

Protecting the Infant Brain During Cardiac Surgery: A Systematic Review

Jennifer C. Hirsch, MD, MS, Marshall L. Jacobs, MD, Dean Andropoulos, MD, Erle H. Austin, MD, Jeffrey P. Jacobs, MD, Daniel J. Licht, MD, Frank Pigula, MD, James S. Tweddell, MD, and J. William Gaynor, MD

Department of Cardiac Surgery, University of Michigan Medical Center, Ann Arbor, Michigan; Department of Pediatric and Congenital Heart Surgery, Cleveland Clinic, Cleveland, Ohio; Departments of Anesthesiology and Pediatrics, Division of Pediatric Cardiovascular Anesthesiology, Texas Children's Hospital/Baylor College of Medicine, Houston, Texas; Department of Surgery, Division of Cardiothoracic Surgery, University of Louisville and Kosair Children's Hospital, Louisville, Kentucky; Department of Cardiac Surgery, The Congenital Heart Institute of Florida (CHIF), All Children's Hospital, University of South Florida College of Medicine, Saint Petersburg and Tampa, Florida; Department of Pediatrics, Division of Neurology, and Department of Surgery, Division of Pediatric Cardiothoracic Surgery, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania; Department of Cardiac Surgery, Children's Hospital of Boston, Boston, Massachusetts; and Department of Surgery, Division of Cardiothoracic Surgery, Medical College of Wisconsin, Milwaukee, Wisconsin

Prevention of brain injury during congenital heart surgery has focused on intraoperative and perioperative neuroprotection and neuromonitoring. Many strategies have been adopted as "standard of care." However, the strength of evidence for these practices and the relationship to long-term outcomes are unknown.

We performed a systematic review (January 1, 1990 to July 30, 2010) of neuromonitoring and neuroprotection strategies during cardiopulmonary bypass (CPB) in infants of age 1 year or less. Papers were graded individually and as thematic groups, assigning evidence-based medicine and American College of Cardiology/American Heart Association (ACC/AHA) level of evidence grades. Consensus scores were determined by adjudication.

Literature search identified 527 manuscripts; 162 met inclusion criteria. Study designs were prospective observational cohort (53.7%), case-control (21.6%), randomized clinical trial (13%), and retrospective observational cohort (9.9%). Median sample size was 43 (range 3 to 2,481).

Primary outcome was evidence of structural brain injury or functional disability (neuroimaging, electroencephalogram, formal neurologic examination, or neurodevelopmental testing) in 43%. Follow-up information was reported in only 29%. The most frequent level of evidence was evidence-based medicine level 4 (33.3%) or ACC/AHA class IIB: level B (42%). The only intervention with sufficient evidence to recommend "the procedure or treatment should be performed" was avoidance of extreme hemodilution during CPB.

Data supporting use of current neuromonitoring and neuroprotective techniques are limited. The level of evidence is insufficient to support effectiveness of most of these strategies. Well-designed studies with correlation to clinical outcomes and long-term follow-up are needed to develop guidelines for neuromonitoring and neuroprotection during CPB in infants.

(Ann Thorac Surg 2012;94:1365-73)

© 2012 by The Society of Thoracic Surgeons

Survival after surgery for complex congenital heart disease (CHD) has improved dramatically over the past 30 years. With improved early survival there has been an increasing focus on longer term mortality and morbidity after surgery for CHD in neonates and infants. This has led to the troubling recognition that neurodevelopmental dysfunction is the most common, and potentially most disabling, complication of CHD and its treatment. Prevention of perioperative brain injury and subsequent developmental disability has become an increasingly important driver for changes in operative management strategies and research efforts.

There are few subjects debated among pediatric cardiac surgeons that are more contentious than the neuromonitoring and neuroprotective strategies used during cardiac surgery in neonates and infants. Any discussion of the use of deep hypothermic circulatory arrest (DHCA) or near-infrared spectroscopy (NIRS) will elicit multiple, very strongly held opinions. Each potential neuroprotective or monitoring strategy has advocates who promote their viewpoint with great fervor along with individuals who hold the counterpoint with equal intensity. Unfortunately, the body of evidence that supports one strategy or viewpoint over another is often limited and inconclusive. In order to determine the role of any intervention or

Presented at the Forty-eighth Annual Meeting of The Society of Thoracic Surgeons, Fort Lauderdale, FL, Jan 28-Feb 1, 2012.

Address correspondence to Dr Hirsch, University of Michigan Congenital Heart Center, C.S. Mott Children's Hospital, Flr 11, Rm 723, 1540 E Hospital Dr, SPC 4204, Ann Arbor, MI 48109-4204; e-mail: jhirsch@umich.edu.

Appendices for this article are available in the Auxiliary Annals section of the STS website <http://www.sts.org/annals-thoracic-surgery/auxiliary-annals>.

Abbreviations and Acronyms

ACC/AHA	= American College of Cardiology/ American Heart Association
BIS	= bispectral index
CHD	= congenital heart disease
CPB	= cardiopulmonary bypass
DHCA	= deep hypothermic circulatory arrest
EBM	= evidence-based medicine
EEG	= electroencephalogram
HLHS	= hypoplastic left heart syndrome
MRI	= magnetic resonance imaging
NIRS	= near-infrared spectroscopy
RCP	= regional cerebral perfusion
RCT	= randomized control trial
SSEP	= somatosensory evoked potentials
STS	= Society of Thoracic Surgeons
TCD	= transcranial Doppler

monitoring in the clinical setting, it is incumbent upon us to understand that everything we do is associated with potential clinical gains, limitations, and potential harm with intervention, as well as increased cost.

While there is an ever increasing body of medical literature related to these topics, controversy persists and it is difficult for the practitioner and health care systems to assimilate and evaluate the existing evidence in order to make decisions about the care of individual patients. The systematic review is a formal process to identify and evaluate primary studies and other research to make determinations about effectiveness of particular interventions or therapies [1, 2]. Traditional narrative reviews, while providing an overview of an issue, are subjective and suffer from potential selection bias and error. In contradistinction, a systematic review relies on a prospectively defined protocol to identify and appraise the relevant evidence. The aim is to completely identify the relevant publications, minimize selection bias, and develop an objective summary of the data through pre-established evaluation and grading criteria.

With the support of the Workforce for Evidence Based Medicine and the Workforce for Congenital Heart Surgery of the Society of Thoracic Surgeons (STS), we performed a systematic review of the literature describing neuromonitoring and neuroprotection strategies during cardiopulmonary bypass (CPB) in infants of age 1 year or less. The goal of the review was to evaluate the strength of the evidence for the effectiveness of each strategy and the relationship to long-term neurodevelopmental outcomes. Secondary goals were the following: (1) to determine if the level of evidence for any strategy was sufficient to make recommendations concerning guidelines for clinical practice; and (2) to provide results which could be helpful in planning future clinical trials or quality improvement initiatives.

Material and Methods

A systematic review of the literature was performed to evaluate all clinical studies describing techniques or

outcomes of neuromonitoring or neuroprotective strategies implemented during pediatric cardiac surgical procedures involving CPB, including studies with extension of monitoring immediately prior to (within 24 hours) or after operations (within 48 hours). A panel of 9 individuals was convened, including specialists in pediatric cardiac surgery (7), pediatric anesthesiology (1), and pediatric neurology (1). Types of interventions in the literature included the following: drug administration with stated purpose of neuroprotection; any monitoring of cerebral blood flow, oxygenation, or electrical activity; any alteration or scheme of management of CPB with the stated purpose of neuroprotection; and any intervention with the stated purpose of neuromonitoring.

Eligibility Criteria

Inclusion criteria for the literature search were limited to human studies of surgery with CPB, English language, and patients of age 1 year or less. Heart transplantation was excluded due to multiple potential confounding variables. Editorials, case reports, duplicates, and pure technique papers were excluded. Narrative reviews as well as all of the references were evaluated to avoid publication bias by identifying any potential additional manuscripts that were not initially identified in the search strategy. However, these reviews were not included in the final analysis as they do not represent primary data.

Search Strategy

The literature was searched from January 1, 1990 to July 30, 2010 using OVID Medline, OVID Medline In-Process, OVID Old Medline, and all evidence-based medicine (EBM) reviews. The dates for eligibility of publications were chosen to include the introduction of regional cerebral perfusion [3] and to encompass the Boston Circulatory Arrest Trial [4]. The search was performed using the keywords and search strategy outlined in Figure 1. A total of 527 manuscripts were initially identified. Review of all abstracts by 2 investigators (J.C.H., J.W.G.) resulted in identification of 187 potentially eligible manuscripts based on the inclusion and exclusion criteria. All of these manuscripts were reviewed by the panel as well as the references and associated narrative reviews. Additional manuscripts were added and excluded based on this process with a final list of 162 manuscripts for the final analysis (Fig 1).

Data Review and Analysis

A standardized web-based data form was created for grading individual manuscripts. All manuscripts were evaluated for study size, design, outcomes measured (specifically whether a neurologic outcome was the primary endpoint), inclusion of neurodevelopmental testing, and long-term follow-up. Manuscripts were divided into categories based on type of neuromonitoring or neuroprotection (Table 1). A group of 4 reviewers were assigned to each category with 2 of the reviewers always being the first and senior author. Each manuscript was assigned 2 grades by each reviewer using the Oxford Evidence Based Medicine scoring system (available at

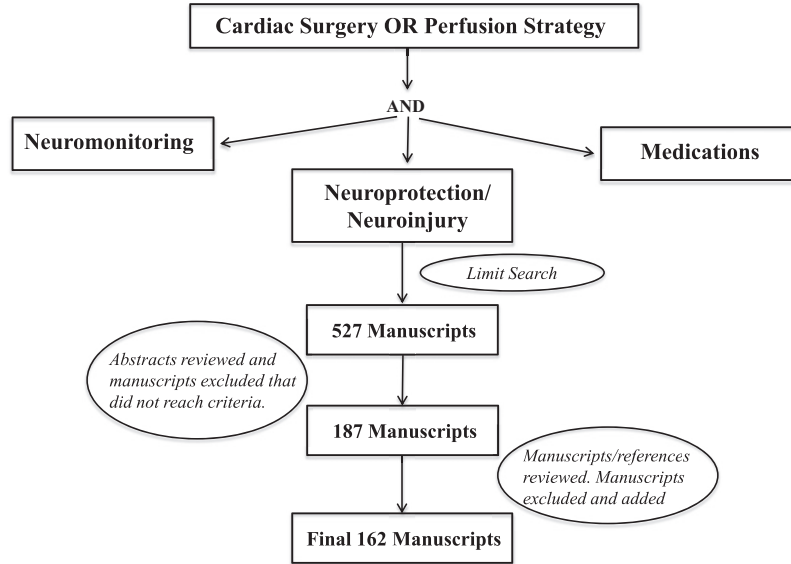


Fig 1. Search strategy and manuscript selection for inclusion. Cardiac surgery/perfusion strategy: Heart surgery, heart disease, cardiac surgery, heart defect, cardiopulmonary bypass, cardiopulmonary bypass, deep hypothermic circulatory arrest, deep hypothermic circulatory arrest, regional cerebral perfusion, antegrade cerebral perfusion, antegrade cerebral perfusion, regional low flow perfusion, regional low-flow perfusion, intermittent cerebral perfusion, intermittent perfusion, selective cerebral perfusion, retrograde cerebral perfusion, cerebral blood flow velocity, cerebral autoregulation, cerebral passive perfusion. Neuromonitoring: Near-infrared spectroscopy, monitoring, transcranial Doppler, electroencephalography, electroencephalogram, bispectral indices, neuromonitoring, oximetry, jugular venous oximetry, jugular bulb oximetry, optical spectroscopy. Neuroprotection/neuroinjury: brain, neurologic, neuroprotection, neurobehavioral, neurocognitive, cerebral protection, stroke, seizure, disability, developmental disability, neurocognitive testing, choreoathetosis, neurologic injury, brain injury, brain ischemia, cerebral embolus, cerebral thrombosis, air embolus, periventricular leukomalacia, sinovenous thrombosis, sagittal sinus thrombosis, white matter injury, grey matter injury, cortical injury, cerebral infarction, arterial ischemic stroke, watershed infarction, intracerebral hemorrhage, intraparenchymal hemorrhage, Bayley scales of infant development. Medications: phenobarbital, erythropoietin, allopurinol, aprotinin, tranexamic acid, steroids, methylprednisolone, hematocrit, dexamethasone, hemodilution, hypothermia, nasopharyngeal temperature. Search limit: humans, clinical trial, meta-analysis, practice guideline, comparative study, controlled clinical trial, guideline, journal article, multicenter study, English, core clinical journals, all infant (birth to 12 months), publication date from January 1, 1990 to July 30, 2010.

<http://www.cebm.net/index.aspx?0=1025>) and the American College of Cardiology/American Heart Association level of evidence grade (available at <http://www.acc.org/qualityandscience/clinical/manual/pdfs/methodology.pdf>).

Table 1. Manuscript Categories (n = 162)

Category	No. of Manuscripts in Group
Biomarkers	17
Bispectral Indices	1
EEG	5
Imaging	9
Monitoring	10
Near-infrared spectroscopy	27
Perfusion	32
Pharmacology	5
Postoperative monitoring	11
Risk factor analysis	33
Transcranial Doppler	8
Technique	3
Temperature/SvO ₂	1

EEG = electroencephalogram; SvO₂ = venous oxygen saturation.

Guidelines were adopted by the panel to provide consistency in grading. In terms of study design, if any comparison group was present, the study was considered to have a “control group” in contrast to studies with a single treatment group only, which was then considered a case-series or cohort study. A study was considered randomized if patients were randomized to different treatment or management groups that were used as the basis for the primary analysis. However, if a study was randomized for a different purpose and the report is a secondary analysis of the randomization, the study was not considered to be randomized. A neurologic outcome considered to be a primary outcome included imaging (magnetic resonance imaging, cranial ultrasound), neurodevelopmental testing, discrete neurologic exam (formal assessment of all patients with a specified exam, excluding simple surveillance for gross clinical abnormalities), and monitoring that detects clinical neurologic abnormalities (for example, electroencephalogram [EEG] assessment for seizures was included, but EEG monitoring for suppression in the operating room was not included). Long-term follow-up needed to be an organized evaluation with 1 of the aforementioned neurologic outcomes and only include assessments after hospital

discharge. In terms of EBM grading, level 4 "case series" included studies with a sample size less than 20; otherwise prospective observational cohorts were considered 2B or 2B- based on quality, and case control studies were considered level 3B. In terms of the American College of Cardiology/American Heart Association (ACC/AHA) level of evidence grade for individual manuscripts, class III (no benefit) was assigned to studies that did not demonstrate any specific benefit or harm (ie, no correlation to clinical outcomes), class IIB was assigned to studies where there was some demonstration of some potential impact on clinical outcomes, level B was used for studies with a large sample size (those that were EBM 2B), and level C was used for studies with a smaller sample size (EBM 4) or poor study design.

After the individual grading, a formal adjudication process was undertaken to reach a consensus score and summary statement for each individual manuscript. A compilation of all the data elements was created independently by the STS technical support staff. The first and senior author performed an initial reconciliation of the overall data by recording consensus where present and averaging submitted scores to reach a preliminary score and concluding statement where consensus was absent. The preliminary report of the reconciled scores was returned to the initial subgroups of 4 reviewers for each category for adjudication. Each reviewer assessed the reconciled scores and reported agreement or disagreement for each manuscript. A list of all manuscripts for which universal agreement of the reconciled scores could not be achieved was compiled. Conference calls with the individuals comprising each review group were undertaken for discussion and final adjudication of the reconciled scores. Manuscripts for which a consensus score could not be determined within a group were submitted to the overall panel of 9 individuals for final adjudication. The final list of all reconciled scores was submitted to the entire panel for final approval. Therefore, all scores represent a consensus of the initial 4 reviewers and approval of the entire panel.

In circumstances where multiple manuscripts addressed closely related subject matter (eg, a management strategy or monitoring modality), that subset of manuscripts were subcategorized into thematic groups to allow for evaluation of the body of literature available on a discrete set of topics in contrast to the initial evaluation that considered the quality of evidence contained within each manuscript independently (Table 2). This subset of manuscripts was submitted to the entire panel for review and scoring by the ACC/AHA level of evidence grade with a separate data retrieval form. Initial reconciliation of scores was performed by the first and senior author after all scoring sheets had been completed. The reconciled scores were distributed to the entire panel for individual approval of summary scores and statements. Summary statements were based on the list of phrases that was developed in 2003 by the ACCF/AHA Task Force on Practice Guidelines (available at <http://www.acc.org/qualityandscience/clinical/manual/pdfs/methodology.pdf>)

for use when writing recommendations. Any areas of discrepancy were submitted for panel adjudication.

Results

Based on the established search criteria, a total of 162 manuscripts were identified that met the eligibility criteria (Fig 1). For initial organization and distribution for scoring, manuscripts were placed into 13 categories based on the primary type of neuromonitoring or neuroprotection being employed (Table 1). The baseline characteristics of the overall body of literature include median sample size of 43 (range 3 to 2,481) with 43% (n = 69) of articles having a neurologic outcome as a primary endpoint and only 29% (n = 47) of articles including some level of formal follow-up after hospital discharge. Prospective observational cohort was the most common study design (53.7%, n = 87), followed by case-control study (21.6%, n = 35), randomized clinical trial (13%, n = 21), retrospective observational cohort (9.9%, n = 16), pre or post study (1.2%, n = 2), and crossover study (0.6%, n = 1). A full list scoring details of individual manuscripts, grouped manuscripts, and category summary scores can be found in Appendices 1-3 (see Appendices in Auxiliary Annals section of the STS website [<http://www.sts.org/auxiliaryannals/Hirsch-2012-94-4-1365-Appendix1.pdf>, <http://www.sts.org/auxiliaryannals/Hirsch-2012-94-4-1365-Appendix2.pdf>, and <http://www.sts.org/auxiliaryannals/Hirsch-2012-94-4-1365-Appendix3.pdf>]).

Adjudication Process

Based on the previously described adjudication process, only 18 of 162 manuscripts were resubmitted for further adjudication. The results were no change in score (6), increase in score by 1 level (8), and decrease in score by 1 level (4).

Summary of Manuscript Categories

A summary of the manuscript category, study design distribution, and scores is presented in Appendix 3 (see Appendix in Auxiliary Annals section of the STS website [<http://www.sts.org/auxiliaryannals/Hirsch-2012-94-4-1365-Appendix3.pdf>]). The distribution of EBM grades and ACC/AHA scores is presented in Figure 2. The majority of papers reviewed were assigned an EBM grade of 4 (33.3%, case series or poor quality cohort or case control studies) or 2B- (24.7%, individual cohort study or low quality randomized control trial [RCT] with wide confidence intervals). The most common ACC/AHA level of evidence score was class IIB: level B (42%, procedure or treatment may be considered, effectiveness is not well established) and class III (no benefit): level C (24.1%, procedure or treatment is not beneficial and is not recommended). Only 4 papers (2.5%) of the manuscripts reviewed were scored as class IIA: level B, which includes procedures or treatments that are reasonable to consider and may be useful. This included a manuscript in the perfusion category that suggests that hypothermia reduces metabolic requirements with no absolute safe level determined and no follow-up data [5], 2 manuscripts in

Table 2. American College of Cardiology/American Heart Association Scores and Conclusions for Grouped Manuscripts

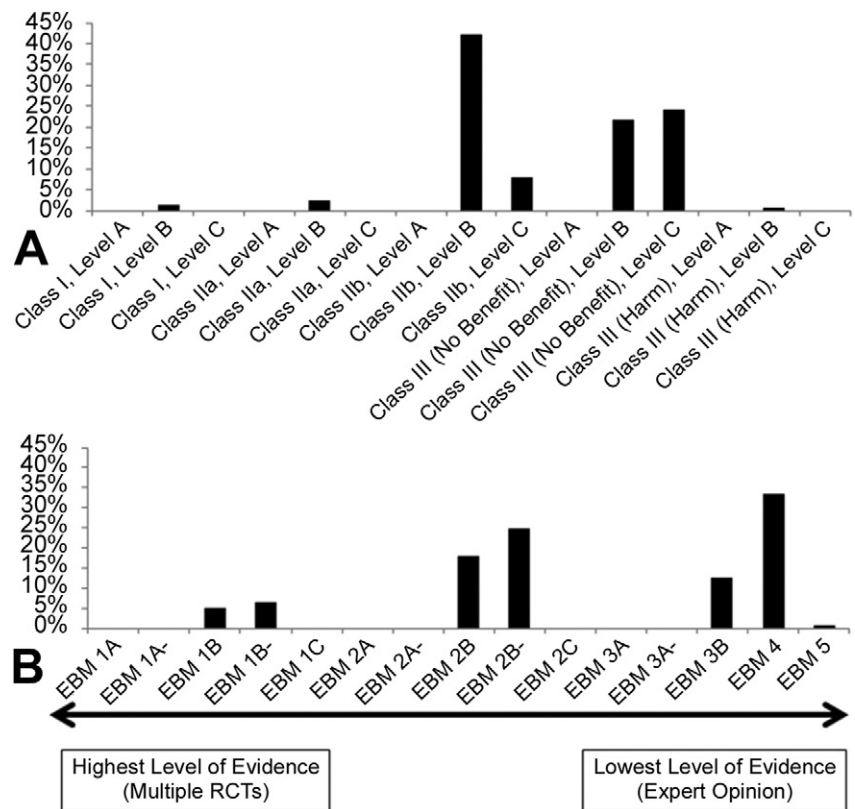
Group	No.	Final Score	Original Distribution of Scores	Final Conclusion
Blood gas management	3	Class IIB, Level B	Class IIB, Level B (7) Class III (no benefit), Level B (2)	No data to demonstrate superiority of alpha vs pH stat blood gas management at long-term neurodevelopmental testing. It is reasonable to use either strategy.
Hematocrit	3	Class IIA, Level A	Class I, Level B (2) Class IIA, Level A (3) Class IIA, Level B (4)	Avoiding extreme hemodilution can be beneficial/effective/useful. An exact lower limit for hematocrit has not been well defined but should probably not go below 24%.
EEG	4	Class III (no benefit), Level C	Class IIB, Level B (4) Class III (no benefit), Level B (2) Class III (no benefit), Level C (3)	No data to show that EEG monitoring is associated with better or worse outcomes. Use of routine EEG monitoring cannot be recommended.
Cooling	5	Class III (no benefit), Level B	Class III (no benefit), Level B (9)	No data to support any difference in outcomes with any specific cooling strategy or duration. No specific cooling strategy can be recommended.
Glycemic control	5	Class III (harm), Level B	Class IIA, Level B (1) Class IIB, Level B (3) Class III (no benefit), Level B (4) Class III (harm), Level B (1)	There is no evidence of benefit for tight glycemic control; however, there may be harm from hypoglycemia. Tight glycemic control is not indicated.
S100 β	12	Class III (no benefit), Level B	Class III (no benefit), Level A (1) Class III (no benefit), Level B (2) Class III (no benefit), Level C (6)	Measuring S100 β has not been demonstrated to identify patients at increased risk of neurologic injury. Measurement of S100 β is not indicated.
Transcranial Doppler (TCD)	15	Class III (no benefit), Level B	Class IIB, Level B (4) Class IIB, Level C (1) Class III (no benefit), Level A (1) Class III (no benefit), Level B (2) Class III (no benefit), Level C (1)	The data concerning TCD is limited in quality. There is no evidence that the use of TCD is associated with improved neurodevelopmental outcomes. TCD monitoring may be considered.
NIRS	35	Class III (no benefit), Level B	Class IIB, Level B (2) Class IIB, Level C (2) Class III (no benefit), Level B (2) Class III (no benefit), Level C (3)	The data concerning NIRS and neurodevelopmental outcomes are limited in quality. There is no consistent evidence that the use of NIRS is associated with improved neurodevelopmental outcomes. NIRS may be considered as a monitoring methodology.
DHCA/LFCPB/RCP	44	Class III (no benefit), Level B	Class IIB, Level B (2) Class III (no benefit), Level B (6) Class III (no benefit), Level C (1)	Among the 3 commonly used perfusion strategies employed for neonatal cardiac surgery using cardiopulmonary bypass and deep hypothermia, none is clearly superior. No specific perfusion strategies can be recommended.

DHCA/LFCPB/RCP = deep hypothermic arrest, low flow cardiopulmonary bypass, regional cerebral perfusion; EEG = electroencephalogram; NIRS = near-infrared spectroscopy.

the pharmacology category with 1 demonstrating increased EEG seizure activity with hypoglycemia [6], another which reported that administration of allopurinol was associated with fewer instances of seizures or coma than was placebo in infants who survived stage 1 surgery for hypoplastic left heart syndrome using deep hypothermic circulatory arrest (DHCA) [7] and a manuscript in the

postoperative category from the Boston Circulatory Arrest Trial showing that body temperature in the postoperative setting did not impact neurodevelopmental outcomes when a strategy targeting normothermia was employed [8]. Two manuscripts (1.3%) were graded class 1: level B evidence, which represents procedures or treatments that are recommended as the benefits clearly

Fig 2. Distribution of (A) American College of Cardiology/American Heart Association classification, and (B) evidence-based medicine grade for individual manuscripts. (RCT = randomized controlled trial.)



outweigh the risks. Both of these studies were in the perfusion category and evaluated the effect of hemodilution on neurodevelopmental outcomes. They contained evidence that extreme hemodilution (no safe threshold identified but may be near 24%) is associated with adverse neurodevelopmental and perioperative outcomes [9, 10].

Summary of Thematic Groups

Nine thematic groups were identified involving common topics within the literature. Within each thematic group, all papers pertained to investigations related to a common neuromonitoring or neuroprotective strategy. One hundred and twenty-six manuscripts were included in this secondary analysis representing 77.8% of the primary manuscript list. Evidence supporting the strategies evaluated in 7 of the 9 thematic groups (representing 95% of the 122 manuscripts in this secondary analysis) was classified as class III (no benefit): level B or C, suggesting these strategies are not recommended and have not been found to be useful. Evidence for strategies of arterial blood gas management, including 3 manuscripts, was classified as class IIB: level B, indicating that differing strategies may be considered; however, effectiveness of any given strategy is not well established. Manuscripts analyzing evidence for hemodilution strategies (n = 3) reached the highest score of class IIA: level A, indicating that it is reasonable to avoid severe hemodilution and that maintenance of higher hematocrit while on bypass can be beneficial, albeit with no clearly defined thresh-

old. The adjudication process for the thematic groups was the most debated, involving multiple iterations. To provide transparency as to the differing opinions, the individual scores are included with the final reconciled score (Table 2). The NIRS thematic group was particularly controversial among the panel, with 4 members recommending a score of IIB (benefits outweigh risk, procedure or treatment may be considered) and 5 members recommending a score of III (no benefit, procedure or treatment is not recommended). Based on majority vote, the final score is III.

In the final analysis, no body of literature on any specific topic reached class I: level A grade for evidence indicative of a strategy or intervention that clearly demonstrates a benefit and is recommended as being effective. Only two independent manuscripts reached class I: level B in terms of avoidance of severe hemodilution; however, when evaluated as a body of literature on the topic by the overall group, the ultimate level of evidence was downgraded to class IIA: level A.

Comment

“Evidence based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research” [11]. The systematic review is a formal process to identify and

evaluate primary studies and other research to make determinations about effectiveness of particular interventions or therapies [1, 2]. The objective of this study was to evaluate the strength of evidence supporting the use of neuroprotective and neuromonitoring strategies, both from the perspective of the individual published studies and the cumulative weight of evidence from the combined literature.

There has been recognition of brain injury and developmental disability in children with CHD even before the development of cardiac surgery, including seizures, stroke, and developmental delay. These deficits were attributed to sequelae of unrepaired CHD. In the 1950s and 1960s, the introduction of CPB made repair of CHD possible; with the prospect of improved survival and better long-term outcomes. However, mortality for small infants and neonates with CHD remained high. In the 1970s, the use of DHCA for the performance of heart surgery in neonates and small infants by Murphy and colleagues [12] and Barratt-Boyes and colleagues [13, 14], among others, was associated with a marked improvement in short-term survival. However, improved survival led to the concerning recognition that neurodevelopmental abnormalities were common among survivors, even after a successful cardiac repair.

Neuropathologic studies of children who died after open heart surgery revealed an alarming prevalence of acquired lesions including embolic stroke and evidence of generalized hypoxic-ischemic injury. These findings led to a concern that brain injury and developmental injury were the result of operative events (emboli, inadequate cerebral perfusion, and cerebral hypoxia-ischemia) and were directly related to the use and management of CPB and DHCA. Because of this concern, it was hypothesized that changes and improvements in bypass circuits and perfusion strategies, including introduction of techniques to monitor cerebral perfusion and function, would reduce the incidence of brain injury and improve neurodevelopmental outcomes.

Many innovative approaches to the conduct of CPB and to monitoring the functional status of the brain and its blood supply have been introduced into clinical practice. Perfusion strategies for infant heart surgery have evolved considerably over the last 2 decades. Many practitioners have adopted practices that allow them to avoid or minimize the use of DHCA, as there was a presumption that neurodevelopmental morbidity was largely related to this technique. In addition, numerous modalities have been introduced for purposes of perioperative and intraoperative monitoring of the brain and its blood supply, metabolism, and electrophysiologic status. While some of these strategies and practices have been widely adopted, and even promoted by their proponents as best practices or "standard of care," the strength of evidence supporting the use of these practices has not been previously assessed in a systematic fashion. Unfortunately, the body of evidence that supports one strategy or viewpoint over another is often limited and not conclusive.

The highest level of evidence supporting a particular

strategy or technique is that replicated by multiple RCTs. Despite a multitude of studies and the dedicated efforts of multiple clinicians and investigators over the past 2 decades, none of the strategies that have been introduced for neuromonitoring or neuroprotection during cardiac surgery in infants are supported by evidence that meets this standard. As described above, most of the evidence used by clinicians to guide their practice is derived from small case series, nonrandomized clinical trials, and observational studies. The quality of evidence available (Table 2) makes it impossible to determine conclusively that a particular strategy is effective or superior in comparison with another. For most of the strategies and techniques evaluated, the ACC/AHA classification was class IIB: level B. This means that limited populations have been evaluated and the evidence is from nonrandomized trials or a single RCT. The recommendation for class IIB: level B according to the guidelines is that evidence suggests that the particular treatment may be considered but effectiveness is uncertain or not well established.

Manuscripts included in this systematic review (Table 2) addressed subjects including arterial blood gas management during CPB (alpha-stat, pH-stat), hemodilution and level of hematocrit during CPB, glycemic control, use of hypothermia and cooling strategies, and various comparisons among and between high-flow CPB, low-flow CPB, DHCA, and regional cerebral perfusion. Among all of these topics, the sole practice for which there is substantial evidence (class IIA-level A) to recommend its use is the avoidance of extreme hemodilution (an exact lower limit for hematocrit has not been well defined but should probably not go below 24%). There was no evidence to support strategies of tight glycemic control; however, there may be harm from hypoglycemia. With respect to blood gas management, cooling strategy, and perfusion strategy (continuous bypass, DHCA, regional cerebral perfusion), there are no data to demonstrate superiority or to recommend any specific practice relative to others.

Another group of manuscripts included in this systematic review (Table 2) addressed the numerous modalities that have been used for purposes of perioperative and intraoperative monitoring of the brain and its blood supply, metabolism, and electrophysiologic status. Among the different technologies described in the manuscripts are NIRS, transcranial Doppler, EEG, bispectral index, and somatosensory evoked potentials. With respect to all of these topics, one retrospective observational study of multimodality neuromonitoring [15] demonstrated an improvement in clinical outcome and was assigned an ACC/AHA score of class IIB: level B. All other neuromonitoring studies were graded the same or lower than this. There is insufficient evidence of an association with improved neurologic outcomes to recommend the use of any single modality as a neuromonitoring strategy during CPB.

As with any analysis, this systematic review has inherent limitations. The analysis ends with manuscripts published before July 30, 2010; therefore, newer papers that have been

published on neuroprotection and neuromonitoring are not included. The analysis was limited to clinical studies, excluding the large body of literature involving animal research on neuromonitoring and neuroprotection. This was intentional to focus on evidence supporting benefit or harm associated with specific strategies currently being employed in the clinical setting. Specific animal studies may demonstrate benefit of any of the aforementioned modalities presented in this analysis, but without application in the clinical setting it is not appropriate to incorporate these findings into recommendations regarding the delivery of care. Two independent scoring systems were used in this analysis as both have inherent limitations. The ACC/AHA level of evidence grade is intended for evaluation of a body of literature rather than individual manuscripts. Therefore its use in the initial analysis is limited for individual manuscripts and is more appropriately applied in the secondary analysis presented in Table 2. The Oxford EBM grades is more suited for individual manuscripts; however, it lacks the associated classification of recommendations that allows application of the analysis to clinical decision making. Finally, the adjudication process involved multiple iterations and attempts to reach final consensus on grades. With respect to the use of NIRS and transcranial Doppler as modalities for neuromonitoring, there is significant diversity of opinion among panel members regarding the adoption of specific phrases suggested by the ACCF/AHA Task Force on Practice Guidelines that correspond to particular levels of evidence.

“What are we to do when the irresistible force of the need to offer clinical advice meets with the immovable object of flawed evidence? All we can do is our best: give the advice, but alert the advisees to the flaws in the evidence on which it is based.” <http://www.cebm.net/index.aspx?0=1025>). Despite the lack of firm evidence, the practicing surgeon must make daily decisions about the management of CPB and the conduct of the operation. Individual clinical expertise must be integrated with the best available clinical evidence to determine practice. This lack of high-level evidence for a particular monitoring modality or clinical approach does not mean that these areas should not be studied further. Rather, it should motivate clinical investigators to perform well-designed prospective trials, with defined longer term neurodevelopmental endpoints, to conclusively answer these questions. There are inherent challenges to these future research endeavors, including the heterogeneous pediatric cardiac surgery population and the significant potential impact of practice pattern variation. In addition, current research efforts are hindered by the lack of sensitive neurologic sequelae within months of the injury rather than years. Prospective randomized trials can be time consuming and costly, making them potentially impractical. However, through the use of existing registries, collaborative efforts among centers, and the growing incorporation of neurodevelopment follow up clinics in many large centers, creation of multiinstitutional projects incorporating retrospective and prospective components may be of significant benefit. Independent of future study design, in order to improve outcomes for our

patients it is essential that we critically evaluate our existing practices as well and develop novel therapies.

The conduct of open heart surgery on neonates and infants has evolved to include a variety of management strategies and neuromonitoring modalities that are used in the hope of minimizing brain injury and adverse neurodevelopmental outcomes. This systematic review demonstrates that data supporting use of these techniques are limited and their effectiveness is uncertain. Except for avoidance of extreme hemodilution there is insufficient evidence to recommend use of any specific neuromonitoring or neuroprotective strategy during CPB. This situation is not the result of failure to investigate the use of these strategies but rather of failure to design studies that have the potential to correlate their use with orderly observations and rigorous data reflecting neurologic outcomes. While the design of this systematic review would not result in downgrading of evidence based upon correlation with short-term outcomes only, it is impossible to overstate the importance of long-term follow-up when attempting to evaluate the efficacy of a treatment strategy or monitoring modality. Future studies on infant brain protection should focus on neurodevelopmental outcomes with long-term follow-up.

This project was supported by The Society of Thoracic Surgeons' Evidence Based Medicine Task Force and Congenital Heart Surgery Workforce.

References

1. Columb MO, Lalkhen A-G. Systematic reviews and meta-analyses. *Current Anaesthesia & Critical Care* 2005;16:391-4.
2. Nightingale A. A guide to systematic literature reviews. *Surgery (Oxford)*. 2009;27:381-4.
3. Pigula FA, Nemoto EM, Griffith BP, Siewers RD. Regional low-flow perfusion provides cerebral circulatory support during neonatal aortic arch reconstruction. *J Thorac Cardiovasc Surg*. 2000;119:331-9.
4. Newburger JW, Jonas RA, Wernovsky G, et al. A comparison of the perioperative neurologic effects of hypothermic circulatory arrest versus low-flow cardiopulmonary bypass in infant heart surgery. *N Engl J Med* 1993;329:1057-64.
5. Kern FH, Ungerleider RM, Reves JG, et al. Effect of altering pump flow rate on cerebral blood flow and metabolism in infants and children. *Ann Thorac Surg* 1993;56:1366-72.
6. de Ferranti S, Gauvreau K, Hickey PR, et al. Intraoperative hyperglycemia during infant cardiac surgery is not associated with adverse neurodevelopmental outcomes at 1, 4, and 8 years. *Anesthesiology* 2004;100:1345-52.
7. Clancy RR, McGaurn SA, Goin JE, et al. Allopurinol neuro-cardiac protection trial in infants undergoing heart surgery using deep hypothermic circulatory arrest. *Pediatrics* 2001;108:61-70.
8. Cottrell SM, Morris KP, Davies P, Bellinger DC, Jonas RA, Newburger JW. Early postoperative body temperature and developmental outcome after open heart surgery in infants. *Ann Thorac Surg* 2004;77:66-71.
9. Jonas RA, Wypij D, Roth SJ, et al. The influence of hemodilution on outcome after hypothermic cardiopulmonary bypass: results of a randomized trial in infants. *J Thorac Cardiovasc Surg* 2003;126:1765-74.
10. Wypij D, Jonas RA, Bellinger DC, et al. The effect of hematocrit during hypothermic cardiopulmonary bypass in infant

- heart surgery: results from the combined Boston hematocrit trials. *J Thorac Cardiovasc Surg* 2008;135:355-60.
11. Sackett DL, Rosenberg WM. The need for evidence-based medicine. *J R Soc Med* 1995;88:620-4.
 12. Murphy JD, Freed MD, Keane JF, Norwood WL, Castaneda AR, Nadas AS. Hemodynamic results after intracardiac repair of tetralogy of Fallot by deep hypothermia and cardiopulmonary bypass. *Circulation* 1980;62(2 Pt 2):I168.
 13. Barratt-Boyes BG. Corrective surgery for congenital heart disease in infants with the use of profound hypothermia and circulatory arrest techniques. *Aust N Z J Surg* 1977;47:737-44.
 14. Barratt-Boyes BG, Neutze JM. Primary repair of tetralogy of Fallot in infancy using profound hypothermia with circulatory arrest and limited cardiopulmonary bypass: a comparison with conventional two stage management. *Ann Surg* 1973;178:406-11.
 15. Austin EH III, Edmonds HL Jr, Auden SM, et al. Benefit of neurophysiologic monitoring for pediatric cardiac surgery. *J Thorac Cardiovasc Surg* 1997;114:707-17.

DISCUSSION

DR JAMES JAGGERS (Aurora, CO): While people are thinking, I know that this has generated a lot of discussion at previous presentations. Just a quick methodologic question. If each one of the nine reviewers scored these, if you will, as far as their level of evidence, was there always agreement amongst the reviewers? And if there was not, how did you manage that?

DR HIRSCH: Surprisingly when we actually looked at the scores, we in general scored them quite closely even though many of us had differing opinions coming in regarding different techniques.

We then averaged the scores for our initial attempt at reaching an adjudication, and then they were sent out to the group for a final consensus.

Any score that people did not agree with, they could submit it back for discussion. We had conference calls where we would review the manuscripts as a group and come to a consensus amongst the initial four reviewers to determine the final score. And if we could not do that amongst the initial four reviewers, it went to the entire panel of nine.

But interestingly, out of 162 manuscripts, there were only 18 manuscripts that were submitted for adjudication, and of those only 50% ended up with a score change.

DR HITENDU DAVE (Zurich, Switzerland): You mentioned that extreme hemodilution was the only factor which came out; so do you have any advice on what is a safe level?

DR HIRSCH: I think with the current evidence in terms of extreme hemodilution, the best that we can say is that it probably should not go below 24%, but an absolute number really has not been clearly defined in the literature. But I think based on the three studies that are out there, a level below 24% probably is clearly associated with adverse outcomes.

DR ERIC N. MENDELOFF (Dallas, TX): Jen, I have a two-part question. One is pretty straightforward. The inclusion criteria you whittled down from 544 to 156. You did not really mention what those criteria were, and just if you could shed a little light on that. That is number one.

And number two, we are left with all these technologies of Doppler monitoring, NIRS [near-infrared spectroscopy], use of regional cerebral perfusion, et cetera. What do you do at University of Michigan when you take a hypoplast to the operating room to do their Norwood?

DR HIRSCH: So in terms of the first question, to get from the initial 527 manuscripts down to the 162, the inclusion criteria included infants less than one year of age. It had to be something that involved monitoring within the operating room. It could be extended to 24 hours before and 48 hours after as well as not being a case report, editorials, or duplicates.

What is remarkable is despite the search filters that you can put into PubMed, how many times you still end up with an animal study, an adult study, or a mixed population. So a lot of studies were excluded in that manner as well as duplicates.

In terms of neuromonitoring, I think what is hard is that there are a lot of things that are out there currently available. And I think our message is not to say do not do anything, but I think currently we are all dependent upon what works best within our own hands that we are using in our own settings. But it is very hard to say that any one thing should become standard of care or is really associated with improved outcomes.

Currently at the University of Michigan, we have NIRS monitors that are placed on our patients by anesthesia, but we do not do anything about them. We do not react to that number at all at this point in time. We do not do any other formal testing in the operating room, and we do deep hypothermic circulatory arrest as our current perfusion strategy.

DR MENDELOFF: And you do not use them in your early postoperative decision making, either?

DR HIRSCH: No. We do not have them in the ICU [intensive care unit]. We have done several clinical trials using NIRS more for mesenteric ischemia and early feeding in single-ventricle patients. We have not used them for monitoring of multisystem organ failure in the postoperative setting or for cerebral monitoring in the postoperative setting.