

# Effects of Intraoperative Hypothermia on Neuropsychological Outcomes after Intracranial Aneurysm Surgery

Steven W. Anderson, PhD,<sup>1</sup> Michael M. Todd, MD,<sup>2</sup> Bradley J. Hindman, MD,<sup>2</sup> William R. Clarke, PhD,<sup>3</sup> James C. Torner, PhD,<sup>4</sup> Daniel Tranel, PhD,<sup>1</sup> Bongin Yoo, PhD,<sup>3</sup> Julie Weeks, MPT,<sup>2</sup> Kenneth W. Manzel, BA,<sup>1</sup> Satwant Samra, MD,<sup>5</sup> and the IHAST Investigators

**Objective:** Subarachnoid hemorrhage and surgical obliteration of ruptured intracranial aneurysms are frequently associated with neurological and neuropsychological abnormalities. We reported that intraoperative cooling *did not* improve neurological outcome in good-grade surgical subarachnoid hemorrhage patients, as assessed by the Glasgow Outcome Scale score or other neurological and functional measures (National Institutes of Health Stroke Scale, Rankin Disability Scale, Barthel Activities of Daily Living). We now report the results of neuropsychological testing in these patients.

**Methods:** A total of 1,001 patients who bled  $\leq 14$  days before surgery were randomly assigned to intraoperative hypothermia ( $t = 33^\circ\text{C}$ ) or normothermia ( $37^\circ\text{C}$ ). Outcome was assessed approximately 3 months after surgery. Patients underwent the Benton Visual Retention, Controlled Oral Word Association, Rey–Osterrieth Complex Figure, Grooved Pegboard, and the Trail Making tests. *T*-scores for each test were calculated from normative data. *T*-scores were averaged to calculate a Composite Score. A test result (or the Composite Score) was considered “impaired” if the *T*-score was two or more standard deviations below the norm. A Mini-Mental State Examination was also performed.

**Results:** Neurological outcome data were available in 1,000 patients. Sixty-one patients died. Of the 939 survivors, 873 completed 3 or more tests (exclusive of the Mini-Mental State Examination). Patients with poor neurological outcomes were less likely to complete testing; only 3.9% of Good Outcome (Glasgow Outcome Scale score = 1) patients were untested, compared with 38.6% of patients with Glasgow Outcome Scale scores of 3 and 4. There were no prerandomization demographic differences between the two treatment groups. For hypothermic patients, 16.8% were impaired from their Composite Score versus 20.0% of patients in the normothermic group ( $p = 0.317$ ). For patients in the hypothermic group, 54.5% were impaired on at least one test, compared with 55.5% of patients in the normothermic group ( $p = 0.865$ ). Similar results were seen in patients with baseline WFNS scores = I. Mini-Mental State Examination scores in the hypothermic and normothermic groups were  $27.4 \pm 3.8$  and  $26.8 \pm 4.5$ , respectively.

**Interpretation:** This is the largest prospective evaluation of neuropsychological function after subarachnoid hemorrhage to date. Testing was completed in a high fraction of patients, demonstrating the feasibility of such testing in a large trial. However, the frequent inability to complete testing in poor-outcome patients suggests that testing may be best used to refine outcome assessments in good-grade patients. Many patients showed impairment on at least one test, with global impairment present in 17 to 20% of patients (18–21% of survivors). This was true even among the patients with the best preoperative condition (WFNS = 1). There was no difference in the incidence of impairment between hypothermic and normothermic groups.

Ann Neurol 2006;60:518–527

Subarachnoid hemorrhage (SAH) is associated with a high mortality rate and a high incidence of neurological<sup>1–4</sup> and neuropsychological deficits among survivors.<sup>5–7</sup> Many methods have been tried to reduce the frequency of these deficits, particularly in patients requiring surgical clipping of their aneurysms.<sup>8–12</sup>

We recently completed a large clinical trial examining the impact of mild intraoperative hypothermia in

patients undergoing surgical therapy to treat (clip) an acutely ruptured intracranial aneurysm.<sup>13</sup> At 3 months after surgery, we were unable to detect any treatment benefit as assessed by the Glasgow Outcome Scale (GOS) score (our primary outcome variable) or by other neurological measures including the Rankin Disability Scale, the National Institutes of Health Stroke Scale (NIHSS), or the Barthel Activities of Daily Liv-

From the Departments of <sup>1</sup>Neurology and <sup>2</sup>Anesthesia, University of Iowa Carver College of Medicine; Departments of <sup>3</sup>Biostatistics and <sup>4</sup>Epidemiology, University of Iowa College of Public Health, Iowa City, IA; and <sup>5</sup>Department of Anesthesiology, University of Michigan College of Medicine, Ann Arbor, MI.

Received Apr 24, 2006, and in revised form Aug 16. Accepted for publication Sep 15.

Published online Nov 27, 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/ana.21018

Address correspondence to Dr Anderson, Department of Neurology, 2155 RCP, University of Iowa Carver College of Medicine, Iowa City, IA 52242. E-mail: steven-anderson@uiowa.edu

ing (ADL) Index. More than 60% of patients in both hypothermic and normothermic groups had “Good Outcomes” (GOS score = 1). Although the trial was large enough to detect a 10% absolute difference in the incidence of “Good Outcomes” (eg, 70 vs 60%), we were concerned from the onset that the GOS,<sup>14</sup> although well validated and widely used in studies of SAH,<sup>3,4,15</sup> might not be sufficiently sensitive.

Neuropsychological evaluations have long been used to aid in the clinical assessment of patients with neurological disease. There is also increasing evidence of the utility of neuropsychological tests for evaluating treatment effects in neurologically focused clinical trials. These include pharmacological interventions for neurodegenerative diseases and epilepsy, and chemotherapy and radiotherapy protocols for central nervous system cancer.<sup>16–19</sup> Moreover, several previous studies have shown that a high fraction of even “Good Outcome” SAH patients may have neuropsychological abnormalities.<sup>5,7,20–25</sup> Based on these considerations, we supplemented our primary and secondary outcome measures with a battery of neuropsychological tests. We now report the results of those measures.

## Subjects and Methods

The Intraoperative Hypothermia for Aneurysm Surgery Trial (IHAST) was a large, multicenter, prospective, randomized, partially blinded clinical trial designed to determine the impact of intraoperative hypothermia (target  $t = 33^{\circ}\text{C}$  vs  $t = 37^{\circ}\text{C}$ ) on neurological outcome in 1,000 patients undergoing surgical aneurysm clipping within 14 days after an acute aneurysmal SAH.

Details on trial design, patient eligibility, and study conduct have been published previously.<sup>13</sup> In brief, nonpregnant adults with a World Federation of Neurologic Surgeons<sup>26</sup> score of I, II, or III, a body mass index of less than  $35\text{kg}/\text{m}^2$ , and no cold-related disorders were eligible. Patients could not be endotracheally intubated at the time of preoperative assessment. Informed consent was obtained either from the patient or their legal representatives. A permuted block randomization scheme was used, stratified by center and time between SAH and surgery (0–7 days; 8–14 days). No more than 2 hours before the planned start of surgery patients were evaluated and enrolled via a telephone-accessed computer system, which directed the anesthesiologist to use a numbered opaque envelope containing the patient’s treatment assignment. The envelope was opened only after the induction of anesthesia. If, before induction, eligibility criteria were no longer met (eg, neurological deterioration), the envelope was not opened and the patient was not considered randomized. All study personnel, except the anesthesiologists involved in intraoperative care, were blinded to treatment assignment.

After endotracheal intubation and positioning, a retrocardiac esophageal temperature probe (Mon-a-therm Esophageal Stethoscope XL; Tyco Mallinckrodt, Pleasanton, CA) was inserted and patients were covered with a forced-air blanket connected to a PolarAir heating/cooling unit (Arizant, Eden

Prairie, MN). The use of a circulating water mattress and/or intravenous cold saline as cooling aids was optional. In patients randomized to hypothermia, esophageal temperature was reduced as quickly as possible, without any delay in the progress of surgery. The goal was to achieve a temperature between  $32.5$  and  $33.5^{\circ}\text{C}$  at the time the clip was applied to the first aneurysm. Temperature in patients randomized to normothermia was kept between  $36$  and  $37^{\circ}\text{C}$ . Rewarming of hypothermic patients began after the last aneurysm had been secured, and continued until normothermia was achieved.

After surgery, patients were extubated at the discretion of the care team. No attempt was made to control postoperative care, but all aspects of treatment were monitored for either 14 days or until discharge (if this occurred before 14 days). Patients were also assessed on the day of discharge.

Patients were contacted by telephone or seen in person 3, 6, and 9 weeks after surgery. A final follow-up examination was conducted approximately 3 months after surgery. At this time, an extensive battery of tests was completed including: (1) modified GOS (this was the primary outcome measure for the trial),<sup>14,27</sup> (2) Rankin Disability Scale,<sup>28</sup> (3) Barthel ADL Index,<sup>29</sup> (4) NIHSS,<sup>30</sup> and (5) a neuropsychology battery (see later). All evaluations were performed by trained examiners certified by the University of Iowa Steering Committee.

## Selection of Neuropsychological Tests

The purpose of the neuropsychological evaluation was to provide a standardized, quantitative assessment of performance in key domains of cognition, including memory, language, visuospatial and visuomotor abilities, and attention. The primary criterion in selecting the tests for the battery was that they be sensitive to cognitive deficits in patients with circumscribed lesions or more global cerebral dysfunction. Secondary criteria included that the battery be brief, relatively free of cultural biases, and amenable to administration in several languages. The tests chosen for our study were selected based on analysis of data from the Patient Registry of the Division of Cognitive Neuroscience, in the Department of Neurology at the University of Iowa. This database includes comprehensive neuropsychological testing and neuroimaging on more than 2,000 patients with focal brain lesions caused by vascular events or neurosurgical procedures. In selecting the IHAST test battery, we analyzed data from 94 patients (not included in IHAST) with SAH who had undergone comprehensive neuropsychological evaluations (generally including 15–30 individual tests). Each patient was classified by the evaluating neuropsychologist as cognitively impaired or not impaired. Each individual test score was coded as impaired or normal based on the criterion of two standard deviations (SDs) below the mean of the normative data. We then examined the sensitivity of various combinations of tests that met the above criteria. A battery of five tests that showed 95% agreement with the conclusion of the comprehensive neuropsychological evaluation was identified. The resulting battery included the Benton Visual Retention Test,<sup>31</sup> Controlled Oral Word Association (COWA),<sup>32</sup> Rey–Osterrieth Complex Figure–Copy,<sup>33</sup> Grooved Pegboard (GPB), and Trail Making Test.<sup>33,34</sup> In

addition, based on its use in other related trials,<sup>35</sup> a Mini-Mental State Examination (MMSE)<sup>36,37</sup> was also performed.

### *Neuropsychological Evaluation*

A neuropsychologist or psychometrician at each participating IHAST center was certified by a review of credentials, with the requirement of documented training at a reputable neuropsychological training program and at least 1 year of experience administering neuropsychological tests. Standardized test materials and instructions were provided to each center. Examiners were unaware of group assignment. If a patient could not be examined at the operating center, arrangements were made for them to be evaluated by trained examiners at another location. Patient age, years of education, and handedness were also recorded. All raw data were sent to the University of Iowa for scoring by a single technician (K.W.M.) who was also unaware of treatment group. Test performances were compared with appropriate age- and education-adjusted normative data<sup>31–34</sup> (J. S. Wefel and K. E. Boward. Normative data for the Complex Figure Test. Unpublished data. Department of Neurology, University of Iowa, Iowa City, IA, 1995). The following results were converted to *T*-scores: Benton Visual Retention Test number correct, COWA, Rey–Osterrieth Complex Figure–Copy, GPB right hand time to completion, GPB left hand time to completion, and Trail Making Test B. Individual *T*-scores were then averaged to provide a Composite Score, which provided an index of global cognitive status.

If all of the tests were not completed, a Composite Score was calculated based on those tests that were completed, provided at least three scores were available. If fewer than three of the tests were completed, a Composite Score was not calculated, but patients were classified as being globally impaired or not impaired by a process of imputation (see later), based on the results of their other neurological assessments.

Impairment on either an individual test or on the Composite Score was defined as a *T*-score of 30 or less, which is 2 SDs below the population norms. Impairment on the MMSE was also defined as a score two or more SDs below the age- and education-adjusted norm.<sup>37</sup>

**MISSING DATA/IMPUTATION OF IMPAIRMENT.** Neuropsychological testing (excluding the MMSE) was not completed in 66 (7%) surviving patients, 31 in the hypothermic group and 35 in the normothermic group. In 54 of these patients, no neuropsychological testing was performed, whereas 10 patients completed a single test (most commonly the COWA) and 2 patients completed two tests. However, on inspection, it was apparent that the characteristics of the untested patients differed substantially between the two treatment groups. For example, of the 31 hypothermic patients without Composite Scores, 18 (58%) were classified as “medically untestable” by the examiner, and 21 (68%) had GOSs of 3 or 4 (severe impairment or vegetative state, respectively). In contrast, only 12 normothermic patients (34%) were considered to be medically untestable or had GOSs of 3 or 4.

Patients without Composite Scores were distributed equally across all participating centers and across the duration of the study; no evident “pattern” suggesting bias could be identified. Nevertheless, neuropsychological impairment is

more likely in patients with severe neurological disability. Because excluding more “poor outcome” patients from one group has the potential to “bias” the results in favor of that group, it was necessary develop a process by which patients without Composite Scores could be included in the outcome assessment process.

To eliminate the possibility of arbitrary assignment and subjective bias, we developed a computerized imputation process to relate our four principal outcome measures (GOS, Rankin Disability Scale, NIHSS, and Barthel ADL) to neuropsychological outcomes; at least two of these test results were available for all patients. The imputation process used data from the 873 patients who did have Composite Scores. We used logistic regression to relate our four neurological outcome measures to the presence or absence of neuropsychological impairment on the Composite Score. Because at least 2 of these neurological outcome measures were available for each of the 66 individuals without Composite Scores, it was possible to determine the likelihood that any given patient would have been classified as impaired, compared with the group of patients with Composite Scores. For example, a patient with GOS = 3, Rankin Disability Scale score = 4, NIHSS = 18, and Barthel ADL = 20 had a 99.99% likelihood of being impaired, whereas an individual with GOS = 1, Rankin Disability Scale score = 1, NIHSS = 0, and Barthel ADL = 100 had only an 11% chance of being impaired.

This algorithm was then applied to the 66 patients with missing Composite Scores. Patients with a *calculated* likelihood of impairment of 80% or more were classified as “impaired.” This cutoff was selected to ensure that only patients with a high likelihood of impairment were so classified; the imputation algorithm, when tested on patients with measured Composite Scores, had a high specificity (99.3%), a relatively low sensitivity (approximately 20.6%), and high positive and negative predictive value (86.8 and 84.8%, respectively). As such, it intentionally underestimates the actual number of impaired patients among individuals with missing Composite Scores.

Note that this process was used *only* to impute the likelihood of impairment on the Composite Score; no effort was made to impute values for actual numerical Composite Scores or impairment on individual test results.

### *Data Analysis*

All data entry and analyses were performed by the Data Management Center at the University of Iowa.

The impact of treatment (hypothermia vs normothermia) on neuropsychological outcome was assessed using three principal measures. First, the distribution of outcomes based on Composite Score (unimpaired, impaired, and dead) were compared using the Cochran–Mantel–Henzel general association test, adjusting for center, stratum, and treatment.<sup>38</sup> As noted earlier, a patient was considered to be “impaired” if his or her Composite Score was 30 or less (2 or more SDs below the norm) or, in the event that a Composite Score was not available, based on the previously noted imputation process.

Second, the distribution of outcomes based on whether a patient was impaired on *one or more* neuropsychological tests was compared using a similar process. A patient with impair-

ment on any single test would be considered as “impaired on one or more tests” regardless of results on any other tests (or the absence of any other test results). If no test results were available, and the imputation process noted above classified the patient as “impaired” on the Composite Score, they were also considered to be “impaired on one or more test.” If they were classified as “unimpaired” by imputation, they were also classified as “unimpaired” on one or more tests.

Third, the results for individual tests (*T*-scores) and composite scores were expressed as mean  $\pm$  SD. No effort was made to impute individual or Composite Scores, and because of imbalances in characteristics of missing patients, no between-groups statistical comparisons were performed. To explore the relative sensitivities of the different tests, we performed within-group comparisons of individual *T*-scores and the incidence of impairment on individual tests using analysis of variance.

Finally, we performed the identical evaluations on the subpopulation of patients with WFNS score of I at the time of enrollment. We focused on this subpopulation because these patients have the least amount of posthemorrhage physiological and functional disturbance. As such, they have the “most to lose” for a result of events during surgery (hypotension, brain retraction, temporary vessel occlusion) or afterward (eg, vasospasm). If intraoperative hypothermia were to confer neurological protection and preserve remaining neurological and/or neuropsychological function, the treatment effect might be most obvious in the WFNS I population.

## Results

Between February 2000 and April 2003, 3,966 patients underwent intracranial aneurysm surgery at 30 participating centers. Of this group, 2,856 had suffered an acute SAH. A total of 1,183 of these patients were eligible, and 1,033 were enrolled. Due to changes in status after enrollment, 32 enrolled patients were not randomized, resulting in a total of 1,001 subjects. Three-month GOS scores were obtained in 1,000 patients (499 hypothermia and 501 normothermia). Sixty-one patients died (29 in the hypothermic group and 32 in the normothermic group). Among the surviving patients, Rankin Disability Scale scores were available for all 939, Barthel ADLs were available for 938, and NIHSS scores were available for 913. At least 1 neu-

ropsychological test was completed by 885 patients, and a Composite Score was available for 873 (93% of survivors; 439 in the hypothermic group and 434 in the normothermic group). MMSEs were completed in 883 patients (448 hypothermia and 435 normothermia).

Baseline characteristics of the efficacy population, as well as nonneuropsychological outcomes, can be found in our previous publication.<sup>13</sup> In the hypothermic and normothermic groups, 65 and 66% of patients were female, mean ( $\pm$ SD) ages were  $52 \pm 12$  and  $51 \pm 13$  years, and average years of education were  $11.6 \pm 2.9$  and  $11.7 \pm 2.7$  years, respectively. The frequencies of premorbid neurological disease, psychiatric disease, and alcohol abuse did not differ between groups (Table 1). Other than the planned differences in intraoperative temperatures ( $33.2 \pm 0.9^\circ\text{C}$  vs  $36.5 \pm 0.7^\circ\text{C}$ , mean  $\pm$  SD), significant intergroup differences were noted for blood glucose at the time of first clip application (hypothermia vs normothermia:  $138 \pm 37$  vs  $127 \pm 31\text{mg/dl}$ ), total operative crystalloid administration (hypothermia vs normothermia:  $3.8 \pm 1.6$  vs  $3.3 \pm 1.5\text{L}$ ), and urine output (hypothermia vs normothermia:  $2.1 \pm 1.3$  vs  $1.7 \pm 1.6\text{L}$ ). A total of 25.1% of hypothermic patients remained intubated at 2 hours after surgery versus 13.2% of normothermic patients. There were no intergroup differences in GOS, Rankin Disability Scale score, Barthel ADL, or NIHSS; 66% of hypothermic group patients had a GOS = 1, versus 63% of normothermic group patients ( $p = 0.32$ ; odds ratio, 1.14 [95% confidence interval, 0.88–1.48]).

### Neuropsychological Outcomes

**IMPAIRMENT ON COMPOSITE SCORE.** As noted, 873 of the 939 surviving patients completed 3 or more individual tests and could be directly classified as “impaired” or “unimpaired” from their Composite Scores (Table 2). Impairment status was imputed in the remaining 66 based on the process described earlier. In the hypothermic group, 16.8% of patients were classified as “impaired” (17.9% of surviving patients) compared with 20.0% in the normothermic group (21.3%

Table 1. Presubarachnoid Hemorrhage History

Premorbid History	Treatment		Total (N = 873)	<i>p</i>
	Hypothermia (n = 439)	Normothermia (n = 434)		
Neurological disease	46 (10.5%)	29 (6.7%)	75 (8.6%)	0.053
Psychiatric disease	7 (1.6)	6 (1.4%)	13 (1.5%)	1.00
Alcohol abuse	34 (7.7%)	33 (7.6%)	67 (7.7%)	1.00

These data represent those subjects who completed sufficient testing to generate a Composite Score. Neurological disease includes prior history of stroke, tumor, head trauma, epilepsy, or other major neurological disease. There were no significant differences between treatment groups based on Fisher’s exact test.

Table 2. Impairment on Composite Score

Impairment	Hypothermia <sup>a</sup>			Normothermia <sup>a</sup>		
	Tested <sup>b</sup>	Untested and Imputed <sup>c</sup>	Total	Tested <sup>b</sup>	Untested and Imputed <sup>c</sup>	Total
Unimpaired, n	375	11	386 (77.4%)	345	24	369 (73.7%)
Impaired, n	64	20	84 (16.8%)	89	11	100 (20.0%)
Dead, n			29 (5.8%)			32 (6.4%)
Total, n	439	31	499	434	35	501

The Composite Score is the average of the individual *T*-scores for the six administered tests.

<sup>a</sup>There was no significant difference in the distribution of outcomes between hypothermia and normothermic groups (CMH general association test,  $p = 0.317$ ).

<sup>b</sup>Patients completed 3 or more tests; impairment indicates a score that is 2 or more standard deviations below the norm ( $T$ -score  $\leq 30$ ).

<sup>c</sup>Patients did not complete neuropsychological testing, and their impairment status was determined using a logistic regression imputation process based on their recorded Glasgow Outcome Scale, Rankin Disability Scale, Barthel's Activities of Daily Living index, and National Institutes of Health Stroke Scale scores.

of surviving patients). There were no intergroup differences in the distribution of outcomes ( $p = 0.317$ ). The relationship between outcome as assessed by GOS and neuropsychological impairment is shown in Table 3.

**IMPAIRMENT ON ONE OR MORE INDIVIDUAL TESTS.** Approximately 50% of patients in both groups were classified as being "impaired" on one or more individual neuropsychological tests (Table 4). There were no differences in the distribution of outcomes ( $p = 0.865$ ).

**INDIVIDUAL TEST RESULTS.** The frequencies of impairment in each cognitive domain assessed, for the two treatment groups combined, are presented in Table 5. The results for individual tests within each treatment group are shown in Table 6. The Composite Score (based only on tested patients) was  $41.0 \pm 10.0$  (mean  $\pm$  SD) in the hypothermic group versus  $38.7 \pm 10.8$  in the normothermic group. Mean *T*-scores on individual tests ranged from  $35.5 \pm 13.3$  (GPB-Left, normothermia) to  $45.2 \pm 11.8$  (COWA, hypothermia). The fraction of patients with impairment on a given test ranged from a low of 9% (COWA, hypothermia) to 38% (GPB-Left, normothermia).

**MINI-MENTAL STATE EXAMINATION.** In the hypothermic group, 448 of the 470 surviving patients completed the MMSE, compared with 435 of the 469 normothermic patients. Mean ( $\pm$ SD) scores for those completing the test were  $27.4 \pm 3.8$  in the hypothermic group versus  $26.8 \pm 4.5$  in the normothermic group. Of the tested survivors in the hypothermic group, 14.5% were classified as impaired versus 18.4% in the normothermic group. However, as for patients with missing Composite Scores, the distribution of outcomes in patients not completing the MMSE dif-

fered between the two treatment groups. Of the 22 hypothermic patients without MMSEs, 4 (18%) had GOSs = 1. By contrast, of the 34 normothermic patients without MMSEs, 20 (59%) had GOSs = 1. As a result, no direct statistical comparison of the two treatment groups was performed. An imputation algorithm was applied to the 56 patients with missing MMSEs, similar to that developed for the Composite Score. Patients with a calculated likelihood of impairment of 80% or more were classified as "impaired" on the MMSE. In the combined group of subjects with MMSE scores and imputed impairment status, there were no intergroup differences in the distribution of outcomes ( $p = 0.386$ ).

**WFNS = I SUBPOPULATION.** Analyses similar to those above were performed on the 660 patients (332 hypothermic and 328 normothermic patients) with the least degree of initial impairment; that is, those with an admission WFNS scores of I. The results were similar to the entire study population. In the hypothermic group, 13.3% of patients were classified as "impaired" on the Composite Score versus 18.3% of normothermic group patients ( $p = 0.099$ ). Likewise, impairment on one or more tests was equivalent between treatment groups (hypothermia: 53.6%; normothermia 52.7%;  $p = 0.970$ ).

## Discussion

The IHAST study tested the hypothesis that mild intraoperative hypothermia would increase the incidence of "good outcomes" in patients undergoing surgical therapy (clipping) for aneurysmal SAH as assessed by the GOS at 3 months after surgery. As reported previously, no treatment effects were detected despite a sufficiently large trial, high protocol compliance, and excellent follow-up. In addition, no differences were seen in several secondary measures of neurological function,

Table 3. Impairment versus Glasgow Outcome Scale

GOS	Hypothermia					Normothermia					Grand Total
	Unimpaired	Impaired	Untested, Impaired	Untested, Unimpaired	Total	Unimpaired	Impaired	Untested, Impaired	Untested, Unimpaired	Total	
1	303	19	0	7	329	262	34	0	18	314	643
2	70	32	1	2	105	73	30	0	5	108	213
3	2	13	18	2	35	10	25	11	1	47	82
4	0	0	1	0	1	0	0	0	0	0	1
5					29					32	61
Total					499					501	

The relationship between Glasgow Outcome Scale (GOS) and Neuropsychological Impairment on the Composite Score. The columns labeled as “untested” are divided into impaired or unimpaired according to the imputation algorithm described in Subjects and Methods.

Table 4. Impairment on One or More Individual Tests

	Hypothermia <sup>a</sup>			Normothermia <sup>a</sup>		
	Tested <sup>b</sup>	Untested and Imputed <sup>c</sup>	Total	Tested <sup>b</sup>	Untested and Imputed <sup>c</sup>	Total
Unimpaired	188	9	197 (39.5%)	166	24	190 (37.9%)
Impaired	258	15	273 (54.7%)	268	9	279 (55.7%)
Dead			29 (5.8%)			32 (6.4%)
Total	446	24	499	436	33	501

<sup>a</sup>There were no significant differences in the distribution of outcomes between hypothermia and normothermic groups (CMH General Association Test,  $p = 0.865$ ).

<sup>b</sup>A “tested” patient was classified as impaired on one or more individual tests if the  $T$ -score for *any* test was two or more standard deviations below the norm, regardless of the actual number of tests completed.

<sup>c</sup>An “untested” patient was classified as impaired *only* if they had failed to take any tests *and* if their imputed Composite Score was impaired.

Table 5. Cognitive Functions Affected

Function	Test	Tested, n	Fraction Impaired	Untested, n
Short-term memory	BVRT	869	27%	70
Cued word finding	COWA	868	12%	71
Visuospatial/constructional ability	CFT	864	28%	75
Visuomotor coordination (RH)	GPB Rt	852	32%	87
Visuomotor coordination (LH)	GPB Lt	838	33%	101
Executive function	TMT-B	852	18%	87

Impairments were common across all cognitive domains assessed. Most frequently affected were visuomotor coordination (assessed with a timed unimanual pegboard task), visuospatial/constructional ability (assessed through paper and pencil copy of a complex geometric figure), and short-term memory (assessed with the immediate reproduction of a series of increasingly complex geometric figures, after brief visual exposure).

BVRT = Benton Visual Retention Test; COWA = Controlled Oral Word Association; CFT = Complex Figure Test; RH = right hand; GPB = Grooved Pegboard; LH = left hand; TMT-B = Trail Making Test Part B.

including the Rankin Disability Scale, the NIHSS, and the Barthel ADL. We can hence conclude that intra-operative cooling does not improve functional and/or neurological outcome in this group of patients.

The GOS was chosen as our primary outcome assessment instrument for IHASt based on its prior use in many SAH-related trials<sup>15,39–44</sup> We chose the Rankin Disability Scale as a secondary outcome for largely the same reasons, that is, its widespread use and its demonstrated ability to respond to treatment interventions.<sup>45</sup> However, both scores are relatively crude,

focusing on general functionality rather than on discrete neurological and/or cognitive abnormalities. They may not be sensitive to small but possibly important differences between groups. Another secondary assessment, the NIHSS,<sup>30</sup> provides considerably more neurological detail but is heavily weighted toward focal motor abnormalities. It was designed to evaluate patients with occlusive stroke, and recent work suggests that it has limitations as an outcome measure in SAH.<sup>46</sup> As a result, we were concerned from the outset that our neurological outcome measures might not be

Table 6. Individual Test Results

Test	Hypothermia				Normothermia			
	Tested, n <sup>a</sup>	T-score <sup>b</sup>	Fraction Impaired <sup>c</sup>	Untested, n	Tested, n <sup>a</sup>	T-score <sup>b</sup>	Fraction Impaired <sup>c</sup>	Untested, n
BVRT	436	42.3 ± 14.1	23%	34	433	38.5 ± 15.0	31%	36
COWA	434	45.2 ± 11.8	9%	36	434	43.5 ± 12.3	15%	35
CFT	437	39.9 ± 18.5	27%	33	427	38.3 ± 19.0	29%	42
GPB Rt	427	37.6 ± 13.2	30%	43	425	35.6 ± 13.9	33%	44
GPB Lt	419	37.1 ± 12.7	28%	51	419	35.3 ± 13.3	38%	50
TMT-B	424	43.7 ± 12.5	13%	46	428	41.6 ± 13.3	22%	41
Composite	439	41.0 ± 10.0	15%	31	434	38.7 ± 10.8	21%	35

Because of the different characteristics of “untested” patients in the hypothermic and normothermic groups, no statistical comparisons were performed on individual test scores or on the numeric Composite Score.

<sup>a</sup>“Tested” indicates the number of patients who completed a given test.

<sup>b</sup>The T-score (mean ± standard deviation [SD]) is based on comparing the patient’s age-, sex-, and education-adjusted raw scores with published norms.

<sup>c</sup>The “fraction impaired” indicates the percentage of patients whose T-score was ≤30 (which is 2 SD below the norm).

BVRT = Benton Visual Retention Test; COWA = Controlled Oral Word Association; CFT = Complex Figure Test; GPB Rt and Lt = Grooved Pegboard, right and left hand; TMT-B = Trail Making Test Part B; Composite = Composite Score, which is the average of all individual T-scores for a patient.

able to detect subtle and yet nevertheless important treatment effects.

We did not choose a neuropsychological outcome as our primary outcome for several reasons. First, neuropsychological testing has not previously been used as a primary outcome measure in SAH. There is no widely accepted “standard” battery of neuropsychological tests for use in this situation. In fact, even the definition of “impairment” on such tests after SAH has not been well defined. Nevertheless, we had reason to believe that neuropsychological testing might be more sensitive to SAH-related injury, and hence treatment. There have been many published studies demonstrating a high incidence of neurocognitive deficits among survivors of SAH, even those with good functional outcomes. For example, Hutter and colleagues<sup>20</sup> demonstrated that 28% to 62% of “Good Outcome” (by GOS) patients studied 6 months after their SAH demonstrated abnormalities on one or more neuropsychological tests. Others have also demonstrated a high incidence of cognitive abnormalities after SAH.<sup>5,7,21–25</sup> In addition, neuropsychological testing has been shown to reflect treatment effects for other disorders.<sup>16–19</sup> For this reason, we included a brief battery of neuropsychological tests.

Our first challenge was to develop a battery of tests that could be administered easily on a large multicenter/multinational scale. Standard neuropsychological testing is often a long and time-consuming procedure, and many tests are language dependent. We believed that to be feasible in a large multicenter trial such as this one, a shorter and more easily administered test battery was essential. It was also important to select a series of tests that covered the major cognitive domains

and that were known to be frequently affected in patients recovering from SAH. This selection process was accomplished by using the Patient Registry of the Division of Cognitive Neuroscience, in the Department of Neurology at the University of Iowa. Based on a review of data from patients with SAH, we selected five well-known and widely used tests whose results correlated well with results of a more comprehensive test battery. We also included the MMSE largely because of its widespread use as an outcome measure in relevant studies.<sup>35,46</sup> The tests were administered in each center by trained neuropsychologists or neuropsychology technicians whose qualifications were all reviewed by the Iowa Steering Committee. We intentionally chose a relatively “conservative” definition of “impairment,” that is, a T-score of two or more SD below the age-, sex-, and education-adjusted norm.

The findings from this study provide support for the feasibility of incorporating neuropsychological testing into the outcome measures for clinical trials in which changes in cognitive function might be expected. However, it also highlights the limitations of such testing in trials where a substantial number of patients might be expected to have severe brain damage. We had no difficulty recruiting trained neuropsychologists and neuropsychology technicians at our participating centers, all of which were major medical centers with neurosurgical units. No potential center was excluded because neuropsychology support was not available. The battery of tests proved to be readily performed with 93% of surviving patients completing testing. This represents the largest prospective assessment of neuropsychological abnormalities after SAH currently published. However, there were problems. In any clinical trial, the

utility of an outcome measure is dependent on the success with which it can be applied to the largest possible fraction of treated patients. Failure to test a significant proportion of randomized patients can introduce bias if the characteristics of those untested patients differ between treatment groups or are dissimilar from the overall study group. This was not a problem for our primary outcome measure (GOS); only 1 patient was lost of a total of 1,001 randomized patients. However, roughly 7% of our surviving patients did not undergo complete neuropsychological testing, and it is clear that this untested group was differed from the overall efficacy population. In the overall trial, 69% of the 939 surviving patients were classified as GOS = 1, 23% as GOS = 2 (moderate disability), and 9% as GOS = 3 (severe disability). By contrast, of the 66 patients with missing Composite Scores, 39% were classified as GOS = 1 and 48% were GOS = 3. Of those patients missing a Composite Score, 45% (30/66) were classified as “medically untestable,” indicting that neuropsychological testing of the sort used in this trial is not readily completed by severely injured patients, even when other outcome measures (GOS, NIHSS) can be assigned.

Use of this form of neuropsychological testing as a primary outcome measure in clinical trials that involve a significant number of severely brain damaged patients could result in a relatively high rate of loss to follow-up. One approach to this problem would be to incorporate simpler cognitive behavioral tests that might be completed by more patients (eg, questions regarding orientation to time and place, repeating digits, simple yes/no questions). However, in this study, the number of patients who completed the MMSE, which uses such simple items, was nearly identical to the number completing at least one neuropsychological test. Furthermore, the MMSE was less sensitive to cognitive deficits than was the neuropsychological testing; of the 151 patients with impaired Composite Scores who also completed the MMSE, only 88 (58%) were found to be impaired on the MMSE.

It is unclear whether testing of cognitive function adds to our assessment of patients with severe damage. A better approach might be to restrict neuropsychological testing a priori to predefined “eligible patients” with otherwise good outcomes according to neurological rating scales. This would preclude the use of neuropsychological testing as a primary outcome measure for many trials, but would take advantage of the unique information provided by neuropsychology as a secondary outcome measure. If we used this approach with our own data, 5.8% of hypothermic GOS = 1 patients were cognitively impaired (by Composite Score) versus 10.8% of normothermic patients; 49.2% and 49.7% of hypothermic and normothermic GOS = 1 patients were impaired on at least one test.

Of even greater importance to this study was that the distribution of neurological outcomes in patients with missing Composite Scores differed significantly between the two treatment groups (see Table 4). We were unable to identify any cause of this maldistribution and concluded that this was a consequence of the randomization process. This maldistribution problem made it critical that, if we wished to use the Composite Score as a global measure of neuropsychological outcome, some method be used to estimate the likelihood of impairment in patients with missing Composite Scores; without doing so, our results would be inconsistent with the primary outcome. That is, if we analyze *only* those patients with Composite Scores, we would conclude that hypothermia resulted in a significantly greater fraction of GOS = 1 patients than normothermia, a finding that conflicts with our primary result.<sup>13</sup> Similarly, we would conclude that neuropsychological outcomes were better in hypothermic group patients. However, we believe that this conclusion would be incorrect and misleading, because it would be due solely to the imbalance in the follow-up procedure, and *not* to a direct treatment effect. For this reason, we developed a logistic, regression-based imputation process. Although one might argue with the cut point for designation of “impairment” on the Composite Score ( $\geq 80\%$  likelihood of impairment), this approach avoided introducing investigator bias or subjectivity into the determination of “impairment” in patients with missing Composite scores. However, we did not consider the imputation process to be sufficiently reliable as to impute more specific data such as: (1) actual Composite Score values in untested patients, or (2) estimate scores for missing individual neurocognitive tests. Thus, because of missing individual scores and the unreliability of imputed individual scores, our ability to reliably compare the effect of intraoperative hypothermia on *individual* neurocognitive test scores was restricted.

Despite these limitations, the neuropsychological battery proved to be sensitive to cognitive deficits that were not apparent by the use of global outcome measures. Although the incidence of impairment on our Composite Score was lower than the incidence of less-than-good outcomes on the GOS (18.4% impairment vs 35.7% for GOS values of  $>1$ ), a large fraction of GOS = 1 patients had some neurocognitive impairment. For example, of the 643 patients categorized as “good outcome” (GOS = 1), 53 (8.2%) were classified as “impaired” on Composite Score, and 318 (49.5%) were “impaired” on at least 1 test. These results are generally consistent with the results of others who have studied SAH.

One contemporary question of interest to clinicians involves the relative importance of SAH itself versus treatment on outcome. Are long-term deficits (neuro-



logical and neuropsychological) due to the SAH itself, or are they modified by treatment? Recent studies suggest a modestly lower morbidity and mortality among patients treated with coiling versus surgery, but it is also impossible to study untreated patients with SAH. The IHAST study was not designed to address this issue, but some insight may be gained by examining outcome in patients with differing “baseline” neurological conditions, that is, those with prerandomization WFNS values of 1 versus poorer outcomes. Theoretically, “surgical” injury would be more evident in patients with the least degree of presurgical dysfunction (and who would be predicted to have the best outcomes). Among our 660 WFNS = 1 patients, 473 (71.7%) were classified as having a “good outcome” (GOS = 1), 25 patients died (3.8%), and 103 (15.6%) were “impaired” on their Composite Score. Conversely, of the 340 patients with baseline WFNS = 2 or 3, 170 (50%) had good outcomes, 36 (10.6%) died, and 88 (20.5%) were impaired. These differences are statistically significant and support the idea that a major fraction of a patient’s outcome is defined by the underlying severity of their SAH.

In summary, our results do not support the hypothesis that intraoperative hypothermia improves neurocognitive outcome in good-grade SAH patients undergoing surgery for the treatment of ruptured intracranial aneurysms. They do, however, demonstrate that neuropsychological testing is a feasible outcome measure. Like other investigators, our results suggest that there is a high incidence of neuropsychological deficits in this population, even among patients considered to have otherwise good outcomes. This result, combined with the problems we encountered with follow-up (particularly in poor-grade patients), suggests that neuropsychological testing may best be used to refine our outcome assessments in good-grade patients rather than as a “global” assessment of outcome.

---

This work is supported by the NIH (National Institute of Neurological Disease and Stroke, RO1 NS38554, M.M.T.).

---

## References

1. Kassell NF, Torner JC, Haley EC, et al. The International Cooperative Study on the timing of aneurysm surgery. Part 1: overall management results. *J Neurosurg* 1990;73:18–36.
2. Saveland H, Hillman J, Brandt L, et al. Overall outcome in aneurysmal subarachnoid hemorrhage. A prospective study from neurosurgical units in Sweden during a 1-year period. *J Neurosurg* 1992;76:729–734.
3. Haley EC, Kassell NF, Torner JC, and the Participants. A randomized controlled trial of high-dose intravenous nicardipine in aneurysmal subarachnoid hemorrhage—a report of the Cooperative Aneurysm Study. *J Neurosurg* 1993;78:537–547.
4. Kassell NF, Haley EC, Apperson-Hansen C, et al. Randomized, double-blind, vehicle controlled trial of tirilazad mesylate in patients with aneurysmal subarachnoid hemorrhage: a cooperative study in Europe, Australia and New Zealand. *J Neurosurg* 1996;84:221–228.
5. Ogden JA, Mee EW, Henning M. A prospective study of impairment of cognition and memory and recovery after subarachnoid hemorrhage. *Neurosurgery* 1993;33:572–587.
6. Hutter BO, Kreitschmann-Andermahr I, Mayfrank L, et al. Functional outcome after aneurysmal subarachnoid hemorrhage. *Acta Neurochir Suppl* 1999;72:157–174.
7. Hackett ML, Anderson CS. Health outcomes 1 year after subarachnoid hemorrhage: an international population-based study. The Australian Cooperative Research on Subarachnoid Hemorrhage Study Group. *Neurology* 2000;55:658–662.
8. Bendtsen AO, Cold GE, Astrup J, Rosenorn J. Thiopental loading during controlled hypotension for intracranial aneurysm surgery. *Acta Anaesthesiol Scand* 1984;28:473–477.
9. Batjer HH, Frankfurt AI, Purdy PD, et al. Use of etomidate, temporary arterial occlusion, and intraoperative angiography in surgical treatment of large and giant cerebral aneurysms. *J Neurosurg* 1988;68:234–240.
10. Olgilvy CS, Rich C, Crowell RM. Mild hypothermia, mannitol, and induced hypotension for cerebral protection during temporary vessel occlusion in aneurysm surgery is safe and effective. *J Neurosurg* 1993;78:367A.
11. Ravussin P, Detribolet N. Total intravenous anesthesia with propofol for burst suppression in cerebral aneurysm surgery—preliminary report of 42 patients. *Neurosurgery* 1993;32:236–240.
12. Pemberton PL, Dinsmore J. The use of hypothermia as a method of neuroprotection during neurosurgical procedures and after traumatic brain injury: a survey of clinical practices in Great Britain and Ireland. *Anaesthesia* 2003;58:371–373.
13. Todd MM, Hindman BJ, Clarke WR, et al. Mild intraoperative hypothermia during surgery for intracranial aneurysm. *N Engl J Med* 2005;353:135–145.
14. Jennett B, Bond M. Assessment of outcome after severe brain damage: a practical scale. *Lancet* 1975;1:480–484.
15. Pickard JD, Murray GD, Illingworth R, et al. Effect of oral nimodipine on cerebral infarction and outcome after subarachnoid haemorrhage: British aneurysm nimodipine trial. *Br Med J* 1989;298:636–642.
16. Kockelmann E, Elger CE, Helmstaedter C. Significant improvement in frontal lobe associated neuropsychological functions after withdrawal of topiramate in epilepsy patients. *Epilepsy Res* 2003;54:171–178.
17. Torres IJ, Mundt AJ, Sweeney PJ, et al. A longitudinal neuropsychological study of partial brain radiation in adults with brain tumors. *Neurology* 2003;60:1113–1118.
18. Tucha O, Smely C, Preier M, et al. Preoperative and postoperative cognitive functioning in patients with frontal meningiomas. *J Neurosurg* 2003;98:21–31.
19. Lu PH, Masterman DA, Mulnard R, et al. Effects of testosterone on cognition and mood in male patients with mild Alzheimer disease and healthy elderly men. *Arch Neurol* 2006;63:177–185.
20. Hutter BO, Gilsbach JM. Which neuropsychological deficits are hidden behind a good outcome (Glasgow = 1) after aneurysmal subarachnoid hemorrhage? *Neurosurgery* 1993;33:999–1005.
21. Hutter BO, Kreitschmann-Andermahr I, Gilsbach JM. Health-related quality of life after aneurysm subarachnoid hemorrhage: impacts of bleeding severity, computerized tomography findings, surgery, vasospasm, and neurologic grade. *J Neurosurgery* 2001;94:241–251.

22. Bjeljac M, Keller E, Regard M, Yonekawa Y. Neurological and neuropsychological outcome after SAH. *Acta Neurochir Suppl* 2002;82:83–85.
23. Mayer SA, Kreiter KT, Copeland D, et al. Global and domain-specific cognitive impairment and outcome after subarachnoid hemorrhage. *Neurology* 2002;59:1750–1758.
24. Powell J, Kitchen N, Heslin J, Greenwood R. Psychosocial outcomes at three and nine months after good neurological recovery from aneurysmal subarachnoid haemorrhage: predictors and prognosis. *J Neurol Neurosurg Psychiatry* 2002;72:772–781.
25. Vilkki JS, Juvela S, Siironen J, et al. Relationship of local infarctions to cognitive and psychosocial impairments after aneurysmal subarachnoid hemorrhage. *Neurosurgery* 2004;55:790–803.
26. Drake CG. Report of World Federation of Neurological Surgeons Committee on a universal subarachnoid hemorrhage grading scale. *J Neurosurg* 1988;68:985–986.
27. The Publications Committee for the Trial of ORG 10172 in Acute Stroke Treatment (TOAST) Investigators. Low molecular weight heparinoid, ORG 10172 (Danaproid), and outcome after acute ischemic stroke. A randomized controlled trial. *JAMA* 1998;279:1256–1272.
28. Rankin J. Cerebral vascular accidents in patients over the age of 60. II. Prognosis. *Scott Med J* 1957;2:200–215.
29. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Md State Med J* 1965;14:61–64.
30. Wityk RJ, Pessin MS, Kaplan RF, Caplan LR. Serial assessment of acute stroke using the NIH stroke scale. *Stroke* 1994;25:362–365.
31. Sivan AB. The Benton Visual Retention Test. 5th ed. San Antonio, TX: The Psychological Corporation, 1992.
32. Benton AL, Hamsher KDS. The Multilingual Aphasia Examination. Iowa City, IA: AJA Associates, 1994.
33. Lezak M. Neuropsychological Assessment. 3rd ed. New York: Oxford University Press, 1995:11:475–480.
34. Heaton RK, Grant I, Matthews CG. Comprehensive norms for an expanded Halstead-Reitan Battery: demographic corrections, research findings and clinical applications. Odessa, FL: Psychological Assessment Resources, 1991.
35. International Study of Unruptured Intracranial Aneurysms Investigators. Unruptured intracranial aneurysms—risk of rupture and risks of surgical intervention. *N Engl J Med* 1998;339:1725–1733.
36. Folstein MF, Folstein SE, McHugh PR. “Mini-Mental State”: a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975;12:189–198.
37. Crum RM, Anthony C, Bassett SS, Folstein MF. Population-based norms for the Mini-Mental State Examination by age and educational level. *JAMA* 1993;269:2386–2391.
38. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959;22:719–748.
39. Haley EC Jr, Kassell NF, Apperson-Hansen C, et al. A randomized, double-blind, vehicle-controlled trial of tirilazad mesylate in patients with aneurysmal subarachnoid hemorrhage: a cooperative study in North America. *J Neurosurg* 1997;86:467–474.
40. Lanzino G, Kassell NF, Dorsch NW, et al. Double-blind, randomized, vehicle-controlled study of high-dose tirilazad mesylate in women with aneurysmal subarachnoid hemorrhage. Part I. A cooperative study in Europe, Australia, New Zealand, and South Africa. *J Neurosurg* 1999;90:1011–1017.
41. Hillman J, Fridriksson S, Nilsson O, et al. Immediate administration of tranexamic acid and reduced incidence of early rebleeding after aneurysmal subarachnoid hemorrhage: a prospective randomized study. *J Neurosurg* 2002;97:771–778.
42. Lanzino G, Kassell NF, Germanson T, et al. Plasma glucose levels and outcome after aneurysmal subarachnoid hemorrhage. *J Neurosurg* 1993;79:885–891.
43. Roos Y. Antifibrinolytic treatment in subarachnoid hemorrhage: a randomized placebo-controlled trial. STAR Study Group. *Neurology* 2000;54:77–82.
44. Shibuya M, Suzuki Y, Sugita K, et al. Effect of AT877 on cerebral vasospasm after aneurysmal subarachnoid hemorrhage. Results of a prospective placebo-controlled double-blind trial. *J Neurosurg* 1992;76:571–577.
45. International Subarachnoid Aneurysm Trial (ISAT) Collaborative Group. International subarachnoid aneurysm trial (ISAT) of neurosurgical clipping versus endovascular coiling in 2143 patients with ruptured intracranial aneurysms: a randomised trial. *Lancet* 2002;360:1267–1274.
46. Kim DH, Haney CL, Van Ginhoven G. Utility of outcome measures after treatment for intracranial aneurysms: a prospective trial involving 520 patients. *Stroke* 2005;36:792–796.