Anesthesia Technique and Outcomes of Mechanical Thrombectomy in Patients With Acute Ischemic Stroke

Kimon Bekelis, MD; Symeon Missios, MD; Todd A. MacKenzie, PhD;
Stavropoula Tjoumakaris, MD; Pascal Jabbour, MD

Background and Purpose—The impact of anesthesia technique on the outcomes of mechanical thrombectomy for acute ischemic stroke remains an issue of debate. We investigated the association of general anesthesia with outcomes in patients undergoing mechanical thrombectomy for ischemic stroke.

Methods—We performed a cohort study involving patients undergoing mechanical thrombectomy for ischemic stroke from 2009 to 2013, who were registered in the New York Statewide Planning and Research Cooperative System database. An instrumental variable (hospital rate of general anesthesia) analysis was used to simulate the effects of randomization and investigate the association of anesthesia technique with case-fatality and length of stay.

Results—Among 1174 patients, 441 (37.6%) underwent general anesthesia and 733 (62.4%) underwent conscious sedation. Using an instrumental variable analysis, we identified that general anesthesia was associated with a 6.4% increased case-fatality (95% confidence interval, 1.9%–11.0%) and 8.4 days longer length of stay (95% confidence interval, 2.9–14.0) in comparison to conscious sedation. This corresponded to 15 patients needing to be treated with conscious sedation to prevent 1 death. Our results were robust in sensitivity analysis with mixed effects regression and propensity score–adjusted regression models.

Conclusions—Using a comprehensive all-payer cohort of acute ischemic stroke patients undergoing mechanical thrombectomy in New York State, we identified an association of general anesthesia with increased case-fatality and length of stay. These considerations should be taken into account when standardizing acute stroke care.
(***Stroke***. 2017;48:361-366. DOI: 10.1161/STROKEAHA.116.015343.)

Key Words: acute ischemic stroke ■ general anesthesia ■ instrumental variable ■ mechanical thrombectomy ■ SPARCS

The evolution of mechanical thrombectomy has revolutionized the treatment of acute ischemic stroke. Despite initial reservations for first-generation devices,[1,2] subsequent clinical trials[3–7] have demonstrated that newer clot retrievers are associated with improved mortality and functional outcomes in appropriately selected patients. The efficacy of this intervention is contingent on timely revascularization of the occluded vessels.[8] Significant efforts are currently underway to optimize emergency medical services associated with stroke, streamline transfers, centralized care, and minimized door-to-needle time. However, most of these initiatives focus on either prehospital care pathways or in-hospital protocols aimed at expediting patient transfer to the angiography suite. Although the method of anesthesia has significant timing implications, the understanding of its impact on the outcomes of mechanical thrombectomy is limited.[9–12] General anesthesia is the preferred method because of the perceptions of improved procedural safety and efficacy.[9] However, conscious sedation, or no sedation, allows continuous neurological monitoring, minimal delays, and potentially improved hemodynamic stability.[10]

Previous observational studies attempting to answer this question have shown mixed results.[13–21] The main limitation of such investigations is not accounting for unmeasured confounding. Patients included in prior retrospective studies have been selected for either anesthesia technique in advance. This selection reflects the different preferences and background of the treating physicians, as well as specific patient characteristics. Administrative databases lack such granularity, limiting the ability to control for these confounders. There has been no prior study attempting to account for these limitations through different analytic approaches in an adult cohort of all ages.

We used the New York Statewide Planning and Research Cooperative System (SPARCS)[22] to study the association of

---

© 2017 American Heart Association, Inc.

Stroke is available at http://stroke.ahajournals.org　　　　　　　　　　　　DOI: 10.1161/STROKEAHA.116.015343

general anesthesia with case-fatality and length of stay (LOS) for patients undergoing mechanical thrombectomy for acute ischemic stroke. An instrumental variable analysis was used to control for unmeasured confounding and simulate the effect of randomization.

Methods

New York Statewide Planning and Research Cooperative System

This study was approved by the Dartmouth Committee for Protection of Human Subjects. All patients undergoing mechanical thrombectomy for acute ischemic stroke who were registered in the SPARCS (New York State Department of Health, Albany, NY)[22] database between 2009 and 2013 were included in the analysis. For these years, SPARCS contains patient-level details for every hospital discharge, ambulatory surgery, and emergency department admission in New York State as coded from admission and billing records. More information about SPARCS is available at https://www.health.ny.gov/statistics/sparcs/.

Cohort Definition

To establish the cohort of patients, we used *International Classification of Disease-9-Clinical Modification* codes to identify patients in the database who underwent mechanical thrombectomy (*International Classification of Disease-9-Clinical Modification* code 39.7) for acute ischemic stroke (*International Classification of Disease-9-Clinical Modification* code 433.x1, 434.x1) between 2009 and 2013.

Outcome Variables

The primary outcome variable was case-fatality during the initial hospitalization after mechanical thrombectomy for ischemic stroke. Secondary outcome was LOS during the initial hospitalization.

Exposure Variables

The primary exposure variable was the anesthesia technique (general versus conscious sedation). General anesthesia required the patient being intubated, whereas conscious sedation included intravenous sedation (with or without local anesthetic) without the use of a breathing tube. Hospitals voluntarily report data on anesthesia care to SPARCS.

Covariates (Table I in the online-only Data Supplement) used for risk adjustment were age, sex, race (Black, Hispanic, Asian, White, other), insurance (private, Medicare, Medicaid, uninsured, other), and administration of intravenous tissue-type plasminogen activator (*International Classification of Disease-9-Clinical Modification* 99.10, V45.88). The comorbidities used for risk adjustment were diabetes mellitus, smoking, chronic lung disease, hypertension, hypercholesterolemia, peripheral vascular disease, congestive heart failure, coronary artery disease, history of transient ischemic attack, alcohol abuse, obesity, chronic renal failure, and coagulopathy. Only variables that were defined as present on admission were considered part of the patient's preadmission comorbidity profile.

Statistical Analysis

The association of anesthesia technique with our outcome measures was examined in a multivariable setting. Patients undergoing general anesthesia or conscious sedation in our cohort were nonrandomly selected for either technique based on provider and patient characteristics. To account for this unmeasured confounding and to simulate the effect of randomization, we used an instrumental variable analysis, an econometric technique (Methods in the online-only Data Supplement).[23] The regional rate of general anesthesia (hospital-level general anesthesia rate) was used as an instrument for the technique received. This advanced observational technique has been used before by clinical researchers to answer comparative effectiveness questions for different interventions. The goal is to simulate randomization, especially when the baseline functional characteristics of the patients (including the functional status

of stroke patients, National Institutes of Health Stroke Scale etc) are unknown (similar to our application). This is an established technique in prior literature for ischemic stroke, anesthesia technique, and other pathologies when several variables are missing from the data set.[24–28]

A good instrument is not associated with the outcome other than through the exposure variable of interest (a requirement known as the exclusion restriction criterion).[29] In our case, it is unlikely that the regional rates of general anesthesia would be associated with case-fatality in any way other than the choice of treatment. A 2-stage least squares method was used for the calculation of the coefficients. The value of the F statistic in the first stage of the 2-stage least squares approach was 125, which is consistent with a strong instrument (F statistic>10), based on a practical rule.[23]

A probit regression was used for the categorical outcomes (case-fatality)[30] and a linear regression for the linear outcomes (LOS). Other models such as general logit were considered. However, we elected to use probit because this is the most widely used and studied model in instrumental variable analysis.[31,32] The covariates used for risk adjustment in these models were age, sex, race, insurance, and all the comorbidities mentioned previously. Because the coefficients produced by the probit function are not interpretable, we used the marginal effects of our independent variables instead. The marginal effects are the partial derivatives of the coefficients and reflect the change in the probability of the dependent variable, for 1 U change in the independent variable, at the average value of all other covariates.

To demonstrate the robustness of our data in a sensitivity analysis, we used standard techniques to account for measured confounding, while accounting for clustering at the hospital level. For categorical outcomes, we used a probit regression model with hospital identity as a random effects variable, while controlling for all the covariates mentioned previously. For continuous outcomes, we performed similar analyses using linear models. LOS demonstrated a positively skewed distribution, and a logarithmic transformation was additionally used in sensitivity analysis. The direction of the observed associations did not change, and therefore, we elected to present the untransformed data for ease of interpretation. In an alternative way to control for confounding for categorical outcomes, we used a propensity-adjusted (with deciles of propensity score) probit regression model. We calculated the propensity score of general anesthesia with a separate probit regression model, using all the covariates mentioned previously. The results were identical (Table II in the online-only Data Supplement). In post hoc analyses, we investigated the association of anesthesia technique with the risks of pneumonia or hemorrhagic transformation using the previously specified instrumental variable analysis.

Regression diagnostics were used for all models. Number needed to treat was calculated when appropriate. All results are based on 2-sided tests, and the level of statistical significance was set at 0.05. This study, based on 1174 patients, has sufficient power (80%) at a 5% type I error rate to detect differences in case-fatality as small as 7.8%. Statistical analyses were performed using Stata version 13 (StataCorp, College Station, TX).

Results

Patient Characteristics

In the selected study period, there were 1174 patients undergoing mechanical thrombectomy for acute ischemic stroke (mean age was 67.3 years, with 52.4% females) who were registered in SPARCS. Four hundred and forty-one (37.6%) patients underwent general anesthesia, and 733 (62.4%) patients underwent conscious sedation. The characteristics of the 2 cohorts at baseline can be seen in Table 1.

Inpatient Case-Fatality

Overall, 126 (25.6%) inpatient deaths were recorded after general anesthesia and 147 (18.1%) after conscious sedation. General anesthesia was associated with increased case-fatality

Table 1. Patient Characteristics

	Total (N=1174)	General Anesthesia (N=441)	Conscious Sedation (N=733)
Age, mean (SD)	67.3 (15.0)	66.5 (15.2)	67.6 (14.9)
Female sex, n (%)	615 (52.4)	219 (49.7)	394 (53.8)
Race			
White, n (%)	782 (66.6)	309 (70.1)	522 (71.2)
Black, n (%)	149 (12.7)	57 (12.9)	83 (11.3)
Hispanic, n (%)	143 (12.2)	40 (9.0)	73 (10.0)
Asian, n (%)	67 (5.7)	21 (4.8)	38 (5.2)
Others, n (%)	33 (2.8)	14 (3.2)	17 (2.3)
Insurance			
Medicare, n (%)	633 (53.9)	233 (52.8)	400 (55.3)
Private insurance, n (%)	419 (35.7)	156 (35.4)	263 (34.9)
Medicaid, n (%)	76 (6.5)	31 (7.0)	45 (6.3)
Uninsured, n (%)	26 (2.2)	15 (3.4)	11 (1.6)
Others, n (%)	20 (1.7)	…*	…*
IV tPA, n (%)	613 (52.2)	219 (49.7)	394 (53.8)
Comorbidities			
Diabetes mellitus, n (%)	289 (24.6)	105 (23.8)	183 (25.0)
Smoking, n (%)	151 (12.9)	40 (9.1)	104 (14.2)
Obesity, n (%)	97 (8.3)	9 (8.8)	59 (8.0)
Transient ischemic attack, n (%)	…*	…*	…*
Coronary artery disease, n (%)	322 (27.4)	125 (28.3)	206 (28.1)
Congestive heart failure, n (%)	313 (26.7)	113 (25.6)	199 (27.1)
Chronic lung disease, n (%)	151 (12.9)	62 (14.1)	97 (13.2)
Coagulopathy, n (%)	48 (4.1)	23 (5.2)	27 (3.7)
Chronic renal failure, n (%)	114 (9.7)	37 (8.4)	78 (10.6)
Hypertension, n (%)	791 (67.4)	297 (67.3)	497 (67.8)
Hypercholesterolemia, n (%)	475 (40.5)	152 (34.5)	335 (45.7)
Alcohol, n (%)	38 (3.2)	…*	…*
Peripheral vascular disease, n (%)	104 (8.9)	32 (7.3)	73 (10.0)

IV tPA indicates intravenous tissue-type plasminogen activator; SD, standard deviation; and SPARCS, Statewide Planning and Research Cooperative System.

*Output suppressed to respect the SPARCS reporting limit of 11 patients.

in comparison to conscious sedation (difference, 7.2%; 95% confidence interval [CI], 2.6%–12.0%) in unadjusted analysis. Likewise, using a probit regression with instrumental variable analysis, we identified that general anesthesia was associated with a 6.4% increased case-fatality (95% CI, 1.9%–11.0%), in comparison to conscious sedation (Table 2). This persisted in a mixed effects probit regression model (adjusted difference, 7.5%; 95% CI, 2.9%–12.1%). This corresponded to 15 patients needed to be treated with conscious sedation to prevent one death.

Length of Stay
The average LOS was 19.6 days (SD 35.0) after general anesthesia and 11.7 days (SD 12.5) after conscious sedation. General anesthesia was associated with increased LOS in

comparison to conscious sedation (difference, 7.9; 95% CI, 5.1–10.7) in the unadjusted analysis. Using a linear regression with instrumental variable analysis, we demonstrated (Table 2) that general anesthesia was associated with 8.4 days longer LOS in comparison to conscious sedation (95% CI, 2.9–14.0). We found similar results in a mixed effects linear regression model (adjusted difference, 7.3; 95% CI, 4.6–10.1).

Post Hoc Analyses
In post hoc analysis, using an instrumental variable analysis, anesthesia technique was not associated with the risk of pneumonia (adjusted difference, 2.3%; 95% CI, −1.2% to 5.7%) or hemorrhagic transformation (adjusted difference, 1.5%; 95% CI, −2.4% to 5.6%).

Table 2. Models Examining the Association of General Anesthesia With Outcomes

	Inpatient Mortality,%*	P Value	Length of Stay†	P Value
Unadjusted analysis, difference (95% CI)	7.2 (2.6–12.0)	<0.001	7.9 (5.1–10.7)	<0.001
Instrumental variable analysis, adjusted difference (95% CI)‡	6.4 (1.9–11.0)	<0.001	8.4 (2.9–14.0)	<0.001
Mixed effects regression, adjusted difference (95% CI)§	7.5 (2.9–12.1)	<0.001	7.3 (4.6–10.1)	<0.001

CI indicates confidence intervals.

*All regressions were based on probit models.

†All regressions were based on linear models.

‡Hospital-level general anesthesia rate was used as an instrument of anesthesia technique.

§Hospital identity was used as a random effects variable.

Discussion

Using a comprehensive all-payer cohort of patients in New York State with acute ischemic stroke, we identified an association of general anesthesia with increased case-fatality and LOS after mechanical thrombectomy. Our results were robust when considering several advanced observational techniques to account for measured and unmeasured confounders. Mechanical thrombectomy has seen explosive growth in recent years and is currently performed by multiple specialties, without standardized perioperative protocols, including anesthesia choice.[11,12] This is contributing to an ongoing debate about the relative effectiveness of different anesthesia techniques during mechanical thrombectomy for acute ischemic stroke.[9,10]

Several observational studies have compared the outcomes of general anesthesia and conscious sedation for this population. Most of the studies have been retrospective analyses of single institution experiences, demonstrating conflicting results with limited generalization, given their inherent selection bias.[15–18,20,21,33] The interpretation of larger multicenter studies is equally limited. McDonald et al,[19] using the MarketScan database, demonstrated that conscious sedation is associated with a survival benefit. They used propensity score matching to balance the covariates among anesthesia groups. However, participation in this commercial database was voluntary, and therefore, it is likely that hospitals incentivized to achieve higher quality standards would be overrepresented. This self-selection introduces significant unmeasured confounding, which the authors did not account for. In another study, Abou-Chebl et al[14] analyzed a national registry of a single thrombectomy device and demonstrated superior results for conscious sedation. The generalizability of their results in the population at large is limited only to the device used and the few hospitals incentivized to participate in this national registry. A post hoc analysis of the anesthesia method used in a randomized controlled trial, examining the effectiveness of mechanical thrombectomy, demonstrated that the benefit of the trial was only seen with conscious sedation.[34] This analysis had the same bias as all retrospective designs. Three

randomized trials are currently underway, specifically looking into the comparative effectiveness of different anesthesia techniques.

These prior analyses have some common methodologic limitations. Multicenter studies are vulnerable to clustering at the hospital level. Most previous authors did not evaluate or adjust for this bias. Most importantly, all the analytical methods used accounted, to some degree, for known confounders. Although this may be adequate in some studies, the selection of patients for either anesthesia technique before the analysis introduces significant unmeasured confounding. Patients may be selected for general anesthesia because of worse functional status, more severe stroke, or heavier comorbidity burden. Physician or patient preference, as well as provider training and specialty, might affect that decision too. Not accounting for this dimension of confounding puts the robustness of their findings into question.

Our study purposefully addresses many of these methodologic limitations. First, we created a cohort of all patients in a major state, giving a true picture of practice in the community. Second, we used advanced observational techniques to control for confounding. Propensity score stratification was used to adjust our analyses for known confounders. The possibility of clustering, which can bias the results of multicenter national studies, was accounted for by using mixed effects methods. Most importantly, an instrumental variable analysis was used to control for unmeasured confounders (mainly the a priori selection of anesthesia technique) and simulate the effects of randomization. The instrumental variable analysis is expected to control for such factors and report results for patients of similar functional status. Results were consistent across techniques, supporting the validity of the observed associations.

Further research into the factors contributing to the superiority of conscious sedation for this pathology is warranted. Theoretical reasons include the neurotoxicity of certain anesthetic agents,[35] the lack of continuous neurological assessment during general anesthesia,[13] and perioperative hypotension.[36] In the latter case, the induction of general anesthesia commonly leads to a reduction in blood pressure, which may be associated with worse outcomes.[36]

Our study has several limitations. Residual confounding could account for some of the observed associations. However, this is minimized to the extent that we are using a good instrument for anesthesia technique. The F statistic in our analysis suggests a strong instrument. In addition, coding inaccuracies will undoubtedly occur and can affect our estimates. However, several reports have demonstrated that coding for stroke has shown nearly perfect association with medical record review.[37,38] Although SPARCS includes all hospitals from the entire New York State, the generalization of this analysis to the US population is uncertain. SPARCS does not provide any clinical information on the functional status of the patients (National Institutes of Health Stroke Scale), which can affect the choice of anesthesia. However, the use of the instrumental variable analysis is attempting to control for unknown confounders such as these and has been used before in stroke patients of this database.[24–26] By comparing our point estimates for the instrumental variable model and

the multivariable model without instrumental variable, we can identify a small difference. That indicates that there is likely a small degree of selection bias in our data before the application of the instrumental variable analysis.

Additionally, we were lacking posthospitalization and long-term data on our patients. Quality metrics (ie, modified Rankin score) are also not available through SPARCS, and therefore, we cannot compare the 2 treatment techniques on these outcomes. The definitive comparison of the 2 techniques on functional outcomes can only be done in prospective registries. In this direction, the NeuroPoint Alliance has created the first module for a cerebrovascular registry, with results expected in the near future.[39] We do not have any information on the availability of neuroanesthesia in the institutions we are studying. Our results might reflect to some degree the difference between neuroanesthesia and conscious sedation, although it is likely that some centers might preferentially offer conscious sedation, despite the availability of neuroanesthesia services. Finally, causality cannot be definitively established based on observational data, despite the use of advanced techniques, such as the instrumental variable analysis.

Conclusions

The impact of anesthesia techniques on the outcomes of mechanical thrombectomy for acute ischemic stroke remains an issue of debate. Using a comprehensive all-payer cohort of patients in New York State with acute ischemic stroke, we identified an association of general anesthesia with increased case-fatality and LOS after mechanical thrombectomy. Our results were robust when considering several advanced observational techniques to account for measured and unmeasured confounders. These considerations should be taken into account when standardizing acute stroke care.

Sources of Funding

Disclosures

None.

References

1. Broderick JP, Palesch YY, Demchuk AM, Yeatts SD, Khatri P, Hill MD, et al; Interventional Management of Stroke (IMS) III Investigators. Endovascular therapy after intravenous t-PA versus t-PA alone for stroke. *N Engl J Med*. 2013;368:893–903. doi: 10.1056/NEJMoa1214300.
2. Ciccone A, Valvassori L, Nichelatti M, Sgoifo A, Ponzio M, Sterzi R, et al; SYNTHESIS Expansion Investigators. Endovascular treatment for acute ischemic stroke. *N Engl J Med*. 2013;368:904–913. doi: 10.1056/NEJMoa1213701.
3. Berkhemer OA, Fransen PS, Beumer D, van den Berg LA, Lingsma HF, Yoo AJ, et al; MR CLEAN Investigators. A randomized trial of intraarterial treatment for acute ischemic stroke. *N Engl J Med*. 2015;372:11–20. doi: 10.1056/NEJMoa1411587.
4. Campbell BC, Mitchell PJ, Kleinig TJ, Dewey HM, Churilov L, Yassi N, et al; EXTEND-IA Investigators. Endovascular therapy for ischemic stroke with perfusion-imaging selection. *N Engl J Med*. 2015;372:1009–1018. doi: 10.1056/NEJMoa1414792.
5. Saver JL, Goyal M, Bonafe A, Diener HC, Levy EI, Pereira VM, et al; SWIFT PRIME Investigators. Stent-retriever thrombectomy after intravenous t-PA vs. t-PA alone in stroke. *N Engl J Med*. 2015;372:2285–2295. doi: 10.1056/NEJMoa1415061.
6. Goyal M, Demchuk AM, Menon BK, Eesa M, Rempel JL, Thornton J, et al; ESCAPE Trial Investigators. Randomized assessment of rapid endovascular treatment of ischemic stroke. *N Engl J Med*. 2015;372:1019–1030. doi: 10.1056/NEJMoa1414905.
7. Jovin TG, Chamorro A, Cobo E, de Miquel MA, Molina CA, Rovira A, et al; REVASCAT Trial Investigators. Thrombectomy within 8 hours after symptom onset in ischemic stroke. *N Engl J Med*. 2015;372:2296–2306. doi: 10.1056/NEJMoa1503780.
8. Goyal M, Menon BK, van Zwam WH, Dippel DW, Mitchell PJ, Demchuk AM, et al; HERMES Collaborators. Endovascular thrombectomy after large-vessel ischaemic stroke: a meta-analysis of individual patient data from five randomised trials. *Lancet*. 2016;387:1723–1731. doi: 10.1016/S0140-6736(16)00163-X.
9. Brekenfeld C, Mattle HP, Schroth G. General is better than local anesthesia during endovascular procedures. *Stroke*. 2010;41:2716–2717. doi: 10.1161/STROKEAHA.110.594622.
10. Gupta R. Local is better than general anesthesia during endovascular acute stroke interventions. *Stroke*. 2010;41:2718–2719. doi: 10.1161/STROKEAHA.110.596015.
11. McDonagh DL, Olson DM, Kalia JS, Gupta R, Abou-Chebl A, Zaidat OO. Anesthesia and sedation practices among neurointerventionalists during acute ischemic stroke endovascular therapy. *Front Neurol*. 2010;1:118. doi: 10.3389/fneur.2010.00118.
12. Molina CA, Selim MH. General or local anesthesia during endovascular procedures: sailing quiet in the darkness or fast under a daylight storm. *Stroke*. 2010;41:2720–2721. doi: 10.1161/STROKEAHA.110.595447.
13. Abou-Chebl A, Lin R, Hussain MS, Jovin TG, Levy EI, Liebeskind DS, et al. Conscious sedation versus general anesthesia during endovascular therapy for acute anterior circulation stroke: preliminary results from a retrospective, multicenter study. *Stroke*. 2010;41:1175–1179. doi: 10.1161/STROKEAHA.109.574129.
14. Abou-Chebl A, Zaidat OO, Castonguay AC, Gupta R, Sun CH, Martin CO, et al. North American SOLITAIRE Stent-Retriever Acute Stroke Registry: choice of anesthesia and outcomes. *Stroke*. 2014;45:1396–1401. doi: 10.1161/STROKEAHA.113.003698.
15. Hassan AE, Chaudhry SA, Zacharatos H, Khatri R, Akbar U, Suri MF, et al. Increased rate of aspiration pneumonia and poor discharge outcome among acute ischemic stroke patients following intubation for endovascular treatment. *Neurocrit Care*. 2012;16:246–250. doi: 10.1007/s12028-011-9638-0.
16. Jumaa MA, Zhang F, Ruiz-Ares G, Gelzinis T, Malik AM, Aleu A, et al. Comparison of safety and clinical and radiographic outcomes in endovascular acute stroke therapy for proximal middle cerebral artery occlusion with intubation and general anesthesia versus the nonintubated state. *Stroke*. 2010;41:1180–1184. doi: 10.1161/STROKEAHA.109.574194.
17. Langner S, Khaw AV, Fretwurst T, Angermaier A, Hosten N, Kirsch M. [Endovascular treatment of acute ischemic stroke under conscious sedation compared to general anesthesia - safety, feasibility and clinical and radiological outcome]. *Rofo*. 2013;185:320–327. doi: 10.1055/s-0032-1330361.
18. Li F, Deshaies EM, Singla A, Villwock MR, Melnyk V, Gorji R, et al. Impact of anesthesia on mortality during endovascular clot removal for acute ischemic stroke. *J Neurosurg Anesthesiol*. 2014;26:286–290. doi: 10.1097/ANA.0000000000000031.
19. McDonald JS, Brinjikji W, Rabinstein AA, Cloft HJ, Lanzino G, Kallmes DF. Conscious sedation versus general anaesthesia during mechanical thrombectomy for stroke: a propensity score analysis. *J Neurointerv Surg*. 2015;7:789–794. doi: 10.1136/neurintsurg-2014-011373.
20. Nichols C, Carrozzella J, Yeatts S, Tomsick T, Broderick J, Khatri P. Is periprocedural sedation during acute stroke therapy associated with poorer functional outcomes? *J Neurointerv Surg*. 2010;2:67–70. doi: 10.1136/jnis.2009.001768.
21. Sugg RM, Jackson AS, Holloway W, Martin CO, Akhtar N, Rymer M. Is mechanical embolectomy performed in nonanesthetized patients effective? *AJNR Am J Neuroradiol*. 2010;31:1533–1535. doi: 10.3174/ajnr.A2091.
22. New York States Department of Health. Statewide Planning and Research Cooperative System (SPARCS). 2015.
23. Staiger D, Stock JH. Instrumental variables regression with weak instruments. *Econometrica*. 1997;65:557–586.
24. Neuman MD, Rosenbaum PR, Ludwig JM, Zubizarreta JR, Silber JH. Anesthesia technique, mortality, and length of stay after hip fracture surgery. *JAMA*. 2014;311:2508–2517. doi: 10.1001/jama.2014.6499.

25. Tan HJ, Norton EC, Ye Z, Hafez KS, Gore JL, Miller DC. Long-term survival following partial vs radical nephrectomy among older patients with early-stage kidney cancer. *JAMA*. 2012;307:1629–1635. doi: 10.1001/jama.2012.475.

26. Xian Y, Holloway RG, Chan PS, Noyes K, Shah MN, Ting HH, et al. Association between stroke center hospitalization for acute ischemic stroke and mortality. *JAMA*. 2011;305:373–380. doi: 10.1001/jama.2011.22.

27. Bekelis K, Marth NJ, Wong K, Zhou W, Birkmeyer JD, Skinner J. Primary stroke center hospitalization for elderly patients with stroke: implications for case fatality and travel times. *JAMA Intern Med*. 2016;176:1361–1368. doi: 10.1001/jamainternmed.2016.3919.

28. Bekelis K, Gottlieb DJ, Su Y, O'Malley AJ, Labropoulos N, Goodney P, et al. Comparison of clipping and coiling in elderly patients with unruptured cerebral aneurysms. *J Neurosurg*. 2016:1–8.

29. Garabedian LF, Chu P, Toh S, Zaslavsky AM, Soumerai SB. Potential bias of instrumental variable analyses for observational comparative effectiveness research. *Ann Intern Med*. 2014;161:131–138. doi: 10.7326/M13-1887.

30. Foster EM. Instrumental variables for logistic regression: an illustration. *Soc Sci Res*. 1997;26:287–504.

31. Wooldridge JM. *Introductory Econometrics: A Modern Approach*. Mason, OH: Thomson/South-Western; 2006.

32. Rassen JA, Schneeweiss S, Glynn RJ, Mittleman MA, Brookhart MA. Instrumental variable analysis for estimation of treatment effects with dichotomous outcomes. *Am J Epidemiol*. 2009;169:273–284. doi: 10.1093/aje/kwn299.

33. Davis MJ, Menon BK, Baghirzada LB, Campos-Herrera CR, Goyal M, Hill MD, et al; Calgary Stroke Program. Anesthetic management and outcome in patients during endovascular therapy for acute stroke. *Anesthesiology*. 2012;116:396–405. doi: 10.1097/ALN.0b013e318242a5d2.

34. Berkhemer OA, van den Berg LA, Fransen PS, Beumer D, Yoo AJ, Lingsma HF, et al; MR CLEAN Investigators. The effect of anesthetic management during intra-arterial therapy for acute stroke in MR CLEAN. *Neurology*. 2016;87:656–664. doi: 10.1212/WNL.0000000000002976.

35. Messick JM Jr, Newberg LA, Nugent M, Faust RJ. Principles of neuroanesthesia for the nonneurosurgical patient with CNS pathophysiology. *Anesth Analg*. 1985;64:143–174.

36. Leonardi-Bee J, Bath PM, Phillips SJ, Sandercock PA; IST Collaborative Group. Blood pressure and clinical outcomes in the International Stroke Trial. *Stroke*. 2002;33:1315–1320.

37. Kokotailo RA, Hill MD. Coding of stroke and stroke risk factors using international classification of diseases, revisions 9 and 10. *Stroke*. 2005;36:1776–1781. doi: 10.1161/01.STR.0000174293.17959.a1.

38. Tirschwell DL, Longstreth WT Jr. Validating administrative data in stroke research. *Stroke*. 2002;33:2465–2470.

39. NeuroPoint Alliance. The National Neurosurgery Quality and Outcomes Database (n²qod). 2015.

Stroke

American Heart Association

American Stroke Association®

Anesthesia Technique and Outcomes of Mechanical Thrombectomy in Patients With Acute Ischemic Stroke

Kimon Bekelis, Symeon Missios, Todd A. MacKenzie, Stavropoula Tjoumakaris and Pascal Jabbour

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/48/2/361

Data Supplement (unedited) at:
http://stroke.ahajournals.org/content/suppl/2017/01/19/STROKEAHA.116.015343.DC1

SUPPLEMENTAL MATERIAL

1. <u>Scoring guidelines for the NIHSS</u>. This is "Chapter 5" from the Manual of Procedures for the original NINDS rt-PA for Acute Stroke Trial.
2. <u>The NIHSS.</u> This form is the only validated version, and includes the instructions for each item printed directly on the scoring sheet. The form has been modified slightly (header items specific to the original trial deleted) to make it more generic and usable.
3. <u>Filming the NIHSS</u>. This document summarizes key facts about the filming of the videos used to train and certify users. These videos are used by all on-line NIHSS certification vendors. During filming, a number of issues were resolved that may be of interest to regulators intending to require NIHSS utilization by non-research health care personnel.

This is the original scoring manual developed for the NINDS Trial of Rt-PA for Acute Stroke. Some changes many have been made since publication (1989).

Chapter 5 - The NIH Stroke Scale

5.1   Overview

        The Stroke Scale is a standardized neurological examination intended to describe the neurological deficits found in large groups of stroke patients participating in treatment trials. The instructions contained in this manual reflect primary concern for reproducibility. The goal is to have multiple examiners at different sites rate patients similarly. It is possible to challenge the scale on sub-items, and competent neurologists will disagree over the "best" method for testing some items in individual patients. Nevertheless, our interest in reproducibility among many observers in a large multi-center study is paramount, and to this end, all examiners at all sites must use the scale uniformly. We recognize that for some examiners, this means that some testing may be done one way for the study, and a different way in usual clinical practice. The consolation for this disparity is the knowledge that the reproducibility among examiners using this scale will (hopefully) be extremely high.

There are four general principals underlying the scale in its present form:

1. The most reproducible response is generally the first response. For example, on LOC questions, the patient is asked to state age and the current month. The patient who initially responds incorrectly, but later corrects himself, is scored as having given an incorrect response. This approach is critical, because we have no way of standardizing the myriad verbal and non-verbal cues that might be given to patients to promote a correction of an initially incorrect response.

2. It is not permissible to coach patients on any item unless specified in the instructions. This contradicts neurological teaching, since we are generally interested in a patient's best possible performance. Again, standardization of coaching is not possible, and coaching must be avoided in the interest of reproducibility.

3. Some items are scored only if definitely present. For example, ataxia is scored as absent in the patient with hemiplegia, because it is not definitely present at the time of examination. Although somewhat counter intuitive to some physicians, the item must be scored this way to avoid ambiguity and ensure reproducibility.

4. Most importantly, record what the patient does, not what you think the patient can do even if the findings appear contradictory. Many times a competent examiner forms an impression of the patient's level of function, but this impression must not influence scoring. Scoring should include prior deficits except for the sensory item (see instructions)

        The patient's scores should be recorded immediately after the examination, and preferably, each item should be coded as you go through the scale. This is especially necessary at baseline. If baseline results are recorded after the patient has received medication, the examiner may be influenced by the patient's response.

5.2 Certification

Any investigator completing the NIH Stroke Scale for the trial must be certified. Any experienced
Clinical Center Personnel (physician or nurse or physician's assistant) may be certified.

5.2.1   Requirements for Certification

Certification requires:

-       Review of the NIH Stroke Scale Training Tape
-Completion of NIH Stroke Scales for the five patients shown on the NIH Stroke Scale
          Certification Tape #1
-Submission of the five completed forms to the Coordinating Center for review
-Approval by the Coordinating Center

5.2.2   Retention of Certification

Retention of certification by a certified investigator requires:

-Completion of the NIH Stroke Scales on the six patients shown on the NIH Stroke Scale
          Certification Tape #2 (approximately six months after the initial certification)
-Submission of the six completed forms to the Coordinating Center for review.
-Approval by the Coordinating Center

5.3 The NIH Stroke Scale (Form 5)

PURPOSE: This form collects data representing the primary endpoints of the trial (difference
          from Baseline at 24 hour and 3 month NIH Stroke Scale data). A more complete
          discussion of purpose is given in the overview (Section 5.1)

WHEN: The Stroke Scale is completed at baseline **PRIOR TO TREATMENT**, at 2 hours $\pm$ 5
          minutes post treatment onset, 24 hours $\pm$ 20 minutes post stroke onset, 7 to 10 days and
          3 months $\pm$ 2 weeks.

PERSON COMPLETING FORM: The Stroke Scale must be completed by a certified trial investigator.

All stroke scales except for the 2-hour scale must be completed by a certified trial investigator. If possible, the 2-hour scale should also be completed by certified trial personnel. If it is impractical to have a certified investigator perform the 2-hour scale, it may be performed by an uncertified person with telephone supervision by a certified investigator.

The 24-hour NIH Stroke Scale **must** be done by a certified investigator who was not present during treatment. ***The 3-month scale must*** be done by a certified investigator who was not present during treatment. (The two time points are the primary endpoints of the trial.)

INSTRUCTIONS: Extensive instructions are included on the NIH Stroke Scale Form. Additional comments follow:

Three items are used to assess the patient's level of consciousness. It is vital that the items be asked in a standardized manner, as illustrated in the Stroke Scale training tape. Responses must be graded based on what the patient does first. Do not give credit if the patient corrects himself/herself and do not give any clues or coaching.

1a. Level of Consciousness

Instructions:
The investigator must choose a response, even if a full evaluation is prevented by such obstacles as an endotracheal tube, language barrier, orotracheal trauma/bandages. A 3 is scored only if the patient makes no movement (other than reflexive posturing) in response to noxious stimulation.

Comments:
Ask the patient two or three general questions about the circumstances of the admission. Also, prior to beginning the scale, it is assumed that the examiner will have queried the patient informally about the medical history. Based on the answers, score the patient using the 4 point scale on the Stroke Scale form. Remember not to coach. A score of 3 is reserved for the severely impaired patient who makes, at best, reflex posturing movements in response to repeated painful stimuli. If it is difficult to choose between a score of 1 or 2, continue to question the patient about historical items until you feel comfortable in assessing level of consciousness.

1b. LOC Questions

Instructions:
The patient is asked the month and his/her age. The answer must be correct - there is no partial credit for being close. Aphasic and stuporous patients who do not comprehend the questions will score 2. Patients unable to speak because of endotracheal intubation, orotracheal trauma, severe dysarthria from any cause, language barrier or any other problem not secondary to aphasia are given a 1. It is important that only the initial answer be graded and that the examiner not "help" the patient with verbal or non-verbal cues.
Comments:
Ask the patient "how old are you now" and wait for a response. Then ask "what month is it now" or "what month are we in now". Count the number of incorrect answers and do not give credit for being "close". Patients who cannot speak are allowed to write. **Do not give a list of**

possible responses from which to choose the correct answer. This may coach the patient. Only the initial answer is graded. This item is never marked "untestable". (Note: On Certification Tape #1 an intubated patient was given a series of responses from which to choose, but the score for this patient would still be 1.) Deeply comatose (1a=3) patients are given a 2.

1c. LOC Commands:

Instructions:
The patient is asked to open and close the eyes and then to grip and release the non-paretic hand. Substitute another one step command if the hands cannot be used. Credit is given if an unequivocal attempt is made but not completed due to weakness. If the patient does not respond to command, the task should be demonstrated to them (pantomime) and score the result (i.e., follows none, one or two commands). Patients with trauma, amputation, or other physical impediments should be given suitable one-step commands. Only the first attempt is scored.

Comments:
Say to the patient "open your eyes...now close your eyes" and then "Make a fist...now open your hand". Use the non-paretic limb. If amputation or other physical impediment prevents the response, use another suitable one step command. The priming phrase is not scored, and these are used only to set the eyes or hand in a testable position. That is, the patient may be asked first to open the eyes if they are closed when you begin the test. Scoring is done on the second phrase "close your eyes". Count the number of incorrect responses and give credit if an unequivocal attempt is made to perform the operative task, but is not completed due to weakness, pain or other obstruction. Only the first attempt is scored and the questions should be asked only once.

Item 2:Best Gaze

Instructions:

Only horizontal eye movements will be tested. Voluntary or reflexive (oculocephalic) eye movements will be scored but caloric testing is not done. If the patient has a conjugate deviation of the eyes that can be overcome by voluntary or reflexive activity, the score will be 1. If a patient has an isolated peripheral nerve paresis (CN III, IV or VI) score a 1. Gaze is testable in all aphasic patients. Patients with ocular trauma, bandages, pre-existing blindness or other disorder of visual acuity or fields should be tested with reflexive movements and a choice made by the investigator. Establishing eye contact and then moving about the patient from side to side will occasionally clarify the presence of a partial gaze palsy.

Comments:
The purpose of this item is to observe and score horizontal eye movements. To this end, use voluntary or reflexive stimuli and record a score of 1 if there is an abnormal finding in one or both eyes. A score of two is reserved for forced eye deviation that cannot be overcome by the oculocephalic maneuver. Do not do caloric testing. In aphasic or confused patients it is helpful to establish eye contact and move about the bed.

This item is an exception to the rules of using the first observable response and not coaching. In the patient who fails voluntary gaze, the oculocephalic maneuver, eye fixation, and tracking with the examiner's face, are used to provide stronger testing stimuli.

Item 3:Visual

Instructions:
Visual fields (**upper and lower quadrants**) are tested by confrontation, using finger counting or visual threat as appropriate. Patient must be encouraged, but if they look at the side of the moving fingers appropriately, this can be scored as normal. If there is unilateral blindness or enucleation, visual fields in the remaining eye are scored. Score 1 only if a clear-cut asymmetry, including quadrantanopia is found. If patient is blind from any cause score 3. Double simultaneous stimulation is performed at this point. If there is extinction patient receives a 1 and the results are also used to answer question 11.

Comments:
Visual fields are tested exactly as demonstrated in the training video. Use finger counting or movement to confrontation and evaluate upper and lower quadrants separately. A score of 3 is reserved for blindness from any cause, including cortical blindness. A score of 2 is reserved for a complete hemianopia, and any partial visual field defect, including quadrantanopia, scores a 1.

Item 4:Facial Palsy

Instructions:
Ask, or use pantomime to encourage the patient to show teeth or raise eyebrows and close eyes. Score symmetry of grimace in response to noxious stimuli in the poorly responsive or non-comprehending patient. If facial trauma/bandages, orotracheal tube, tape or other physical barrier obscures the face, these should be removed to the extent possible.

Comments:
Ask the patient "Show me your teeth...now raise your eyebrows...now close your eyes tightly". Assess the response to noxious stimulation in the aphasic or confused patient. A useful approach to scoring may be as follows: score a 2 for any clear cut upper motor neuron facial palsy. Normal function must be clearly demonstrated to obtain the score of 0. Anything in between, including flattened nasolabial fold, is scored a 1. The severely obtunded or comatose patient; patients with bilateral paresis, patients with unilateral lower motor neuron facial weakness would receive a score of 3.

Items 5 & 6: Motor Arm and Leg

Instructions:
Each limb is tested in turn, beginning with the non-paretic arm, if known. The limb is placed in the appropriate position: extend the arm (palm down) 90 degrees (if sitting) or 45 degrees (if supine) and the leg 30 degrees (always tested supine). Drift is scored if the arm falls before 10 seconds or the leg before 5 seconds. The aphasic patient is encouraged using urgency in the voice and pantomime but not noxious stimulation. Only in the case of amputation or joint fusion at the shoulder or hip may the score be "untestable" and the examiner must clearly write the explanation for scoring as a "untestable".

Comments:
Perform the test for weakness as illustrated in the video. When testing arms, palms must be down. Count out loud to the patient, until the limb actually hits the bed or other support. The score of 3 is reserved for the patient who exhibits no strength whatsoever, but does minimally move the limb on command when it is resting on the bed. The aphasic patient may understand

what you are testing if you use the non-paretic limb first. Do not test both limbs simultaneously. Be watchful for an initial dip of the limb when released. Only score abnormal if there is a drift after the dip.

Do not coach the patient verbally. Count out loud in strong voice and indicate count using your fingers in full view of the patient. Begin counting the instant you release the limb. (Note that on some of the video illustrated patients, the examiners erroneously delay seconds before beginning to count).

When testing motor leg the patient must be in the supine position to fully standardize the effect of gravity. Note that the examiner is no longer asked to identify the paretic arm or leg. The examiner's assessment of the side of the stroke is given on the Treatment Form (Form 7).

Item 7:Limb Ataxia

Instructions:
This item is aimed at finding evidence of a unilateral cerebellar lesion. Test with eyes open. In case of visual defect, insure testing is done in the intact visual field. The finger-nose-finger and heel-shin tests are performed on both sides, and ataxia is scored only if present out of proportion to weakness. Ataxia is absent in the patient who cannot understand or is paralyzed. Although the use of untestable is discouraged, in the case of amputation, joint fusion or some fractures, the item may be scored "untestable", and the examiner must clearly write the explanation for not scoring. In case of blindness test by touching nose from extended arm position.

Comments:
Ataxia must be clearly present out of proportion to any weakness. Using the finger-nose-finger and the heel-test, count the number of ataxic limbs, up to a maximum of two. The aphasic patient will often perform the test normally if first the limb is passively moved by the examiner. Otherwise, the item is scored 0 for absent ataxia. If the weak patient suffers mild ataxia, and you cannot be certain that it is out of proportion to the weakness, give a score of 0. Remember this is scored positive only when ataxia is present.

Item 8:Sensory

Instructions:
Sensation or grimace to pin prick when tested, or withdrawal from noxious stimulus in the obtunded or aphasic patient. Only sensory loss attributed to stroke is scored as abnormal and the examiner should test as many body areas [arms (not hands), legs, trunk, face] as needed to accurately check for hemisensory loss. A score of 2, "severe or total," should only be given when a severe or total loss of sensation can be clearly demonstrated. Stuporous and aphasic patients will therefore probably score 1 or 0. The patient with brain stem stroke who has bilateral loss of sensation is scored 2. If the patient does not respond and is quadriplegic score 2. Patients in coma (item 1a=3) are arbitrarily given a 2 on this item.

Comments:
Do not test limb extremities, i.e., hands and feet when testing sensation because an unrelated neuropathy may be present. Do not test through clothing.

Item 9: Best Language

Instructions:
A great deal of information about comprehension will be obtained during the preceding sections of the examination. The patient is asked to describe what is happening in the attached picture, to name the items on the attached naming sheet, and to read from the attached list of sentences. Be complete. Have the patient name <u>all</u> items on the naming sheet and read <u>all</u> phrases on the two reading sheets. Comprehension is judged from responses here as well as to all of the commands in the preceding general neurological exam. If visual loss interferes with the tests, ask the patient to identify objects placed in the hand, repeat, and produce speech. The intubated patient should be asked to write. The patient in <u>coma</u> (question 1a=3) will arbitrarily score 3 on this item. The examiner must choose a score in the patient with stupor or limited cooperation but a score of 3 should be used only if the patient is mute and follows no one step commands.

Comments:
It is anticipated that most examiners will be ready to score this item based on information obtained during the history taking and the 8 prior items. The attached picture and naming sheet therefore should be used to confirm your impression. It is common to find unexpected difficulties when the formal testing is done, and therefore every patient must be tested with the picture, naming sheet, and sentences. The score of 3 is reserved for the globally mute or <u>comatose</u> patient. Mild aphasia would score a 1. To choose between a score of 1 or 2 use all the provided materials; it is anticipated that a patient who missed more than two thirds of the naming objects and sentences or who followed only very few and simple one step commands would score a 2.

    This item is an exception to the rule that the first response is used, since several different tools are used to assess language. The stroke scale form contains lengthy examples of the defects associated with each score because of the great potential for variability in answering this question.**Item 10:Dysarthria**

Instructions:

If the patient is thought to be normal, an adequate sample of speech must be obtained by asking patient to read or repeat words from the attached list. If the patient has severe aphasia, the clarity of articulation of spontaneous speech can be rated. Only if the patient is intubated or has other physical barrier to producing speech, may the item be scored "untestable", and the examiner must clearly write an explanation for not scoring. Do not tell the patient why he/she is being tested.

Comments:

Use the attached word list in all patients and do not tell the patient that you are testing clarity of speech. It is **common** to find slurring of one or more words in patients one might otherwise score as normal. The score of 0 is reserved for patients who read all words without any slurring. <u>Aphasic</u> patients and patients who do not read may be scored based on listening to the speech that they do produce or by asking them to repeat the words after you read them out loud. The score of 2 is reserved for the patient who cannot be understood in any meaningful way, or who is mute.

On this question, normal speech must be identified to score a 0, so the unresponsive patient receives the score of 2.

Item 11:Extinction and inattention

Instructions:

Sufficient information to identify neglect may be obtained during the prior testing. If the patient has a severe visual loss preventing visual double simultaneous stimulation, and the cutaneous stimuli are normal, the score is normal. If the patient has aphasia but does appear to attend to both sides, the score is normal. The presence of visual spatial neglect or anosagnosia may also be taken as evidence of abnormality. Since the abnormality is scored only if present, the item is never untestable.

Comments:

This item is open to significant variation among examiners, and all neurologists have slightly different methods of assessing neglect. Therefore, to the extent possible, test only double simultaneous stimulation to visual and tactile stimuli and score 2 if one side extinguishes to both modalities, a 1 if only to one modality. If the patient does not extinguish, but does show other well developed evidence of neglect, score a 1.

5.4 Coma

   A patient with a 3 on Item 1a (Level of Consciousness) is considered to be in a coma. A patient suspected to be in coma should be stimulated by rubbing on the chest or by using a painful stimulus. A 3 is scored only if the patient makes no movement (other than reflexive posturing) in response to the noxious stimulation. **Patients who appear to be in coma and who score less than 3 must be tested on all items.**

   For patients scoring a 3 on Item 1a, the remaining items should be scored as follows:

Item 1b (LOC Questions) - Score 2.

Item 1c (LOC Commands) - Score 2.

Item 2 (Best Gaze) - Patient can be in coma and have gaze palsy that can be overcome by moving the head. Thus the occulocephalic maneuver must be done and the patient scored.

Item 3 (Visual) - Test using bilateral threat.

Item 4 (Facial Palsy) - Score 3.

Items 5 and 6 (Motor Arm and Leg) - This item is interpreted as the voluntary ability to attain a posture. Score 4 for both arm and leg.

Item 7 (Limb Ataxia) - Scored only if present, out of proportion to weakness. Score 0.

Item 8 (Sensory) - Score 2 (arbitrary).

Item 9 (Best Language) - Score 3.

Item 10 (Dysarthria) - Score 2.

Item 11 (Extinction and inattention) - Coma implies loss of all cognitive abilities. Score 2.

5.5 Persons Who Refuse to Cooperate

In the event that a patient refuses to perform the tasks in the course of the examination resulting in an item untested, a detailed explanation must be clearly written on the form. All untested items will be reviewed by the medical monitor and discussed with the examiner if necessary.

5.6 Calculating a Score

In computing a score, the following items should **not** be added to the total:

■ For Item 7 (Limb Ataxia) codes for affected sides (right and/or left arm and leg;
     1 = yes, 2 = no, 9 = untestable).

■ Distal Motor Function.

■ Any 9's.

5.7 Outliers

    There are questions in the certification Tapes 1 and 2 that do not have a single answer. Thus the distribution of responses from those who have completed the certification is used. A response given by 12% or fewer examiners is considered an outlier. Any examiner having 10 or more outliers for Tape 1 or 12 or more outliers for Tape 2 is not certified to do stroke scales for the trial. An examiner who is not certified must redo the certification before they can perform stroke scale evaluations for the trial. They should carefully review the training tape before repeating the certification. Examiners having 6 to 9 outliers for Tape 1 or 7 - 11 outliers for Tape 2 are required to repeat the certification but can continue to do stroke scales for the trial in the interim.

N I H STROKE SCALE

Time: ___ ___:___ ___   [ ]am   [ ]pm

Person Administering Scale _____

Administer stroke scale items in the order listed. Record performance in each category after each subscale exam. Do not go back and change scores. Follow directions provided for each exam technique. Scores should reflect what the patient does, not what the clinician thinks the patient can do. The clinician should record answers while administering the exam and work quickly. Except where indicated, the patient should not be coached (i.e., repeated requests to patient to make a special effort).

Instructions	Scale Definition	Score
1a. Level of Consciousness: The investigator must choose a response if a full evaluation is prevented by such obstacles as an endotracheal tube, language barrier, orotracheal trauma/bandages. A 3 is scored only if the patient makes no movement (other than reflexive posturing) in response to noxious stimulation.	0 = **Alert;** keenly responsive. 1 = **Not alert**; but arousable by minor stimulation to obey, answer, or respond. 2 = **Not alert**; requires repeated stimulation to attend, or is obtunded and requires strong or painful stimulation to make movements (not stereotyped). 3 = Responds only with reflex motor or autonomic effects or totally unresponsive, flaccid, and areflexic.	_____
1b. LOC Questions: The patient is asked the month and his/her age. The answer must be correct - there is no partial credit for being close. Aphasic and stuporous patients who do not comprehend the questions will score 2. Patients unable to speak because of endotracheal intubation, orotracheal trauma, severe dysarthria from any cause, language barrier, or any other problem not secondary to aphasia are given a 1. It is important that only the initial answer be graded and that the examiner not "help" the patient with verbal or non-verbal cues.	0 = **Answers** both questions correctly. 1 = **Answers** one question correctly. 2 = **Answers** neither question correctly.	_____
1c. LOC Commands: The patient is asked to open and close the eyes and then to grip and release the non-paretic hand. Substitute another one step command if the hands cannot be used. Credit is given if an unequivocal attempt is made but not completed due to weakness. If the patient does not respond to command, the task should be demonstrated to him or her (pantomime), and the result scored (i.e., follows none, one or two commands). Patients with trauma, amputation, or other physical impediments should be given suitable one-step commands. Only the first attempt is scored.	0 = **Performs** both tasks correctly. 1 = **Performs** one task correctly. 2 = **Performs** neither task correctly.	_____

Patient Identification. ___ _____ ___ _____ ___ ___

Date of Exam ___ ___/___ ___/___ ___

2. Best Gaze: Only horizontal eye movements will be tested. Voluntary or reflexive (oculocephalic) eye movements will be scored, but caloric testing is not done. If the patient has a conjugate deviation of the eyes that can be overcome by voluntary or reflexive activity, the score will be 1. If a patient has an isolated peripheral nerve paresis (CN III, IV or VI), score a 1. Gaze is testable in all aphasic patients. Patients with ocular trauma, bandages, pre-existing blindness, or other disorder of visual acuity or fields should be tested with reflexive movements, and a choice made by the investigator. Establishing eye contact and then moving about the patient from side to side will occasionally clarify the presence of a partial gaze palsy.	0 = **Normal.** 1 = **Partial gaze palsy;** gaze is abnormal in one or both eyes, but forced deviation or total gaze paresis is not present. 2 = **Forced deviation,** or total gaze paresis not overcome by the oculocephalic maneuver.	_____
3. Visual: Visual fields (upper and lower quadrants) are tested by confrontation, using finger counting or visual threat, as appropriate. Patients may be encouraged, but if they look at the side of the moving fingers appropriately, this can be scored as normal. If there is unilateral blindness or enucleation, visual fields in the remaining eye are scored. Score 1 only if a clear-cut asymmetry, including quadrantanopia, is found. If patient is blind from any cause, score 3. Double simultaneous stimulation is performed at this point. If there is extinction, patient receives a 1, and the results are used to respond to item 11.	0 = **No visual loss.** 1 = **Partial hemianopia.** 2 = **Complete hemianopia.** 3 = **Bilateral hemianopia** (blind including cortical blindness).	_____
4. Facial Palsy: Ask – or use pantomime to encourage – the patient to show teeth or raise eyebrows and close eyes. Score symmetry of grimace in response to noxious stimuli in the poorly responsive or non-comprehending patient. If facial trauma/bandages, orotracheal tube, tape or other physical barriers obscure the face, these should be removed to the extent possible.	0 = **Normal** symmetrical movements. 1 = **Minor paralysis** (flattened nasolabial fold, asymmetry on smiling). 2 = **Partial paralysis** (total or near-total paralysis of lower face). 3 = **Complete paralysis** of one or both sides (absence of facial movement in the upper and lower face).	_____

5. Motor Arm: The limb is placed in the appropriate position: extend the arms (palms down) 90 degrees (if sitting) or 45 degrees (if supine). Drift is scored if the arm falls before 10 seconds. The aphasic patient is encouraged using urgency in the voice and pantomime, but not noxious stimulation. Each limb is tested in turn, beginning with the non-paretic arm. Only in the case of amputation or joint fusion at the shoulder, the examiner should record the score as untestable (UN), and clearly write the explanation for this choice.	0 = **No drift;** limb holds 90 (or 45) degrees for full 10 seconds. 1 = **Drift;** limb holds 90 (or 45) degrees, but drifts down before full 10 seconds; does not hit bed or other support. 2 = **Some effort against gravity;** limb cannot get to or maintain (if cued) 90 (or 45) degrees, drifts down to bed, but has some effort against gravity. 3 = **No effort against gravity;** limb falls. 4 = **No movement.** UN = **Amputation** or joint fusion, explain: _____ **5a. Left Arm** **5b. Right Arm**	 _____ _____
6. Motor Leg: The limb is placed in the appropriate position: hold the leg at 30 degrees (always tested supine). Drift is scored if the leg falls before 5 seconds. The aphasic patient is encouraged using urgency in the voice and pantomime, but not noxious stimulation. Each limb is tested in turn, beginning with the non-paretic leg. Only in the case of amputation or joint fusion at the hip, the examiner should record the score as untestable (UN), and clearly write the explanation for this choice.	0 = **No drift;** leg holds 30-degree position for full 5 seconds. 1 = **Drift;** leg falls by the end of the 5-second period but does not hit bed. 2 = **Some effort against gravity;** leg falls to bed by 5 seconds, but has some effort against gravity. 3 = **No effort against gravity;** leg falls to bed immediately. 4 = **No movement.** UN = **Amputation** or joint fusion, explain: _____ **6a. Left Leg** **6b. Right Leg**	 _____ _____

|		
7. Limb Ataxia: This item is aimed at finding evidence of a unilateral cerebellar lesion. Test with eyes open. In case of visual defect, ensure testing is done in intact visual field. The finger-nose-finger and heel-shin tests are performed on both sides, and ataxia is scored only if present out of proportion to weakness. Ataxia is absent in the patient who cannot understand or is paralyzed. Only in the case of amputation or joint fusion, the examiner should record the score as untestable (UN), and clearly write the explanation for this choice. In case of blindness, test by having the patient touch nose from extended arm position.	0 = **Absent.** 1 = **Present in one limb.** 2 = **Present in two limbs.** UN = **Amputation** or joint fusion, explain: _____	_____
8. Sensory: Sensation or grimace to pinprick when tested, or withdrawal from noxious stimulus in the obtunded or aphasic patient. Only sensory loss attributed to stroke is scored as abnormal and the examiner should test as many body areas (arms [not hands], legs, trunk, face) as needed to accurately check for hemisensory loss. A score of 2, "severe or total sensory loss," should only be given when a severe or total loss of sensation can be clearly demonstrated. Stuporous and aphasic patients will, therefore, probably score 1 or 0. The patient with brainstem stroke who has bilateral loss of sensation is scored 2. If the patient does not respond and is quadriplegic, score 2. Patients in a coma (item 1a=3) are automatically given a 2 on this item.	0 = **Normal;** no sensory loss. 1 = **Mild-to-moderate sensory loss;** patient feels pinprick is less sharp or is dull on the affected side; or there is a loss of superficial pain with pinprick, but patient is aware of being touched. 2 = **Severe to total sensory loss;** patient is not aware of being touched in the face, arm, and leg.	_____
9. Best Language: A great deal of information about comprehension will be obtained during the preceding sections of the examination. For this scale item, the patient is asked to describe what is happening in the attached picture, to name the items on the attached naming sheet and to read from the attached list of sentences. Comprehension is judged from responses here, as well as to all of the commands in the preceding general neurological exam. If visual loss interferes with the tests, ask the patient to identify objects placed in the hand, repeat, and produce speech. The intubated patient should be asked to write. The patient in a coma (item 1a=3) will automatically score 3 on this item. The examiner must choose a score for the patient with stupor or limited cooperation, but a score of 3 should be used only if the patient is mute and follows no one-step commands.	0 = **No aphasia;** normal. 1 = **Mild-to-moderate aphasia;** some obvious loss of fluency or facility of comprehension, without significant limitation on ideas expressed or form of expression. Reduction of speech and/or comprehension, however, makes conversation about provided materials difficult or impossible. For example, in conversation about provided materials, examiner can identify picture or naming card content from patient's response. 2 = **Severe aphasia;** all communication is through fragmentary expression; great need for inference, questioning, and guessing by the listener. Range of information that can be exchanged is limited; listener carries burden of communication. Examiner cannot identify materials provided from patient response. 3 = **Mute, global aphasia;** no usable speech or auditory comprehension.	_____

10. Dysarthria: If patient is thought to be normal, an adequate sample of speech must be obtained by asking patient to read or repeat words from the attached list. If the patient has severe aphasia, the clarity of articulation of spontaneous speech can be rated. Only if the patient is intubated or has other physical barriers to producing speech, the examiner should record the score as untestable (UN), and clearly write an explanation for this choice. Do not tell the patient why he or she is being tested.	0 = **Normal.** 1 = **Mild-to-moderate dysarthria;** patient slurs at least some words and, at worst, can be understood with some difficulty. 2 = **Severe dysarthria;** patient's speech is so slurred as to be unintelligible in the absence of or out of proportion to any dysphasia, or is mute/anarthric. UN = **Intubated** or other physical barrier, explain:_____	_____
11. Extinction and Inattention (formerly Neglect): Sufficient information to identify neglect may be obtained during the prior testing. If the patient has a severe visual loss preventing visual double simultaneous stimulation, and the cutaneous stimuli are normal, the score is normal. If the patient has aphasia but does appear to attend to both sides, the score is normal. The presence of visual spatial neglect or anosagnosia may also be taken as evidence of abnormality. Since the abnormality is scored only if present, the item is never untestable.	0 = **No abnormality.** 1 = **Visual, tactile, auditory, spatial, or personal inattention** or extinction to bilateral simultaneous stimulation in one of the sensory modalities. 2 = **Profound hemi-inattention or extinction to more than one modality;** does not recognize own hand or orients to only one side of space.	_____

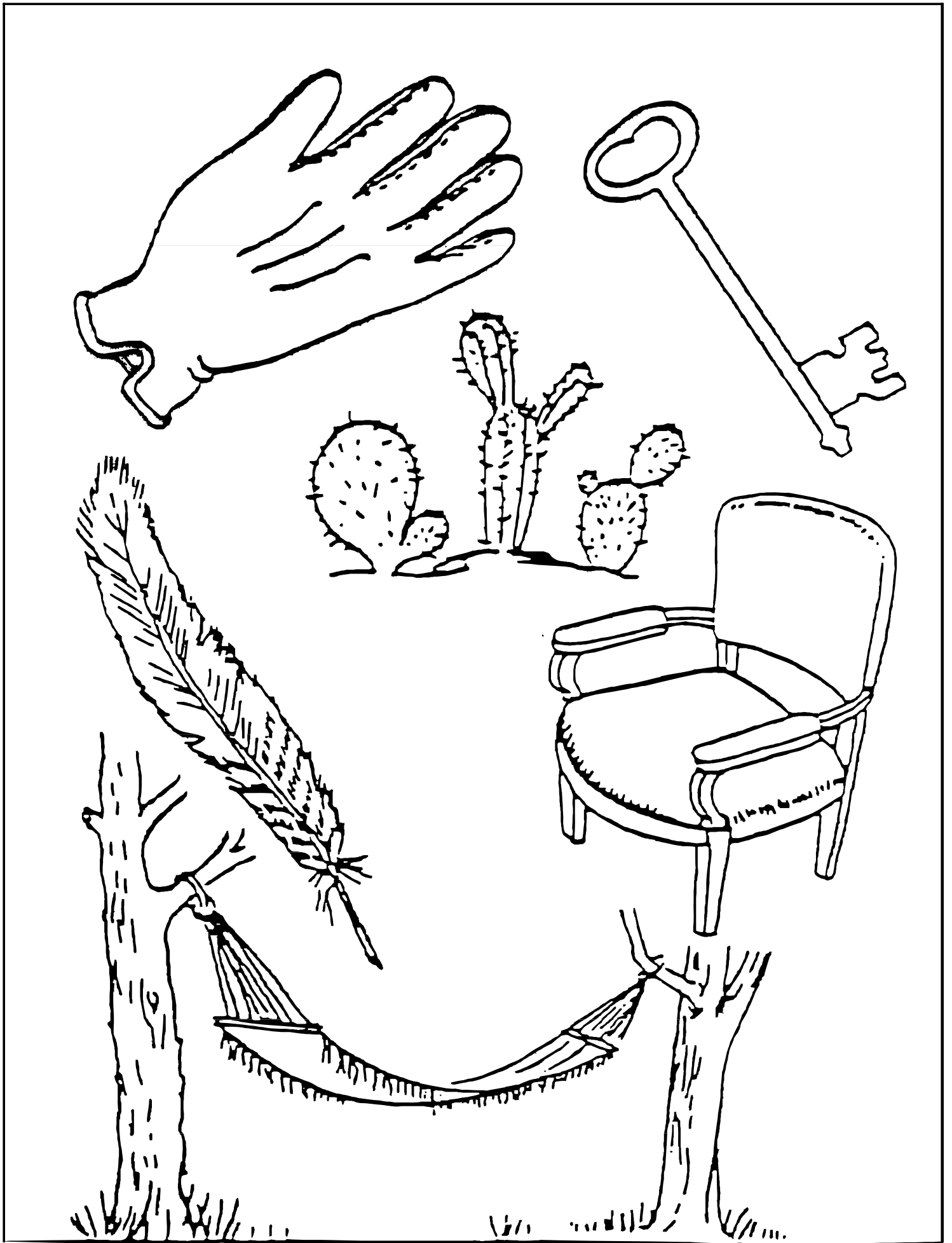## TOTAL   ____
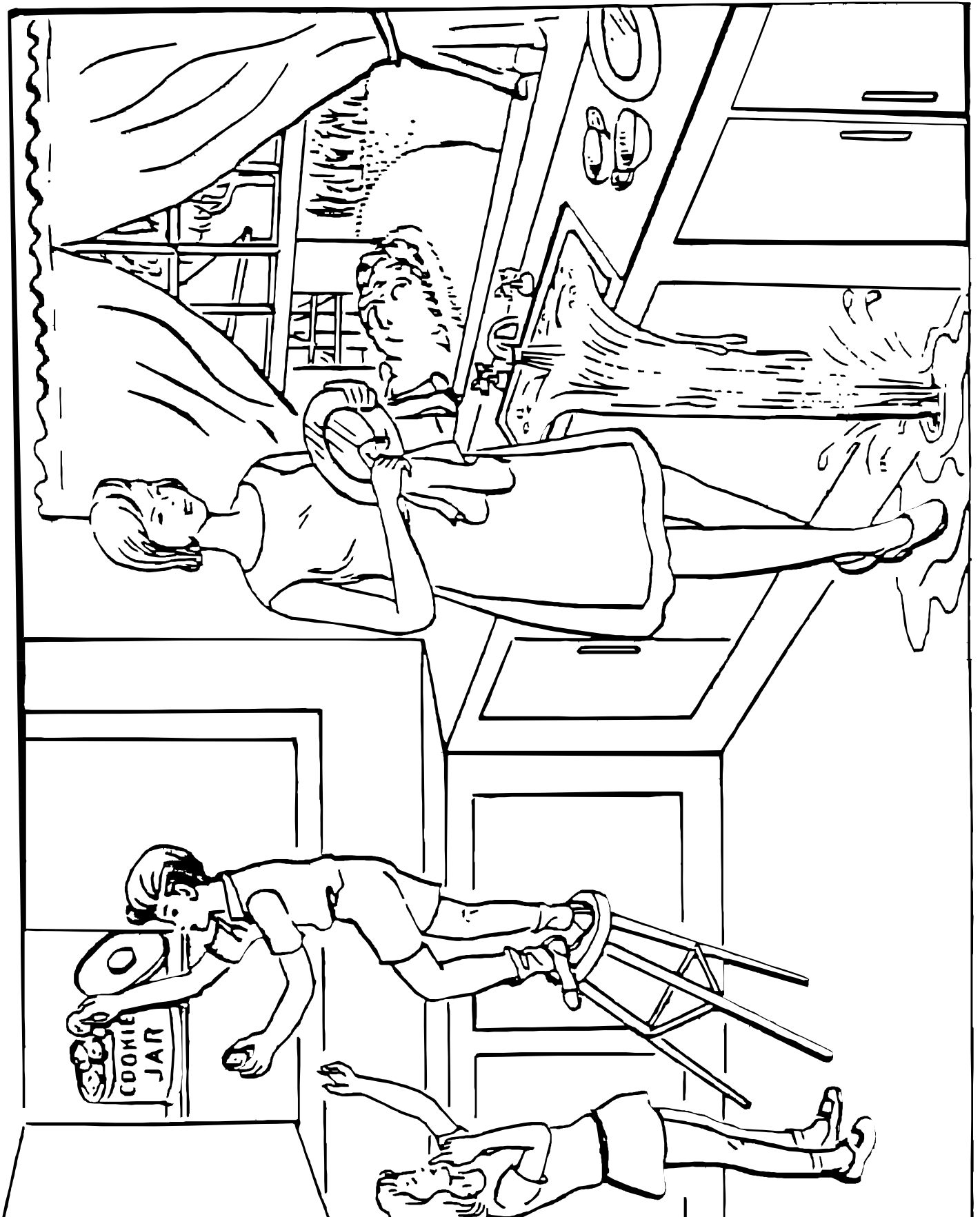
You know how.

Down to earth.

I got home from work.

Near the table in the dining room.

They heard him speak on the
    radio last night.

MAMA

TIP – TOP

FIFTY – FIFTY

THANKS

HUCKLEBERRY

BASEBALL PLAYER

CATERPILLAR

Filming the NIHSS

This document summarizes the process used in filming the NIHSS training and certification videos for use in the NINDS rt-PA for Acute Stroke Trial. These videos are now used widely for training and certification, and thus are of interest to those seeking to understand the strengths and limitations of NIHSS video training and certification.

The first step in validating a stroke deficit rating scale is to train users. Prior to the video era, medicine was taught at the bedside. Clinical trial investigator meetings included break-out sessions to train investigators to properly use the outcome assessment tools. Although unimaginable to today's students, creating a training video involved specialized, expensive cameras and editing equipment; engineering a "master" videotape that was used to then produce videotape reproductions; and physical-mail distribution of VHS videotapes[1]. This creative process involved collaborations with video production teams who brought with them extensive experience, including a sensibility and training that was counterproductive to our purpose. For example, the camera operator with no knowledge of "facial paresis" would light and frame the video shot in a pleasing manner that tended to obscure the subtle flattening of the nasolabial fold that was of greater interest to a neurologist. Thus, in our first attempt to produce an accurate training video, we found it essential to have one of the neurologists behind the camera, directing the videographer to tighten the frame, alter the lighting, or change focus to highlight each clinical deficit. For eye movements and facial weakness, the "close-up" required an ultra-tight shot that violated the film-school rules these professionals knew.

In a live demonstration, the student can observe both the examiner and the patient simultaneously. In a recorded video, the only way to show both the examiner and the patient is to widen the frame, but then the details of the response cannot be seen well. In consultation with our video production colleagues, we developed a 2-camera approach to the production of the video[1]. The video editor could later insert a close-up of the patient response into the frame, allowing the viewer to appreciate both the examiner's technique, and simultaneously a detailed view of the response (Figure). Prior to beginning the NINDS rt-PA for Acute Stroke Trial (the Trial), we produced a training video and 2 certification videos at Henry Ford Hospital. Actual patient volunteers were recruited from the large stroke ward managed at that time by Dr. K.M.A. Welch. Several of the NINDS investigators volunteered to perform aspects of the NIHSS in front of the camera, while others watched from behind the camera to assure accuracy. The collaborating professionals from the video production industry reluctantly responded to our repeated requests to "forget what you learned in film school". Hundreds of hours of tape—including multiple "takes" of each shot needed to illustrate each examination finding—were then edited over a 2-week period in Detroit to produce the needed master videotape. To assure the greatest neurological accuracy across all patient vignettes, one of us (PL) sat with the editor and selected each shot for each patient demonstration. The plan was to create one "training" tape and two "certification" tapes that were to be

viewed after training. The 2 different certification groups were intended to comprehensively test all possible responses to all items on the NIHSS; however, since the videotaping was completed in 2 separate 1-week sessions, and since the only available subjects were whatever patients happened to be admitted during those two weeks, it was impossible to cover all the possible responses, a significant shortcoming of the videotapes[1]. To allow the user time to write a score on the paper answer sheet, pauses were inserted by the video editor during which the screen displayed the graphic message "Record Your Score". After mastering and reproduction, the tapes were distributed to the study sites; all participants in the Trial were required to view the training tape and then one certification tape and score each certification patient. On Certification Tape 1 there were 5 patients and on Tape 2 there were 6 patients, and each patient was shown doing 15 tasks, for a total of 165 tasks followed by "Record Your Score" moments. Once the test was completed, scores were reviewed centrally for "grading", and the successful student-viewer was sent a coffee mug imprinted with the phrase "Record Your Score". Only after passing central review and approval was any investigator allowed to begin enrolling patients into the Trial.

Certification and training

Video training and certification introduces bias into the learning and testing process. The student-viewer needs to be able to see the finding accurately, apply the scoring rule correctly, and then derive the correct score. Unfortunately, even our innovative 2-camera shooting paradigm could not accurately depict every finding on every patient. Subtle findings on sensory and ataxia items were especially problematic[2, 3]. We thus could not create a certification scoring system in which we compared the student-viewer scores to the actual, known deficits in that patient. Therefore, we designed a scoring system that accounted for the artificial limitations of the video viewing process[1].

First, we asked 3 highly experienced and interested investigators from the Trial Steering committee to view and score all 11 certification patients; reproducibility was excellent. As expected, however, even expert users did not see every finding accurately, even though they agreed with each other. Therefore, we created a scoring algorithm: after all (n=162) investigators at all study sites returned their scoring sheets, the mode response was determined for each test item for each certification patient. We required that the mode be clearly identifiable and where more than one response qualified as a "mode" then both, or in some cases 3, responses were allowed as "correct". Once the mode responses were known, each response sheet was scored: any responses more than 1 response from the mode response was scored "outlier". We used the rule that any student-viewer could score one outlier per patient. Thus, no more than 5 outliers were allowed for Certification Tape 1, and no more than 6 outliers were allowed for Certification Tape 2. Users who scored less than the allowed number of outliers on Tape 1 were "certified" and allowed to begin entering patients into the trial. Users who scored more than the allowed number of outliers were asked to re-watch the training video and try again. After 6 months all users were asked to view and score Certification Tape 2 to assure continued proficiency with the scale. Exactly as for Tape 1, mode responses were tabulated and outliers identified.

One intended consequence of the Tapes was that new investigators could be added easily to the trial at any time; this design solved a chronic problem for trials run in the 1980s and 1990s: rather than delay certification of new investigators to the next investigator's training workshop, new team members at the trial study sites were asked to view the Training Tape and then the Certification Tape 1. New users were scored for outlier responses and certified as was done with the original investigators. During the course of the Trial, dozens of new investigators, including MDs and RNs, were added to the trial using these tapes, an innovation for clinical trials at the time. However, an unintended consequence of this scoring system was that there were a number of "correct" responses to many of the case scenarios, creating an impression of leniency in the scoring[4].

Use in trials

Following the publication of the NINDS rt-PA for Acute Stroke Trial in 1995, and regulatory approval of the drug in 1996, clinical trialists expressed interest in using the NIHSS for their clinical trials. Soon there were requests for the Training and Certification Videotapes. Given the culture of the 1990s, and the fact that the tapes had been produced under NIH auspices, all of these requests were granted. Both NIH and industry-sponsored studies were allowed to purchase copies of the Tapes at cost from Henry Ford Hospital—hundreds of videotapes were produced and shipped. Scoring of the user test results were all done by the statisticians at Henry Ford Hospital; the "answer sheet" has never been released or published. Published trials at the time stated in their methods sections that their investigators were trained and certified by NINDS using the Tapes.

After a few years, it became obvious that limitations of the original videotapes were unacceptable. The public communications team at NINDS (Marion Emr and Margo Warren) commissioned a re-shoot of the videos using professionals from the video production industry. The re-shoot was planned to correct many of the deficiencies in the first round of Tapes. Rather than attempt shooting in a clinic or hospital room as was done for the first set of Tapes, a professional set was built in the television studio at the University of California, San Diego, to allow more accurate lighting and careful camera positioning. Again the 2-camera arrangement was used to allow simultaneous recording of the examiner and the patient-response (Figure). The set included several features to enhance demonstration of the patient findings and overcome limitations discovered in the first production. For example, to optimize the visualization of limb drift (Items 5 and 6, Table One of the main manuscript) we placed horizontal blinds over the faux-window on the set: the drifting limb could easily be observed against the horizontal blinds in the background. Special 'fill-in' lighting was designed to optimize visualization of the eye movements and the facial asymmetry. Again, video production professionals were asked to ignore their rules of good photography craft, and instead frame the patient overly tightly or focus too close or too far away. Outpatients from the UCSD Stroke Center and inpatients at several area hospitals were selected so as to assure that every single response on every NIHSS Item was illustrated in both the Training and the Certification videos. By the time of the re-shoot, videotape technology had been replaced by digital video disks (DVD), which facilitated the certification process: since the student-viewer could select each patient item in turn, there was no need for the repetitive pause screen "Record Your Score".

The re-shoot again included many investigators from NINDS-sponsored trials who appeared on camera to illustrate the correct performance of each NIHSS Item. Video recording was completed over 2 weeks in February 2003. Again, one of us (PL) sat with the editing team to select the shots that best illustrated each finding for each item on all patients; editing was completed over several weeks in Washington D.C. A Training videodisk was created that presented each NIHSS scale Item and its scoring rules in detail using the video demonstrations. Then, the recordings of Certification patients were divided into 3 groups carefully created such that each group contained a balance of mild/severe and right/left hemisphere strokes. Over all 3 groups, each response on each Item was shown at least once. Today, Groups A, B, and C are presented sequentially, although in some implementations a group is picked at random for certification. After digital mastering, the Training Program and the Certification Patients were copied onto DVDs. The NINDS took over responsibility for distributing the DVDs to interested groups who were organizing a number of large clinical trials and UCSD provided review and grading services, either directly or via a website[5-8].

Once again, the scoring algorithm developed during the rt-PA for Acute Stroke Trial was used to certify users. Since there was no pool of trial investigators, however, it was necessary to validate the scoring rules in stages[9]. First, the scoring system was "seeded" with correct responses based on 51 responses from expert users at 3 leading stroke centers (UCSD, UT Houston, and University of Cincinnati). The first 50 student-viewers were scored using the 'outlier' method (no more than one outlier per patient) but in place of modal responses the seeded responses were used. After 50 users were certified, we tabulated the modal responses and created a new set of accepted scores by finding the mode response; as before, on some questions there were 2 or even 3 modal responses. Certifications then proceeded as before by scoring outlier responses and allowing no more than one outlier per patient in the certification group. Occasionally over the first 2 years of certification the modal responses were re-summarized, and occasional adjustments were made to the scoring algorithm. As with the Tapes, no "answer sheet" has ever been released or published: the online scoring vendors use the modal responses generated and reviewed by one investigator (PL) and further adjustments may be made from time to time to assure consistency over time.

After a year or 2, depending on regulatory requirements, users return to the DVD or a website, and re-certify by watching the next Certification Group in sequence. Once users have re-certified on all 3 groups, additional re-certifications pick one of the 3 groups to use over again. Some vendors have implemented a 'random' selection process for the 4th and subsequent certification group; others simply start over at Group A. Thus, Group "D" will be one of the original 3 certification groups. It is not possible to create new groupings of Certification patients without disrupting the careful balance of mild/severe and left/right hemisphere strokes among all 3 groups. In the future, a better implementation plan would be that each time a user certifies, a unique constellation of cases would be selected, based on that user's prior test-taking history and seeking to balance mild/moderate patients during each certification experience.

Citations

1.   Lyden P, Brott T, Tilley B, Welch KM, Mascha EJ, Levine S, et al. Improved reliability of the NIH Stroke Scale using video training. NINDS TPA Stroke Study Group. *Stroke*. 1994;25:2220-2226
2.   Lyden P, Claesson L, Havstad S, Ashwood T, Lu M. Factor analysis of the National Institutes of Health Stroke Scale in patients with large strokes. *Arch Neurol*. 2004;61:1677-1680
3.   Lyden PD, Lu M, Levine SR, Brott TG, Broderick J. A modified National Institutes of Health Stroke Scale for use in stroke clinical trials: preliminary reliability and validity. *Stroke*. 2001;32:1310-1317
4.   Hills NK, Josephson SA, Lyden PD, Johnston SC. Is the NIHSS certification process too lenient? *Cerebrovasc Dis*. 2009;27:426-432
5.   Lees KR, Zivin JA, Ashwood T, Davalos A, Davis SM, Diener HC, et al. NXY-059 for acute ischemic stroke. *N Engl J Med*. 2006;354:588-600
6.   Lyden P, Shuaib A, Ng K, Levin K, Atkinson RP, Rajput A, et al. Clomethiazole Acute Stroke Study in ischemic stroke (CLASS-I): final results. *Stroke*. 2002;33:122-128
7.   Johnston KC, Li JY, Lyden PD, Hanson SK, Feasby TE, Adams RJ, et al. Medical and neurological complications of ischemic stroke: experience from the RANTTAS trial. *Stroke*. 1998;29:447-453
8.   Lees KR, Asplund K, Carolei A, Davis SM, Diener HC, Kaste M, et al. Glycine antagonist (gavestinel) in neuroprotection (GAIN International) in patients with acute stroke: a randomised controlled trial. GAIN International Investigators. *Lancet*. 2000;355:1949-1954
9.   Lyden P, Raman R, Liu L, Grotta J, Broderick J, Olson S, et al. NIHSS training and certification using a new digital video disk is reliable. *Stroke*. 2005;36:2446-2449
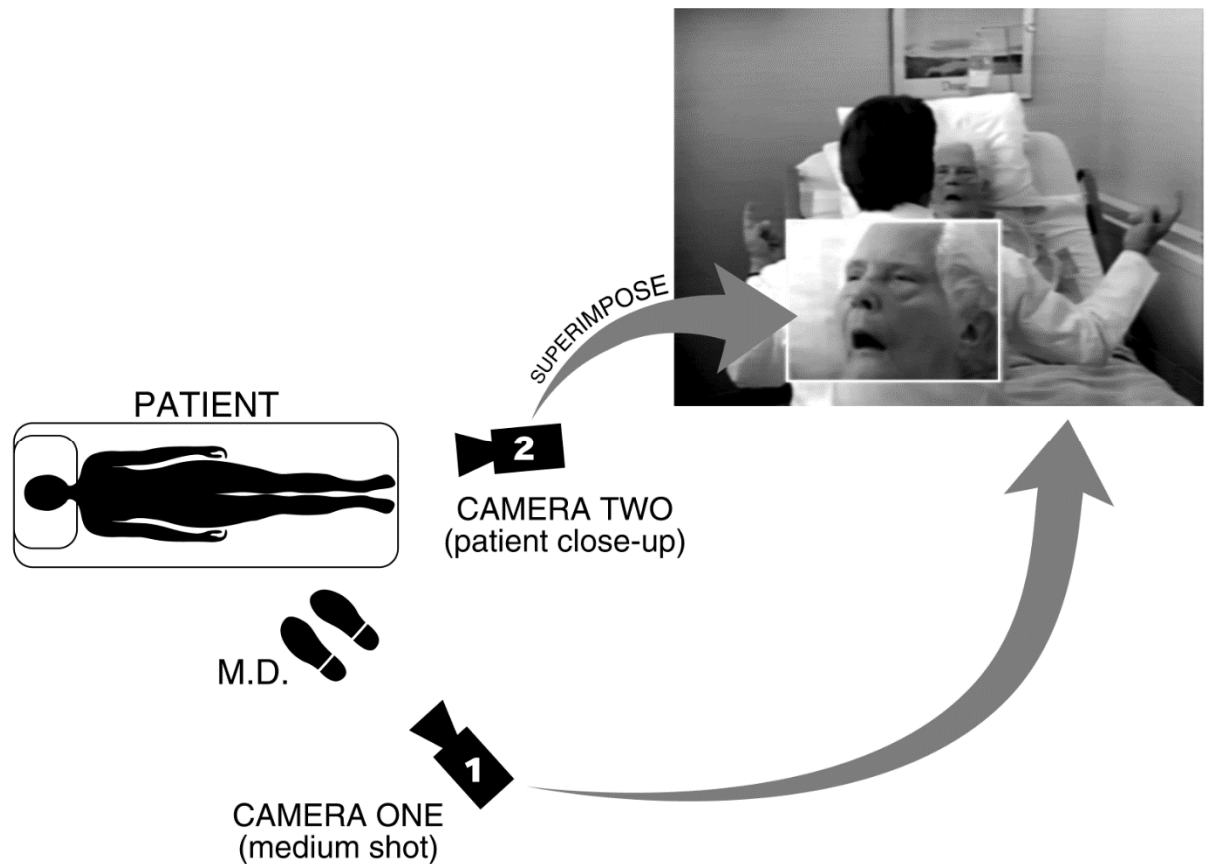
Figure. Framing the test examiner and the subject. Neurological responses can be subtle, but the examiner's stimulus may be large. To show both the larger body movements and commands of the examiner, but also show the possibly subtle responses of the patient, a 2-camera solution was devised. One camera records the examiner showing correct technique, while the second camera records the subtlety of the response, usually in extreme close-up. In post-production, the two images can be merged by the editor to make a single view that demonstrates clearly the method and the response. (Figure reprinted with permission[1]).