1]; Olvert A. Berkhemer[1,3,6]; Anna M. M. Boers[3,39]; J. Huguet[3]; P. F. C. Groot[3]; Marieke A. Mens[3]; Katinka R. van Kranendonk[3]; Kilian M. Treurniet[3]; Ivo G. H. Jansen[3]; Manon L. Tolhuisen[3,39]; Heitor Alves[3]; Annick J. Weterings[3], Eleonora L. F. Kirkels[3], Eva J. H. F. Voogd[11]; Lieve M. Schupp[3]; Sabine Collette[28,29]; Adrien E. D. Groot[4]; Natalie E. LeCouffe[4]; Praneeta R. Konduri[39]; Haryadi Prasetya[39]; Nerea Arrarte-Terreros[39]; Lucas A. Ramos[39].

List of affiliations

Department of Neurology[1], Radiology[2], Public Health[40], Erasmus MC University Medical Center;

Department of Radiology and Nuclear Medicine[3], Neurology[4], Biomedical Engineering & Physics[39], Amsterdam UMC, University of Amsterdam, Amsterdam;

Department of Neurology[5], Radiology[6], Maastricht University Medical Center and Cardiovascular Research Institute Maastricht (CARIM);

Department of Neurology[7], Radiology[8], Sint Antonius Hospital, Nieuwegein;

Department of Neurology[9], Radiology[10], Leiden University Medical Center;

Department of Neurology[11], Radiology[12], Rijnstate Hospital, Arnhem;

Department of Radiology[13], Neurology[14], Haaglanden MC, the Hague;

Department of Neurology[15], Radiology[16], HAGA Hospital, the Hague;

Department of Neurology[17], Radiology[18], University Medical Center Utrecht;

Department of Neurology[19], Neurosurgery[20], Radiology[27], Radboud University Medical Center, Nijmegen;

Department of Neurology[21], Radiology[26], Elisabeth-TweeSteden ziekenhuis, Tilburg;
Department of Neurology[22], Radiology[23], Isala Klinieken, Zwolle;
Department of Neurology[24], Radiology[25], Reinier de Graaf Gasthuis, Delft;
Department of Neurology[28], Radiology[29], University Medical Center Groningen;
Department of Neurology[30], Radiology[31], Atrium Medical Center, Heerlen;
Department of Neurology[32], Radiology[33], Catharina Hospital, Eindhoven;
Department of Neurology[34], Radiology[35], Medical Spectrum Twente, Enschede;
Department of Radiology[36], Amsterdam UMC, Vrije Universiteit van Amsterdam, Amsterdam;Department of Radiology[37], Noordwest Ziekenhuisgroep, Alkmaar;
Department of Radiology[38], Texas Stroke Institute, Texas.

ARTICLE INFORMATION

Received September 1, 2022; accepted January 25, 2023.

Affiliations

Department of Public Health, Erasmus University Medical Center, Rotterdam, The Netherlands (M.A., H.F.L., S.J.d.H., N.v.L.); Erasmus School of Health Policy and Management, Erasmus University Rotterdam, Rotterdam, The Netherlands (F.E.); Department of Radiology and Nuclear Medicine (S.J.d.H., B.R.) and Department of Neurology (S.J.d.H., B.R., D.D.), Erasmus University Medical Center, Rotterdam, The NetherlandsDepartment of Neurology, Maastricht University Medical Center and School for Cardiovascular Diseases, Maastricht, The Netherlands (S.G.H.O., W.v.Z.); Department of Radiology, Rijnstate, Arnhem, The Netherlands (J.M.); Department of Neurology and Neurosurgery, Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands (B.v.d.W.); Department of Radiology and Nuclear Medicine, Maastricht University Medical Center, Cardiovascular Research Institute Maastricht, Maastricht, The Netherlands (W.v.Z.); Department of Radiology, Medical Imaging Center, University Medical Center Groningen, Groningen, The Netherlands (A.v.d.H.); and Department of Radiology and Nuclear Medicine, Amsterdam University Medical Center, Amsterdam, The Netherlands (S.D.R.).

Acknowledgments

We thank the MR CLEAN Registry Investigators–group authors.

Sources of Funding

No funding was received in support of this study. The MR CLEAN Registry was partly funded by TWIN Foundation, Erasmus MC University Medical Center, Maastricht University Medical Center, and Amsterdam UMC.

Disclosures

None.

Supplemental Material

Figures S1–S3

REFERENCES

1. Lingsma HF, Steyerberg EW, Eijkemans MJ, Dippel DW, Scholte Op Reimer WJ, Van Houwelingen HC, Netherlands stroke survey investigators. Comparing and ranking hospitals based on outcome: results from The Netherlands Stroke Survey. *QJM*. 2010;103:99–108. doi: 10.1093/qjmed/hcp169
2. Tsai TC, Joynt KE, Orav EJ, Gawande AA, Jha AK. Variation in surgical-readmission rates and quality of hospital care. *N Engl J Med*. 2013;369:1134–1142. doi: 10.1056/NEJMsa1303118
3. Pringle M, Wilson T, Grol R. Measuring "goodness" in individuals and healthcare systems. *BMJ*. 2002;325:704–707. doi: 10.1136/bmj.325.7366.704
4. Arah OA, Klazinga NS, Delnoij DM, ten Asbroek AH, Custers T. Conceptual frameworks for health systems performance: a quest for effectiveness, quality, and improvement. *Int J Qual Health Care*. 2003;15:377–398. doi: 10.1093/intqhc/mzg049
5. Krumholz HM, Wang Y, Mattera JA, Wang Y, Han LF, Ingber MJ, Roman S, Normand SLT. An administrative claims model suitable for profiling hospital performance based on 30-day mortality rates among patients with an acute myocardial infarction. *Circulation*. 2006;113:1683–1692. doi: 10.1161/CIRCULATIONAHA.105.611186
6. Stein LK, Mocco J, Fifi J, Jette N, Tuhrim S, Dhamoon MS. Correlations between physician and hospital stroke thrombectomy volumes and outcomes: a nationwide analysis. *Stroke*. 2021;52:2858–2865. doi: 10.1161/STROKEAHA.120.033312
7. Olthuis SGH, den Hartog SJ, van Kuijk SMJ, Staals J, Benali F, van der Leij C, Beumer D, Lycklama à Nijeholt GJ, Uyttenboogaart M, Martens JM, et al. Influence of the interventionist's experience on outcomes of endovascular thrombectomy in acute ischemic stroke: results from the MR CLEAN Registry. *J Neurointerv Surg*. 2022:neurintsurg-2021-018295
8. Lingsma HF, Dippel DW, Hoeks SE, Steyerberg EW, Franke CL, van Oostenbrugge RJ, de Jong G, Simoons ML, Op Reimer WJMS, Netherlands Stroke Survey investigators. Variation between hospitals in patient outcome after stroke is only partly explained by differences in quality of care: results from The Netherlands Stroke Survey. *J Neurol Neurosurg Psychiatry*. 2008;79:888–894. doi: 10.1136/jnnp.2007.137059
9. Amini M, van Leeuwen N, Eijkenaar F, Mulder MJHL, Schonewille W, Lycklama à Nijeholt G, Hinsenveld WH, Goldhoorn RJB, van Doormaal PJ, Jenniskens S, et al. Improving quality of stroke care through benchmarking center performance: why focusing on outcomes is not enough. *BMC Health Serv Res*. 2020;20:998. doi: 10.1186/s12913-020-05841-y
10. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke*. 1988;19:604–607. doi: 10.1161/01.STR.19.5.604
11. Kuhrij LS, Wouters MW, van den Berg-Vos RM, de Leeuw FE, Nederkoorn PJ. The Dutch Acute Stroke Audit: benchmarking acute stroke care in The Netherlands. *Eur Stroke J*. 2018;3:361–368. doi: 10.1177/2396987318787695
12. Goyal M, Menon BK, van Zwam WH, Dippel DW, Mitchell PJ, Demchuk AM, Antoni D, Majoie CBLM, van der Lugt A, de Miquel MA, et al. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet*. 2016;387:1723–1731. doi: 10.1016/S0140-6736(16)00163-X
13. Saver JL, Altman H. Relationship between neurologic deficit severity and final functional outcome shifts and strengthens during first hours after onset. *Stroke*. 2012;43:1537–1541. doi: 10.1161/STROKEAHA.111.636928
14. Jansen IGH, Mulder M, Goldhoorn RB, MR CLEAN Registry Investigators. Endovascular treatment for acute ischaemic stroke in routine clinical practice: prospective, observational cohort study (MR CLEAN Registry). *BMJ*. 2018;360:k949. doi: 10.1136/bmj.k949
15. Alegiani AC, Dorn F, Herzberg M, Wollenweber FA, Kellert L, Siebert E, Nolte CH, von Rennenberg R, Hattingen E, Petzold GC, et al. Systematic evaluation of stroke thrombectomy in clinical practice: the German Stroke Registry Endovascular Treatment. *Int J Stroke*. 2019;14:372–380. doi: 10.1177/1747493018806199
16. Spertus JA, Eagle KA, Krumholz HM, Mitchell KR, Normand SL. American College of Cardiology and American Heart Association methodology for the selection and creation of performance measures for quantifying the quality of cardiovascular care. *Circulation*. 2005;111:1703–1712. doi: 10.1161/01.CIR.0000157096.95223.D7
17. Noser EA, Shaltoni HM, Hall CE, Alexandrov AV, Garami Z, Cacayorin ED, Song JK, Grotta JC, Campbell MS. Aggressive mechanical clot disruption: a safe adjunct to thrombolytic therapy in acute stroke? *Stroke*. 2005;36:292–296. doi: 10.1161/01.STR.0000152331.93770.18
18. Abdul-Rahim AH, Fulton RL, Sucharew H, Kleindorfer D, Khatri P, Broderick JP, Lees KR, SITS-MOST Steering Committee. National Institutes of Health Stroke Scale item profiles as predictor of patient outcome. *Stroke*. 2015;46:395–400. doi: 10.1161/STROKEAHA.114.006837
19. Chalos V, van der Ende NAM, Lingsma HF, Mulder M, Venema E, Dijkland SA, Berkhemer OA, Yoo AJ, Broderick JP, Palesch YY, et al. National Institutes of Health Stroke Scale: an alternative primary

outcome measure for trials of acute treatment for ischemic stroke. *Stroke*. 2020;51:282–290. doi: 10.1161/STROKEAHA.119.026791
20. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen E, Tanabe K, Kitagawa G, eds. *Selected Papers of Hirotugu Akaike. Springer Series in Statistics*. New York, NY: Springer; 1998:199–213. doi: 10.1007/978-1-4612-1694-0_15
21. Little RJA. A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc*. 1988;83:1198–1202. doi: 10.1080/01621459.1988.10478722
22. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. *Stat Med*. 1991;10:585–598. doi: 10.1002/sim.4780100410
23. Palmer RH. Using health outcomes data to compare plans, networks and providers. *Int J Qual Health Care*. 1998;10:477–483. doi: 10.1093/intqhc/10.6.477
24. Streiner DL, Norman GR, Cairney J. *Health Measurement Scales: A Practical Guide to their Development and Use*. Oxford Medical Publications; 2008. doi: 10.1093/acprof:oso/9780199231881.001.0001
25. Walsh K, Gompertz PH, Rudd AG. Stroke care: how do we measure quality? *Postgrad Med J*. 2002;78:322–326. doi: 10.1136/pmj.78.920.322
26. Aiken LH, Clarke SP, Cheung RB, Sloane DM, Silber JH. Educational levels of hospital nurses and surgical patient mortality. *JAMA*. 2003;290:1617–1623. doi: 10.1001/jama.290.12.1617
27. Behme D, Gera RG, Tsogkas I, Colla R, Liman J, Maier IL, Liebeskind DS, Psychogios MN. Impact of time on thrombolysis in cerebral infarction score results. *Clin Neuroradiol*. 2020;30:345–353. doi: 10.1007/s00062-019-00786-0
28. Steyerberg EW, van Veen M. Imputation is beneficial for handling missing data in predictive models. *J Clin Epidemiol*. 2007;60:979. doi: 10.1016/j.jclinepi.2007.03.003
29. van der Heijden GJ, Donders AR, Stijnen T, Moons KG. Imputation of missing values is superior to complete case analysis and the missing-indicator method in multivariable diagnostic research: a clinical example. *J Clin Epidemiol*. 2006;59:1102–1109. doi: 10.1016/j.jclinepi.2006.01.015
30. Rosenman R, Tennekoon V, Hill LG. Measuring bias in self-reported data. *Int J Behav Healthc Res*. 2011;2:320–332. doi: 10.1504/IJBHR.2011.043414

SUPPLEMENTAL MATERIAL

Figure S1. Specialized EVT hospitals in the Netherlands.

Figure S2. Flowchart of patient selection in the MR CLEAN Registry.