Primary Outcomes for Resuscitation Science Studies: A Consensus Statement From the American Heart Association

Circulation. published online October 3, 2011;
Circulation is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2011 American Heart Association, Inc. All rights reserved.
Print ISSN: 0009-7322. Online ISSN: 1524-4539

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circ.ahajournals.org/content/early/2011/09/30/CIR.0b013e3182340239

Data Supplement (unedited) at:
http://circ.ahajournals.org/content/suppl/2011/10/18/CIR.0b013e3182340239.DC1.html

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

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

Subscriptions: Information about subscribing to Circulation is online at:
http://circ.ahajournals.org/subscriptions/
Primary Outcomes for Resuscitation Science Studies

A Consensus Statement From the American Heart Association

Lance B. Becker, MD, FAHA, Chair; Tom P. Aufderheide, MD, FAHA; Romergryyko G. Geocadin, MD; Clifton W. Callaway, MD, PhD; Ronald M. Lazar, PhD, FAHA; Michael W. Donnino, MD; Vinay M. Nadkarni, MD, FAHA; Benjamin S. Abella, MD, MPhil; Christophe Adrie, MD; Robert A. Berg, MD, FAHA; Raina M. Merchant, MD, MS; Robert E. O’Connor, MD, MPH, FAHA; David O. Meltzer, MD, PhD; Margo B. Holm, PhD; William T. Longstreth, MD; Henry R. Halperin, MD, MA, FAHA; on behalf of the American Heart Association Emergency Cardiovascular Care Committee and the Council on Cardiopulmonary, Critical Care, Perioperative and Resuscitation

Background and Purpose—The guidelines presented in this consensus statement are intended to serve researchers, clinicians, reviewers, and regulators in the selection of the most appropriate primary outcome for a clinical trial of cardiac arrest therapies. The American Heart Association guidelines for the treatment of cardiac arrest depend on high-quality clinical trials, which depend on the selection of a meaningful primary outcome. Because this selection process has been the subject of much controversy, a consensus conference was convened with national and international experts, the National Institutes of Health, and the US Food and Drug Administration.

Methods—The Research Working Group of the American Heart Association Emergency Cardiovascular Care Committee nominated subject leaders, conference attendees, and writing group members on the basis of their expertise in clinical trials and a diverse perspective of cardiovascular and neurological outcomes (see the online-only Data Supplement). Approval was obtained from the Emergency Cardiovascular Care Committee and the American Heart Association Manuscript Oversight Committee. Preconference position papers were circulated for review; the conference was held; and postconference consensus documents were circulated for review and comments were invited from experts, conference attendees, and writing group members. Discussions focused on (1) when after cardiac arrest the measurement time point should occur; (2) what cardiovascular, neurological, and other physiology should be assessed; and (3) the costs associated with various end points. The final document underwent extensive revision and peer review by the Emergency Cardiovascular Care Committee, the American Heart Association Science Advisory and Coordinating Committee, and oversight committees.

Results—There was consensus that no single primary outcome is appropriate for all studies of cardiac arrest. The best outcome measure is the pairing of a time point and physiological condition that will best answer the question under study. Conference participants were asked to assign an outcome to each of 4 hypothetical cases; however, there was not complete agreement on an ideal outcome measure even after extensive discussion and debate. There was general consensus that it is appropriate for earlier studies to enroll fewer patients and to use earlier time points such as return of spontaneous circulation, simple “alive versus dead,” hospital mortality, or a hemodynamic parameter. For larger studies, a longer time point after arrest should be considered because neurological assessments fluctuate for at least 90 days.

The American Heart Association makes every effort to avoid any actual or potential conflicts of interest that may arise as a result of an outside relationship or a personal, professional, or business interest of a member of the writing panel. Specifically, all members of the writing group are required to complete and submit a Disclosure Questionnaire showing all such relationships that might be perceived as real or potential conflicts of interest.

The opinions expressed in this manuscript are those of the authors and are not necessarily those of the editors or the American Heart Association.

Expert peer review of AHA Scientific Statements is conducted at the AHA National Center. For more on AHA statements and guidelines development, visit http://my.americanheart.org/statements and select the “Policies and Development” link.

Permissions: Multiple copies, modification, alteration, enhancement, and/or distribution of this document are not permitted without the express permission of the American Heart Association. Instructions for obtaining permission are located at www.heart.org/HEARTORG/General/Copyright-Permission-Guidelines_UCM_300404_Article.jsp. A link to the “Copyright Permissions Request Form” appears on the right side of the page.

© 2011 American Heart Association, Inc.

Circulation is available at http://circ.ahajournals.org

DOI: 10.1161/CIR.0b013e3182340239

Downloaded from http://circ.ahajournals.org/ by guest on June 16, 2012
Cardiac arrest, in which all mechanical activity of the heart suddenly stops, strikes ≈295,000 Americans outside the hospital each year. Medical treatment strategies for cardiac arrest are codified in broad guidelines published throughout the world as recommendations for basic, advanced, and pediatric life support.

These international recommendations, which guide the treatment of cardiac arrest, have been developed through a careful evidence evaluation process that formally assesses the effectiveness of various resuscitation therapies reported in the literature. The validity of the evidence evaluation process depends on the use of an appropriate outcome within each individual study to evaluate the superiority or inferiority of many possible treatments. Determination of the most appropriate primary outcome measurement in the care of cardiac arrest is not simple, however, and has been a topic of much consideration, debate, and controversy. This consensus statement provides a state-of-the-art review on the selection of a primary outcome measurement in studies of cardiac arrest, along with consensus recommendations by leaders in the field. Terminology, neurological and nonneurological physiological end points, cost considerations, and a series of general consensus recommendations are discussed.

To consider and weigh the evidence on the issues surrounding primary outcomes, the American Heart Association Emergency Cardiovascular Care Committee commissioned a group of experts to facilitate in-depth discussions between international experts and governmental agencies from multiple fields and policy backgrounds (see the Appendix in the online-only Data Supplement). These experts were charged with defining the important questions related to selecting primary outcomes, preparing preliminary position statements, and convening small-group meetings and a national consensus conference. The goals were (1) to bring together a diverse group of national stakeholders, researchers, regulatory leaders, scientists, and thought leaders to consider the most important issues relating to assignment of a primary outcome in cardiac arrest studies; (2) to identify areas of consensus and areas without consensus in the assignment of a primary outcome; (3) to better understand tools that researchers may use for measuring outcomes; and (4) to provide guidance and consensus recommendations to the research community on optimal primary outcome measurements in specific settings.

Why Is Selection of a Primary Outcome Measurement a Problem?

Many initial studies of cardiac arrest designated a relatively early immediate physiological process variable such as presence versus absence of a pulse during attempted resuscitation, elimination of ventricular fibrillation (VF), or temporary reestablishment of blood pressure for the primary outcome. Other cardiac arrest studies used relatively short-term survival end points such as survival to hospital admission as a primary outcome. Some studies have reported longer-term outcomes such as 30-day or 1-year survival with assessment of neurological function as the primary outcome. Some experts have suggested long-term outcome with neurological assessment for all studies of cardiac arrest. However, there are significant associated costs, including incomplete follow-up requiring intensive efforts to locate missing data, follow-up bias, and lack of feasibility if extensive testing or face-to-face interviews are required. Further complicating the decision, the definition of a good neurological outcome remains uncertain, as does whether the additional cost of obtaining long-term neurological data is worth the expense, particularly if the primary goal of the study is to identify superior and inferior therapies used in a medical transport or emergency setting. In studies of treatments delivered in the field or on an ambulance to help restart an arrested heart, the positive impact of such treatment on long-term effectiveness may be dwarfed by the confounding effects of subsequent therapies delivered in the emergency department, hospital, or intensive care unit. Recent data suggest that survival rates for cardiac arrest vary by nearly 500%, depending on location. Because site differences may be as large as or greater than the effects of interventions, studies should account for possible wide variation in survival rates and care between sites.

Funding for studies of cardiac arrest is limited, and there are significant obstacles to performing such studies because of lack of consent. It is easy to understand why the resuscitation research community has asked for additional guidance in the selection of outcome measurements.

Principles Underlying Selection of an Outcome: When, What, and How Much?

Selection of primary and secondary outcomes reported by clinical resuscitation studies in the literature varies widely. The consensus of the experts was that 3 major factors must be considered when defining a primary outcome measurement:

1. To bring together a diverse group of national stakeholders, researchers, regulatory leaders, scientists, and thought leaders to consider the most important issues relating to assignment of a primary outcome in cardiac arrest studies.
2. To identify areas of consensus and areas without consensus in the assignment of a primary outcome.
3. To better understand tools that researchers may use for measuring outcomes.
4. To provide guidance and consensus recommendations to the research community on optimal primary outcome measurements in specific settings.
for a given study: the time point after cardiac arrest to determine outcome, the physiology assessed at the time point, and the costs associated with use of the outcome. These are the when, what, and how much questions that must be answered when selecting a primary outcome measurement.

**When: The Time Point After Cardiac Arrest**
A primary outcome must first be determined at a specific time point during or after a cardiac arrest. Early time points have been used in many studies, and a wide range of intermediate and long-term outcome time points may be appropriate for a given study.

**What: The Physiological Assessment or Finding Measured at the Time Point**
Survival versus death has been the primary physiological assessment in many studies of cardiac arrest because it is simple, unambiguous, and straightforward to determine. However, survival alone does not tell us if the survivor is in good health, is neurologically impaired, or is in a persistent vegetative state. Assessment of cardiovascular, neurological, and psychological function provides much more information about the survivor. For a given therapy to be considered successful, patient-centered research suggests that patients do not just want to be alive; they also want to have good restoration of neurological function. An array of testing, imaging, and measurement is available to researchers and may provide important insight into a patient’s physiology after cardiac arrest. Multiple assessment tools can be used in resuscitation studies; most have important advantages and limitations. Although several of these tools appear to be useful, none is considered the gold standard.

**How Much: Cost of Use of an Outcome Measurement**
Cost may not seem an important consideration during the selection of a primary end point, but it is a critical component in the responsible conduct of research. Resources for research are limited. Excessive costs reduce the ability of investigators to conduct additional lifesaving research. If the cost of a study exceeds the funds allocated for it, researchers risk losing the ability to find answers to other important research questions. Conversely, when too little is spent to obtain sufficient data, researchers risk being unable to answer the primary question, which is a waste of resources. As a practical note, because outcome measurement has cost implications for the conduct of a study, a realistic budget can be created only if the outcome parameter is specified and costs are estimated on that basis. An unexpected change in primary outcome measurement can prove disastrous for a funded study with a fixed budget.

**Definitions and Concepts**

**Consensus on Definitions**
An ad hoc group of experts met to discuss the most useful definitions in consideration of outcome measurements, summarized below.

**Primary Outcome**
The primary outcome is the outcome chosen by investigators to represent the overall and most significant indicator of the success or failure of the intervention. It is used to power the study for significance to accept or reject the null hypothesis or to evaluate therapeutic equipoise. Primary outcome determines the number of patients required for enrollment in the study and is a primary parameter followed by the data safety monitoring board for early curtailment of a study as a result of clear superiority, harm, or futility. Primary outcome includes a time point or follow-up period for repeated measures and physiological state or parameter at the time point.

**Time Points for Primary Outcome**
The words used to define time points have different meanings in different studies. The following terms are useful:

- “Intra-arrest” or “immediate” implies a time point occurring during or within minutes of achieving return of spontaneous circulation (ROSC). For example, defibrillation with termination of VF that lasts ≥5 seconds has been used in many studies of defibrillation efficacy and is viewed as an immediate time point.
- “Short term” describes a time point occurring more than a few minutes after ROSC and before hospital discharge.
- “Intermediate” refers to a time point reported during the hospital stay or at discharge.
- “Long term” describes a time point that follows hospital discharge such as ≥30 days after cardiac arrest.

**Physiological Condition Measured at the Time Point of the Primary End Point**
These conditions include simple survival versus death, neurological status, cardiovascular status, and other indexes such as renal function or biomarkers. Molecular and genomic alterations are other possible conditions of interest, although they are usually more appropriate for secondary end points or for generating a new hypothesis. More in-depth considerations are described below.

**What Are the Costs of the Primary Study Outcome?**
The selection of the primary outcome is an important factor in the overall cost and budget of the study. Researchers need to be familiar with hidden costs such as losing patients to long-term follow-up assessment and the cost to society of not performing the research. Cost-effectiveness analysis is a tool that can be applied to resuscitation therapies and may involve consideration of long-term costs, including care for survivors with neurological injuries, marginal costs, hospital costs, and future therapies. Some important cost considerations are discussed below.

**Considerations in Outcomes Related to Cardiovascular and Nonneurological Function**
Circulatory arrest results in injury that may affect all organ systems. In general, cardiovascular performance determines immediate and short-term survival, whereas neurological performance affects longer-term survival. Other organ systems contribute to patient morbidity, and previous investigations in critically ill populations suggest a stepwise increase in mortality for each additional organ that fails. Hemodynamic instability can vary independently from and contribute to neurological injury. Even if the patient survives, dysfunction of nonneurological organ systems such as the heart, liver, or kidneys affects quality of life and use of healthcare resources.
Whenever possible, outcome measures must be appropriate for the specific disease processes under study and relevant to the patient. For example, a critically ill patient may receive a treatment that dramatically reduces the risk of organ failure. If the patient dies as a result of the devastating neurological injury, the potentially beneficial intervention targeted to organ failure may remain undetected if the only end point for the study is death. Therefore, the selection of appropriate outcome measurements depends on the stage of development of the drug or device and its appropriateness for the specific disease process being studied.

**Cause and Stratification of Patients With Cardiac Arrest**

The underlying cause of circulatory arrest has a profound impact on definitive therapy, likelihood of recurrence, and overall prognosis and thus may be important in the design of intervention trials.8–12

**Importance of Standardized Definitions and Reliable Incidence Data**

The rate of particular outcomes relies on accurate estimation of the incidence of cardiac arrest. Significant challenges in the epidemiology of sudden cardiac arrest remain, leading to wide variations in reported incidence and outcomes.6,13 This variation is a result of the lack of consistent standardized definitions of sudden cardiac arrest and the use of surrogate data such as deaths caused by coronary heart disease.1

**Measuring Cardiac Dysfunction After Resuscitation**

Acute cardiovascular failure defines cardiopulmonary arrest. Both short-term dysfunction and long-term damage to the heart may result from cardiac arrest. Cardiovascular outcome measures can be divided into those that assess the immediate reversal of cardiac arrest, short-term characteristics of cardiovascular function, and long-term cardiovascular function or survival period free of cardiovascular symptoms and events.

**Immediate Reversal of Cardiac Arrest**

The immediate reversal of cardiopulmonary arrest is the most important prerequisite for long-term survival and requires 2 separate components: restoration of organized electric activity and restoration of mechanical cardiac activity. Therefore, electrophysiological outcomes such as restoration of electric activity may be considered a primary outcome for studies of purely electric interventions (such as defibrillation). For example, termination of VF might be a reasonable outcome for establishing the efficacy of defibrillation. Both asystole and restoration of organized rhythm represent termination of VF. However, restoration of organized electric activity is more likely to lead to better survival, making this outcome superior to asystole.14 The electric activity that follows termination of VF varies over the seconds and minutes after a defibrillation shock, and any outcome must be specified in terms of assessment time after a shock.15

Restoring effective mechanical activity (spontaneous circulation) is a more desirable outcome than restoration of electric activity alone. However, restoration of mechanical activity is also influenced by duration of cardiac arrest, factors related to the intrinsic health of the heart, and the type and quality of immediate care provided before and after a study intervention. These factors must be controlled or adjusted for in assessments of outcomes. The outcome term ROSC has been defined inconsistently in many prior studies. For example, the exact time duration for return of pulses or blood pressure was not specified in the initial Utstein definitions,16 whereas the 2004 Utstein definitions specified 20 minutes of pulses as a criterion for the new term sustained ROSC.17 Although experts noted that changes in immediate care could be recommended on the basis of changes in successful restoration of pulses, many interventions that promote restoration of cardiac activity often fail to improve longer-term outcomes.18,19

The following immediate cardiovascular function outcomes are recommended in increasing order of importance:

- Termination of VF/ventricular tachycardia for a specified duration
- Restoration of organized electric activity for a specified duration
- Restoration of circulation for a specified duration
- Sustained ROSC for a specified duration

**Short-Term Myocardial Function**

- Shock reversal: time from initiation of vasopressor to vasopressor discontinuation
- Mixed or central venous oxygenation
- Ejection fraction
- Cardiac output
- Filling pressures (central venous pressure, pulmonary capillary wedge pressure)
- Mixed or central venous oxygenation
- Lactate clearance

**Lactate clearance**

Consensus guidelines describe strategies for identification and diagnosis of heart failure.25 Symptoms include decreased exercise tolerance or fluid retention as evaluated by 2-dimensional echocardiography with Doppler flow, nuclear ventriculography, and circulating levels of...
Creatinine clearance, changes in glomerular filtration rate, and testing, number of ventilator-free days, and oxygen independence. Long-term myocardial recovery may be particularly sensitive to post-ROSC interventions such as revascularization, pharmacotherapy, and implantable cardioverter-defibrillators. For example, survival until a second episode of VF and sudden death must be defined differently for patients with or without implantable cardioverter-defibrillators. Likewise, functional capacity may have little relation to immediate resuscitation interventions for the patient who undergoes a heart transplantation.

The following long-term cardiovascular function outcomes are recommended:

- New York Heart Association heart failure classification
- Echocardiographic or nuclear assessment of ventricular function
- Circulating levels of brain natriuretic peptide

### Assessing Pulmonary, Renal, and Gastrointestinal Injury or Dysfunction

Other organ systems may be dysfunctional after cardiac arrest, and various scales can be used to quantify this damage, although not all studies should measure all organ systems. Rather, these other measures should be considered when there is specific interest in a specific organ. For example, researchers may wish to examine lung injury after arrest, which can result in an increased alveolar-arterial oxygen gradient (A-a gradient) and is reflected by a decreased ratio of arterial oxygen to fraction of inspired oxygen (PaO2/FiO2). A PaO2/FiO2 ratio of <300 is the criterion for acute lung injury; a ratio of <200 is the criterion for acute respiratory distress syndrome. Of note, other causes of hypoxia such as pulmonary edema or aspiration may be present (either causal or secondary to arrest). Longer-term lung recovery may be measured by pulmonary function testing, number of ventilator-free days, and oxygen independence.

Injury to the kidneys can be measured by the use of creatinine clearance, changes in glomerular filtration rate, and creatinine and urine output. Serum creatinine and urine output are easily measured serially. Ultimately, the need for renal replacement therapy or renal transplantation is an important long-term outcome.

Hepatic injury also occurs after cardiac arrest, and elevated ammonia levels are a very poor prognostic finding. However, data are sparse about the usefulness of other gastrointestinal measurements after cardiac arrest.

Therefore, the following measures are recommended to track potential injury to other organs:

The pulmonary function outcomes are

- PaO2/FiO2 ratio
- Number of ventilator-free days

The renal function outcomes are

- Need for renal replacement therapy
- Urine output
- Creatinine clearance changes

### Measures of Microcirculatory Failure After Cardiopulmonary Resuscitation

Previous investigation in sepsis and injury indicates that microcirculation disturbances lead to organ dysfunction. Even when systemic blood pressure, perfusion pressures, pressor requirements, and systemic oxygen delivery are normalized, cells may still have poor oxygen uptake. Microvascular collapse, intra-vascular thrombogenesis, and endothelial damage occur after ischemia/reperfusion. Direct observation of microvasculature with technologies such as orthogonal polarographic spectral videography may be able to detect dynamic microvascular derangements. Delayed clearance of lactate and reduced mixed (or central) venous oxygen saturation (SvO2) also may reflect microcirculatory dysfunction. At this time, however, all direct and indirect measurements of microcirculatory function can be recommended only as investigational.

### Severity of Illness and Multiple Organ Dysfunction

Severity scores provide a structured method of routinely integrating data about multiple organ systems and are generally calibrated to predict survival (but not neurological outcome). Modern severity scale scores predict survival after many critical illnesses with an area under the receiver-operating characteristics curve of approximately ≥0.80. These scores may be useful for measuring the severity of cardiac arrest illness. Only the most current version of severity scores should be used for research purposes. Clinical practice evolves, and scales lose calibration over time. Scores also lose calibration when applied to different geographic populations or other time points in disease (eg, in the emergency department rather than the intensive care unit).

Because the first 24 hours after restoration of pulses is a dynamic time, severity scores that require 24 hours of information for an initial score (eg, Acute Physiology and Chronic Health Evaluation) may be less useful in cardiac arrest studies because of that dynamic instability. Ideally, a score uses only data from a narrow time window.

The Mortality Prediction Model I through III, Simplified Acute Physiology Score 3, and Sequential Organ Failure Assessment use information available on admission and could be used in cardiac arrest studies. The Mortality Prediction Model III includes antecedent cardiopulmonary resuscitation (CPR) as a variable and thus has been calibrated to the post–cardiac arrest population. The Sequential Organ Failure Assessment scale was designed to monitor patients with sepsis but has been used as a repeat measurement for monitoring patients over time. In 1 study, a score was derived specifically for the cardiac arrest population (derived from 120 patients in 1 center and validated in 210 patients in 4 centers in France) to predict survival with good neurological outcome (out-of-hospital cardiac arrest score). These scores are generally useful to describe severity of illness in a cohort and are not designed for decision making for specific patients.

### Biological Markers of Severity of Illness and Disease Progression

Biological markers in blood would be useful for following disease progression and response to therapy. Ideal biomarkers should have high sensitivity and specificity; be readily
Health Condition  
(Disorder or Disease)

Body Structures
Body Functions
Impairments

Activities
Activity Limitations

Participation
Participation Restrictions

Environmental Factors
Personal Factors

reproduced across different laboratories; be independent for age, race, or sex; and add information above the usual clinical or electrophysiological data already available. Many studies of biomarkers have been conducted in small cohorts with variable cutoffs, however, making it difficult to obtain a low false-positive rate (with low confidence intervals) for predicting poor outcome. In cardiac arrest care, prediction of a poor outcome with a high degree of certainty in this clinical context may lead to withdrawal of care and death. Therefore, near-zero false-positive rates are required in prognostication biomarkers.

Neuron-specific enolase (NSE) has been used to evaluate neuronal injury. NSE levels >33 μg/L at any time from day 1 to 3 appear to predict poor outcome. This NSE level identified patients not recognized by abnormal somatosensory evoked potentials (bilateral absence of N20). Finally, NSE levels are decreased in subjects treated with hypothermia over time (to 48 hours), which suggests that this biomarker may be helpful in monitoring the efficacy of a treatment. However, a study examining the variable results obtained from 7 NSE kits found that false-positive rates ranged from 0% to 54%, which would argue against the use of absolute levels of NSE for predicting prognosis after cardiac arrest.

Nonneurological markers associated with unfavorable outcomes include lactate (or rate of lactate clearance), blood creatinine, liver enzymes, serum cytokines, soluble selectins, coagulation abnormalities with a constant disseminated intravascular coagulation, and hyperglycemia. The inflammatory profile, along with the presence of endotoxinemia and dysregulated leukocyte production of cytokines (so-called endotoxin tolerance), resembles the profile of patients with severe sepsis or traumatic brain injury. Markers such as these may be nonspecific, particularly when vaspressors are required or when death results from multiple organ failure. Therefore, these markers may be useful for monitoring the efficacy of therapeutic strategies to the same extent that they are useful in any other group of critically ill patients.

Cardiac enzyme markers (eg, troponins, creatine kinase-MB) are frequently elevated after cardiac arrest, but these changes may not be specific because of the multiple causes of heart injury, including myocardial infarction, rapid ventricular rhythm (VF or ventricular tachycardia), chest compres-

Figure 1. Adaptation of the World Health Organization international classification of functioning, disability, and health. Adapted from Classification of Functioning, Disability and Health (ICF). Copyright © 2001, the World Health Organization.

sion, administration of epinephrine, and defibrillation. No current post–cardiac arrest biomarker measure has adequate specificity or sensitivity to follow recovery or to predict outcome after cardiac arrest. Candidate biomarkers can be recommended only as an area for further research.

Considerations for Assessing Neurological Functional Outcome After Cardiac Arrest

Until recently, successful resuscitation from cardiac arrest was defined by mortality, physiological measures, or surrogate diagnostic tests. Such measures became outcomes in clinical trials. As resuscitation interventions have become more successful, however, there is an increasing need to reconsider the patient-centered outcomes associated with survival.

The World Health Organization’s International Classification of Functioning, Disability and Health (ICF) (Figure 1), linked to the International Classification of Diseases, is 1 rubric that categorizes domains that have an impact on health conditions. The international classification of functioning categorizes body structures, functions, and impairments. It also includes everyday activities and limitations, societal participation and restrictions, and the impact of human and physical environment and personal factors. The international classification of functioning model can serve to organize the types of data that need to be collected in resuscitation from cardiac arrest. In survivors of cardiac arrest, for example, body structures and functions can include neuroanatomical structures and electrophysiological measurements that are the focus of periresuscitation outcomes, whereas activities, participation, environmental factors, and personal factors are relevant to patient status after discharge.

Neurological Outcome After CPR

Many clinical trials on neuroprotection in the era of modern resuscitation failed to show significant outcome benefit. Because newer interventions targeted the preservation of brain function, the Cerebral Performance Category (CPC) was developed as the central nervous system outcome measure. In the context of Utstein-style reporting, the CPC scale became the most commonly used standard for postresuscitation outcome measurement. The CPC was adapted from the Glasgow Outcome Scale for traumatic head injury. The strengths of the CPC are extensive use, simplicity, and separation into good and
poor outcomes. Despite the historical importance, intuitive appeal, and widespread use of the CPC, however, no validity or reliability studies have ever been conducted for any postresuscitation time points for which it has been used.

The CPC purports to assess domains of functioning after CPR, with scores ranging from 1 (good cerebral performance/normal life) to 5 (brain death). Each CPC score, however, includes multiple domains of function. For example, a CPC score of 2 represents 3 domains of function: impairment (eg, the presence of hemiplegia, seizures, dysarthria, or permanent memory or mental changes), level of activities performed (eg, ability to dress independently, to travel by public transportation, to prepare food), and level of participation (sufficient cerebral function to work part-time in a sheltered environment). Unfortunately, it has not been established that the CPC has sufficient sensitivity to assess all domains. Moreover, rater bias as to the domain that is the primary focus may differ if the CPC score is derived in the emergency department (consciousness), 1 month later in a rehabilitation facility (cognition), or 1 year after discharge at home (activities of daily living). The CPC has been used as an outcome measure in multiple follow-up studies of cardiac arrest, targeting those who have survived ≥6 months after cardiac arrest and often dichotomizing scores into good (CPC score 1 or 2) and poor (CPC score 3–5) neurological outcomes. For example, studies have examined the relationship between CPC score 1 or 2 and poor (CPC score 3–5) neurological outcomes at discharge.75 A direct comparison of the CPC and quality-of-life outcomes at good recovery (CPC score 1) at the time of discharge.69 Cardiopulmonary Resuscitation showed that 86% of patients had continued to show that outcomes remain unclear. These studies, for example, those by Raina et al70 and Hsu et al,64 which used discharge CPC scores to project outcomes at 1 month after cardiac arrest, raise concerns about overestimation of positive longer-term outcomes.65,70 Likewise, Tianen et al66 reported that among 93% of participants who were classified as having a good outcome (CPC score 1 or 2), 5 months after cardiac arrest, 34% had moderate or severe deficits in standardized neuropsychological measures. A systematic review of 28 studies examining cognitive impairment ≥3 months after out-of-hospital cardiac arrest found impairment (mainly memory, attention, and executive function) in 6% to 100% of patients.71 In the same report,71 the 3 largest prospective studies show high rates of impairment, ranging from 42%72 to 50%73 to 60%74 at 3 months. Another study showed that up to 74% of survivors have low societal participation at 3 years.75 A direct comparison of the CPC and the Health Utilities Index showed that the CPC is an important tool indicating broad functional outcome categories that are useful for a number of key clinical and research applications but should not be considered a substitute for the Health Utilities Index.45 Collectively, these studies suggest that although the patient may survive, some neurological dysfunction is perhaps more common than realized.

Neurological Considerations to Improve Assessment of Recovery After CPR

The ultimate goal of CPR is the preservation of the prearrest level of function. Given that neurological status is a major determinant of overall functional outcome, either the primary or secondary measures of a good study should include some measure of neurological recovery. It has been determined that not all unresponsive patients during and immediately after successful resuscitation have comparable neurological injury or clinical outcome. At present, predictive factors for poor functional outcome are apparent 3 days after cardiac arrest.76 As better prognostic factors are identified, stratification before randomization or control for imbalances after randomization in future clinical studies will be possible (Table 1). There is also a need to develop objective monitors that track the real-time and subsequent effects of interventions on neurological recovery as measures of efficacy and safety in clinical trials. Opportunities to develop better monitoring may be derived from advances in clinical neuroelectrophysiology, neuroimaging, and biomarkers. There is growing consensus that the absence of such measurements may have contributed to the failure of neuroprotection trials in acute ischemic stroke,77,78 and a similar process may be contributing to the lack of success in many trials of neuroprotection after cardiac arrest.

Other factors may further complicate neurological outcome assessment. The underlying disease or condition that brought about cardiac arrest in itself may worsen central nervous system function or even result in death. In addition, decisions about withdrawal of life support in patients before potential recovery pose complex analytic questions. Future studies are needed to integrate such factors into the quantitative models for natural history studies and for clinical trials with intent-to-treat designs.

Functional Outcome: When and What to Assess

Recovery of neurological function may begin soon after ROSC and continue months after the injury. In the acute admission period, neurological factors have been used to predict poor functional outcome.76 These prognostic variables are components of the neurological examination that defines level of responsiveness, cranial nerve function, and motor response. Diagnostic tests such as electroencephalograms, evoked potentials, neuroimaging, and serum and cerebrospinal biomarkers have also been used as outcome measures (Table 1). The main objective in neurological assessment during the acute postresuscitation period is not only to determine ongoing injury but also to establish the patient’s recovery from unresponsiveness with the ability to follow commands or speak comprehensibly. This recovery is currently tracked with a standard neurological examination or the CPC, which is an adaptation of the Glasgow Coma Scale.79–81 Recovery from coma, among the most important steps in neurological recovery, remains poorly understood.82 For those who remain unresponsive, prediction of poor outcome may be established by clinical examination and is enhanced by diagnostic measures.71 In those who regain wakefulness, the evaluation of cognitive function is added to the standard neurological assessments. Tests such as the Glasgow Coma Scale,79–81 the Mini-Mental State Examination,82,83 and the National Institutes of Health Stroke Scale are added to the bedside neurological assessment.84–86 Other clinical manifestations that have been used as outcomes during this period are neurological complications such as prolonged seizures87 and myoclonic epilepsy.76
Table 1. Measures Related to Neurological Prognosis and Outcome After Cardiac Arrest

<table>
<thead>
<tr>
<th>Measures</th>
<th>Awake, With Comprehensible Speech, or Following Commands</th>
<th>Not Awake, Without Comprehensible Speech, or Not Following Commands</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Periresuscitation</td>
<td>Early Hospital Course</td>
</tr>
<tr>
<td>Neurological examination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spontaneous breathing activity</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Pupillary light response</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Other brainstem reflexes</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Motor response (from GCS)</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Verbal response (from GCS)</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>NIHSS</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Comprehensive neurological examination</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Preexisting conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demographics</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Comorbidities</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Genetic</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Biomarkers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serum (NSE, S100B)</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Cerebrospinal fluid (CK isoenzymes)</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Electrophysiological tests</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEG</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Evoked potentials</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Imaging</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>MRI</td>
<td>S</td>
<td>R</td>
</tr>
<tr>
<td>PET</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interventions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatments</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>No code orders</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Withdrawal of life-sustaining</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>treatments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Complications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recurrent cardiac arrest</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Cardiogenic shock</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Seizures</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Dichotomous outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Survival and time of death</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Awakening and time of awakening</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Independence (“In the last 2 weeks, did you require help from another person for everyday activities?” From simple questions for stroke)</td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Full recovery (“Do you feel that you have made a complete mental recovery from the coma that followed your cardiac arrest?” Modified from simple questions for stroke)</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

(Continued)
Neurological assessment, especially of cognitive function, is susceptible to many physiological and pharmacological perturbations during the acute period. It is therefore unlikely that any single measurement will be sufficient to establish neurological status. However, recent advances may provide opportunities to establish better predictive models, to improve care, and to promote recovery.

Optimal times for follow-up after cardiac arrest have yet to be established. In Table 1, early hospitalization signifies the period when active intervention may alter potential outcome (i.e., the first 72 hours to 7 days); later hospitalization refers to the period in which neurological injury has become more established (i.e., >7 days). The hospital discharge date is well defined, but the duration from time of cardiac arrest varies widely. At the time of discharge, the patient’s neurological status is assessed to determine further care and discharge disposition. The discharge date is defined as the day the patient is declared neurologically stable enough to leave the hospital.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Periresuscitation</th>
<th>Early Hospital Course</th>
<th>Later Hospital Course</th>
<th>Discharge</th>
<th>Postdischarge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awake, With Comprehensible Speech, or Following Commands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not Awake, Without Comprehensible Speech, or Not Following Commands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global ordinal outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOS</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>CPC</td>
<td></td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>GOS-E</td>
<td></td>
<td>R</td>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified Rankin Scale</td>
<td></td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Functional activities measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modified BI (basic ADL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FAQ (instrumental ADL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitive measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td></td>
<td>R</td>
<td>R</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>Memory (RAVLT)</td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Executive functioning (Trail Making Test, parts A and B)</td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Attention (DSST)</td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Depression measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Center for Epidemiologic Studies Depression Scale</td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Hamilton Rating Scale for Depression</td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Zung Self-Rating Depression Scale</td>
<td></td>
<td></td>
<td></td>
<td>R</td>
<td>R</td>
</tr>
<tr>
<td>Quality-of-life measures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health Utilities Index Mark 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical Outcomes Study 36-Item Short-Form Health Survey</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RAND-12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nottingham Health Profile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NIH Toolbox for Assessment of Neurological and Behavioral Function</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Satisfaction measure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consumer Assessment of Healthcare Providers and Systems</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

R indicates research; S, standard care; GCS, Glasgow Coma Scale; NIHSS, National Institutes of Health Stroke Scale; NSE, neuron-specific enolase; CK, creatine kinase; EEG, electroencephalogram; CT, computed tomography; MRI, magnetic resonance imaging; PET, positron emission tomography; GOS, Glasgow Outcome Scale; CPC, Cerebral Performance Category; GOS-E, Glasgow Outcome Scale–Extended; BI, Barthel Index; ADL, activities of daily living; FAQ, Functional Activities Questionnaire; MMSE, Mini-Mental State Examination; RAVLT, Rey Auditory Verbal Learning Test; DSST, Digit Symbol Substitution Test; SF-12, 12-Item Short Form Survey Instrument; and NIH, National Institutes of Health.

The population (awake or not awake), interval from cardiac arrest, and purpose (standard care or research) in which the measures are most useful are indicated. Measures that are accepted as standard of care at a given time point may also be used as a research measure. The use of certain measures for research is provided on the basis of perceived relevance of the measure at that time or experience from previous or ongoing investigations.
discharge, most patients are awake, defined as being able to follow commands or to produce comprehensible speech, but a few patients have awakened up to 3 months after cardiac arrest. Nevertheless, those who awaken during hospitalization have typically not achieved their highest recovery by the time of discharge. At present, discharge status is not a good predictor of a patient’s future level of function.

With so much patient- and treatment-related variability in the time between cardiac arrest and hospital discharge, standardized assessment of neurological outcomes measurement points has yet to be established. Indeed, few natural history studies have characterized neurological function serially after cardiac arrest. In 1 study, the Mini-Mental State Examination score improved after cardiac arrest but was almost identical at 3 months and 1 year after arrest. For patients in a persistent vegetative state 1 month after a nontraumatic injury, recovery of consciousness is extremely unlikely after 3 months. As a point of comparison in focal brain ischemia, the follow-up of primary outcome in some key neuroprotection trials is often at 3 months, and a consensus statement supports that duration. For understanding neurological processes after cardiac arrest, a 3-month postdischarge period would balance the opportunity for recovery and minimize loss of patients to follow-up and competing end points, including death.

Types of Functional Outcome Measures

Global Summative and Nonsummative Measures

The Glasgow Coma Scale, Glasgow Outcome Scale, Glasgow Outcome Scale—Extended, CPC, and modified Rankin Scale have 1-digit ordinal scales that reflect increasing or decreasing levels of overall functional status. The Glasgow Coma Scale, which has established validity and reliability, addresses eye opening, motor responses, and verbal responses, with the lowest score reflecting the poorest response. Among the nonsummative measures, the Glasgow Outcome Scale addresses level of consciousness, recovery, and disability on a 5-point scale ranging from 1 (death) to 5 (good recovery) and has adequate reliability and validity in non–cardiac arrest populations. The Glasgow Outcome Scale—Extended also has adequate reliability and validity in non–cardiac arrest populations. The modified Rankin Scale addresses symptoms, disability, and level of care on a 7-point scale ranging from 0 (no symptoms) to 6 (death). The most commonly used outcome scale for disability after stroke, the modified Rankin Scale has demonstrated construct validity and adequate interobserver reliability. Standardized instructions ensure reliability of administration.

The Barthel Index and the Functional Activities Questionnaire are summative measures used to measure disability in basic activities of daily living and instrumental activities of daily living in patients with neurological impairments in non–cardiac arrest populations. The modified Barthel Index is used to rate independence in 10 basic activities of daily living (bowel and bladder control, toileting, transfers, wheelchair mobility, stairs, grooming, bathing, feeding, and dressing). The Functional Activities Questionnaire is designed to be completed by a proxy of the patient and rates the patient’s level of dependence on a 4-point scale ranging from 0 (independent) to 3 (dependent) for 10 independent activities of daily living. It also has established validity and reliability.

Quality-of-Life Measures

Health-related quality-of-life measures usually include multiple constructs such as perceived health status (physical and psychological), functional status (activities of daily living and occupational status), social interactions, symptoms, costs, and burden. The 2 major types of health-related quality-of-life tools are profiles, which attempt to measure multiple constructs and can be used across populations or are designed for specific populations, and utility measures, which usually reflect health status and the value of health status. Table 1 lists various health-related quality-of-life measures available for use. Properties deemed essential to a quality-of-life measure include reliability, validity, responsiveness, and sensitivity.

The Health Utilities Index Mark 3 offers functional classification, generic health-related quality of life, and preference-based scores. It has been evaluated in the cardiac arrest population and can be completed in face-to-face or telephone interviews. The Health Utilities Index Mark 3 and the Nottingham Health Profile can be used to address quality-of-life issues relevant to survivors of cardiac arrest.

Cognitive Measures

Implementation of the Chain of Survival, use of disease-modifying therapies, and refinement of implantable cardioverter-defibrillators have led to increased survival from cardiac arrest but with greater attention to the nature and extent of the sequelae of severe brain ischemia. As with other acute ischemic brain injuries, recent studies have shown that higher intellectual function may be a more sensitive measure of brain integrity than motor or sensory function or coordination. With the recognition that CPCs have never been tested for validity and reliability, researchers began administering domain-driven (eg, memory, executive function) standardized measures of neuropsychological function to survivors of cardiac arrest. A comprehensive battery of neuropsychological tests was given as part of a clinical trial to assess the effectiveness of nimodipine versus placebo in treating out-of-hospital cardiac arrest. With no treatment difference across the 2 groups, the combined neuropsychological outcomes at 3 and 12 months showed moderate to severe deficits in 60% and 48%, respectively, in delayed (memory) recall, manual dexterity, calculations, skilled motor movement, planning, initiation, attention, motivation, and depression, a level sufficient for qualification for US Social Security disability income under the classification of organic mental disorders.

Numerous studies have successfully used neuropsychological tests to assess survivors of cardiac arrest. The most common impairments of cardiac arrest survivors are memory, followed by executive dysfunction and attention, consistent with recent findings from magnetic resonance imaging of the brain that indicate that atrophic changes are found globally and are not restricted to the hippocampus. Neuropsychological measures consistent with this impairment triad are the Rey Auditory Verbal Learning Test (memory), Digit Symbol Substitution Test (attention), and Trail Making Tests A and B (executive functioning). For assessment of cognitive measures, it was the consensus of the experts that at least 1 measure of memory, attention/processing speed, and executive function be incorporated into a brief cognitive battery administered at targeted study.
We propose that VOI calculations may be particularly helpful in estimating the costs of not performing the research. There are data, however, demonstrating adequately reliable use of chart-based assessments of neurological function in the outpatient setting. Under these circumstances, telephone measures of memory, attention, and executive function can serve as valid indexes of cognitive function. It is important to note that cognitive measures are sensitive to dysfunction but are not disease specific. The goal therefore is the use of instruments most likely to be affected by relevant neuropathological mechanisms.

What Are the Costs of Primary Outcomes? This section characterizes costs that are frequently hidden but associated with conducting resuscitation research; many costs are related to the primary outcome. There are also public health costs to society when we fail to conduct research or when effective therapies are implemented slowly. The resuscitation researcher needs to consider direct protocol costs, hidden costs, and human costs and must understand cost-effectiveness analysis. We recommend that researchers become familiar with newer concepts, the value of information (VOI) calculations, which can estimate the costs of not performing the research.

Economic Value of Performing Clinical Resuscitation Research: Value of Research Calculation We propose that VOI calculations may be particularly helpful to resuscitation researchers because these calculations can serve to estimate the costs to society if research is not performed and may be a potent argument in favor of conducting research. VOI is a recently developed approach to use decision analysis methods to prospectively assess the expected value of research. These calculations involve initially characterizing the uncertainty of the value of a medical intervention and the optimal decision in the presence of that uncertainty. The approach then asks how additional information would help reduce that uncertainty and how decision making would change given that information. The expected gains in health (such as quality-adjusted life-years) or improvements in costs resulting from that information can then be assessed. VOI calculations can and have been used to establish research agendas, to inform healthcare policy decisions, and to argue for the importance of proposed research.

No studies to date have used VOI calculations to examine the value of cardiac resuscitation research studies. However, consider this example: At a conservative valuation of $50 000 per life-year, the 1 million life-years lost annually to cardiac arrest place the cost of sudden cardiac death at about $50 billion annually. The net present value of research that could reduce cardiac arrest mortality by 10% would therefore be about $75 billion.

Resuscitation research should incorporate VOI methods to better establish the overall impact of resuscitation research and prioritize the resuscitation research agenda. An approach to encourage and disseminate the use of these methods should be established.

Measuring Costs in Resuscitation Studies Not all studies need to perform cost-effectiveness analysis, but it is important to establish the economic effects of new programs and therapies before their widespread dissemination. This assessment requires an accurate comparison of outcomes achieved with the best available alternative, which is best measured in comparative clinical trials. It is useful to measure economic end points in the same trial that establishes clinical efficacy to provide an unbiased measure of the cost-effectiveness of the new therapy. We do not recommend cost as a primary outcome for all studies; however, we do encourage researchers to consider the inclusion of cost as a part of good study design.

Economic outcomes can be measured alongside clinical outcomes in a trial if data collection is designed appropriately. A common contemporary method is to measure the use of key resources (eg, days in intensive care) and then to use a set of standard cost weights for each resource to determine overall costs. A randomized design is particularly valuable in assessing cost because a randomized trial provides an unbiased assessment of outcomes and the attendant “unplanned costs.”

To generate acceptance and dissemination of effective interventions, data collection for economic outcomes, whenever possible, should be incorporated into the initial study design, including postresuscitation strategies. Health gains and losses associated with interventions should be included in the cost-effectiveness analysis. Longer follow-up (3–12 months after hospital discharge) is crucial for valid economic evaluation, as is consideration of costs that go beyond initial emergency medical services or hospital costs and include long-term societal costs. The time horizon should be long enough to capture all major consequences of interventions, especially unplanned costs.

Cost of Delays in the Adoption of Effective Postresuscitation Therapies Researchers should be attuned to the fact that slow adoption of lifesaving methods into clinical practice raises the issue of years of potential life lost as a result. This has not been well explored in resuscitation literature. Therapeutic hypothermia serves as a model to illustrate these important concepts as they relate to resuscitation research.

The cost of slow implementation of therapeutic hypothermia into clinical practice can be described in terms of years of potential life lost prevented at the current adoption rate relative to the optimal adoption rate. Current estimates suggest that only 20% of US hospitals have implemented the use of therapeutic hypothermia in the treatment of survivors of cardiac arrest. Assuming that a modest 20% of patients with ROSC are comatose after arrest and that the number needed to treat is 6, therapeutic hypothermia (at the current adoption rate) prevents 5100 years of potential life lost. If full implementation of therapeutic hypothermia were to occur, an additional 15 300 years of potential life lost could be prevented. The cost of age-adjusted loss in productivity can also be calculated from the years of potential life lost and used to describe the economic impact of delays in implementation of an effective intervention for cardiac arrest.
Hidden Costs and Delays in the Implementation of Protocols Have an Impact on Costs
Studies with longer time points for the primary outcome will generally cost more to conduct. Delays in implementation of resuscitation protocols and other hidden costs are numerous in resuscitation research. Development of a research protocol includes initiation by the investigator, revision by the protocol review committee, and additional revision by the data safety monitoring board. Because of safety concerns, regulatory agencies may require that the entire process be repeated, resulting in significant delays, as well as unexpected and substantially increased study costs. The process of exception to informed consent under emergency circumstances, including community consultation, public notification, and multiple institutional review board approvals, adds to delay. These processes are often interdependent, requiring that the study be stopped and the process repeated each time a significant revision is requested by a regulatory agency or committee. Sample size calculations and “stopping rules” depend on the primary outcome measure, and sample size will affect cost. Another consideration associated with longer-term time points is losing the patient between the time of discharge and time of follow-up. This introduces bias and should be avoided.

Patient safety is the consummate concern of regulators and investigators. The US Food and Drug Administration is congressionally mandated to assure the public of the safety of new drugs and devices. Investigators should communicate with regulatory agencies very early in the development of protocol to avoid delays and increased costs. Resuscitation research requires achieving a balance between an optimal protocol to avoid delays and increased costs. Resuscitation protocols and other hidden costs are numerous in resuscitation research. Development of a research protocol includes initiation by the investigator, revision by the protocol review committee, and additional revision by the data safety monitoring board. Because of safety concerns, regulatory agencies may require that the entire process be repeated, resulting in significant delays, as well as unexpected and substantially increased study costs. The process of exception to informed consent under emergency circumstances, including community consultation, public notification, and multiple institutional review board approvals, adds to delay. These processes are often interdependent, requiring that the study be stopped and the process repeated each time a significant revision is requested by a regulatory agency or committee. Sample size calculations and “stopping rules” depend on the primary outcome measure, and sample size will affect cost. Another consideration associated with longer-term time points is losing the patient between the time of discharge and time of follow-up. This introduces bias and should be avoided.3

Integrating Lessons Learned, Looking Back, and Thinking Forward
During the consensus conference, the experts conducted a series of panels to assist in deliberations through group exercises that examined lessons learned from prior resuscitation studies, looked back at the primary outcomes used by studies that some thought had a major impact on the last set of CPR guidelines, and then practiced “thinking forward” by using 4 mock trial scenarios to determine the primary end point recommended by experts for each hypothetical trial.

Lessons Learned From Prior Studies
Important aspects of selection of outcome were noted from the following studies, which were deemed particularly interesting by the experts. The cited studies demonstrate the need for extreme caution in extrapolating recommendations from animal studies, pilot studies, and short-term survival studies and highlight that some significant differences may not be observable until longer-term time points are evaluated.

High-Dose Epinephrine Studies
Several high-dose epinephrine studies were launched when compelling data revealed that short-term survival rates improved with administration of high-dose epinephrine over standard-dose epinephrine.110 However, a randomized multicenter study showed no improvement in longer-term survival in adults,111 and another trial showed a clear disadvantage in survival to hospital discharge in a pediatric population.112

Amiodarone for VF Studies
The antiarrhythmic amiodarone was recommended for shock-refractory VF, and its use versus placebo or lidocaine was studied in 2 clinical trials. Although there was improvement in rates of successful hospital admission with the use of amiodarone, no improvement in hospital discharge rates was demonstrated.113,114

CPR Without Ventilations: The Osaka Study
An observational study from Japan reported improved neurological outcome among 4068 adults with primary cardiac arrest of cardiac origin when bystanders performed continuous compressions compared with compressions with ventilations.115 However, it is noteworthy that no significance was seen until the end point was examined 1 year after the event. An earlier end point would have shown a nonsignificant trend only.

Looking Back at High-Impact Studies: What Was the Primary End Point?
Another approach to determining the primary end point is to ask what end point was used by prior high-impact studies that weighed heavily in the creation of the Emergency Cardiovascular Care guidelines. To explore this approach, several resuscitation experts generated a nonvalidated list of 10 high-impact studies that were thought to have a major influence on the 2005 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care.115a The list of studies was analyzed for factors related to the primary end point (time point and physiological state). Of interest, 3 observational studies of CPR process variables were cited by experts as having a “high impact on guidelines” but did not include any survival-related outcomes. Seven “high-impact” studies were reviewed for time point and physiological status of the primary outcome.109,115–120 The outcome information from these 7 studies is displayed on a 2-dimensional graph wherein time is represented on the x axis and physiological state is represented on the y axis (Figure 2). It was recognized that current high-impact studies will typically map onto the lower left quadrant (relatively early time points paired with simple outcomes such as “alive versus dead”). In contrast, many experts would like to see future studies populate the opposite quadrant in the upper-right portion of the map with more detailed physiological and neurological assessment and a long-term time point. The distribution of primary outcomes for studies currently considered high impact was noted by many experts as being less than optimal.

Thinking Forward: An Exercise in Assignment of Primary Outcome Measurements
The experts considered 4 mock scenarios for a hypothetical resuscitation study. Each scenario had an assigned panel of experts who discussed the scenario and suggested a primary outcome measurement. The audience was encouraged to participate with questions and comments. After 90 minutes of discussion and debate, all participants were polled on what they...
thought would be the optimal primary end point for each specific study. Each participant selected an optimal time point and an optimal physiological condition for each of the 4 scenarios.

**Resuscitation Scenario 1**

In the first study scenario, an easy-to-deploy ventilatory device was shown to improve 30-day survival in 3 small nonrandomized trials of pediatric respiratory arrest. The investigator proposed implementing the use of the device in 3 well-developed emergency medical services systems to test whether use of the device improved outcome over standard care (no device).

**Results:** The group of experts, which numbered between 31 and 38 during these discussions, was divided on the results. Approximately 45% (14 of 31) believed that the optimal time point was hospital discharge, whereas 26% (8 of 31) supported a time point of 30-day survival. Twenty-three percent (7 of 31) said that a time point ≥6 months should be the primary outcome. Six percent (2 of 31) supported the use of earlier time points. In the selection of a physiological state, 43% (13 of 30) supported the use of a simple CPC/modified Rankin Scale score for assessment of physiological state, whereas 27% (8 of 30) supported the simple “alive versus dead” end point, and 20% (6 of 30) recommended neurocognitive assessments beyond the modified Rankin Scale score. Ten percent (3 of 30) recommended quality-of-life scoring.

**Resuscitation Scenario 2**

In the second study scenario, a unique defibrillation waveform tested only in animals was found to decrease pulseless electric activity and to increase ROSC. A randomized clinical trial was proposed to compare use of the new waveform with a biphasic waveform in a large emergency medical services system. It was hypothesized that this waveform would improve immediate and long-term outcome.

**Results:** The experts were more uniform in their recommendations for this scenario. Eighty-seven percent (34 of 39) thought that the optimal time point was simple ROSC, whereas 10% (4 of 39) supported a 24-hour survival time point, and 3% (1 of 39) supported a hospital discharge time point. For physiological state, 86% (25 of 29) supported the use of the simple “alive versus dead” physiological end point.

**Resuscitation Scenario 3**

In the third study scenario, a small, early, randomized, blinded, single-center, controlled trial of a new medication was proposed. The medication had been shown to improve the rates of short-term ROSC, and a favorable trend had been reported in the hemodynamic profile of patients in the postresuscitation period. The investigators hypothesized that patients treated with medication versus placebo would have better outcomes.

**Results:** Among the experts, 71% (24 of 34) felt that the optimal time point was hospital discharge, whereas 26% (9 of 34) supported a 30-day time point. For physiological assessment, 64% (21 of 33) supported a CPC/modified Rankin Scale score, and 33% (11 of 33) supported neurocognitive or more advanced assessment.

**Resuscitation Scenario 4**

In the fourth study scenario, a chemical hibernation nanoagent had been shown to improve 30-day survival rates and neurological status in patients with non-VF cardiac arrest in a large randomized European study. A national consortium proposed a definitive randomized controlled trial to determine whether this agent should become the standard of care.

**Results:** Among the experts, 82% (28 of 34) thought that 3 months was the optimal time point, 12% (4 of 34) supported a 6-month time point, and 6% (2 of 34) supported a simple hospital discharge time point. For physiological assessment, 52% (16 of 31) supported the use of a combined neurocognitive/quality-of-life assessment, whereas 35% (11 of 31) supported the use of a CPC/modified Rankin Scale score.

**Conclusions From the Consensus Exercises**

It is important to note that there was substantial variation among resuscitation experts on the assignment of primary outcome when presented with the hypothetical scenarios. Even after 2 days of discussion among the experts, true consensus is imperfect.

**Final Recommendations**

**Consensus on Guiding Principles**

The process for preparation of this statement used multiple techniques to approach consensus, including circulation of preconference position papers; participation by a diverse group of experts at the conference; reviews of published literature; ad hoc reviews by experts and conference participants; discussions among experts in the “looking back,” “thinking forward,” and summary presentations; and creation of the final document, which underwent multiple reviews of content and recommendations. The potentially useful outcome measures for specific organ assessment are summarized in Table 2. During the consensus process, the following guiding principles were developed and recommended for determination of primary study outcomes:

1. No single recommended primary outcome is appropriate for all studies of cardiac arrest or resuscitation.
Careful selection of the primary outcome variable is an essential and vital part of every study. There was not complete consensus among experts when real scenarios were tested.

2. The 3 major considerations when creating a primary end point for a study are the time point after arrest to determine outcome, the physiology assessed at the time point, and the costs associated with use of the outcome. These are the when, what, and how much questions that must be answered.

3. The best primary outcome for a given study is the pairing of a time point and physiological condition that will answer the question and thus allow a valid determination of the superiority or inferiority of the specific therapy under study while minimizing the cost of the study.

4. Although no single primary outcome is appropriate for all studies, the following general themes have emerged to guide researchers:

a. For very early studies that are more likely to have fewer patients, it is appropriate to use early endpoints, including ROSC, simple “alive versus dead,” hospital mortality, or a hemodynamic parameter as a primary outcome.

b. For larger studies, it is appropriate to use endpoints that are at a longer time point after the arrest. Because the patient’s neurological condition fluctuates widely during the first 90 days after arrest, stable neurological evaluations should consider a 90-day time point for assessment.

c. For very large randomized controlled trials that may have a major impact on public health policy, longer-term end points such as 90 days coupled with neurocognitive and quality-of-life assessments should be considered. In these trials, some data may have to be excluded because of potential loss of patient follow-up, which must be considered during the design of the trial. This recommendation may require future validation because it is not evidence based.

5. For studies in which the need for regulatory involvement can be anticipated, early discussions with regulatory agencies to consider the primary end point (and other issues) are recommended.

6. Barriers to resuscitation research will ultimately result in human deaths. Thus, it is vital that research be performed. Patients will die unnecessarily when resuscitation research is delayed or not performed. These factors need to be considered in the conduct of resuscitation research by regulatory bodies and the public.

7. There are significant potential costs for not performing research, including substandard clinical care and potential years of life lost because of preventable conditions. VOI analysis may assist in quantifying the costs of not performing resuscitation research. These human costs need to be recognized by all stakeholders in resuscitation research.

Future Directions

There was full consensus that the lack of an easy-to-administer, validated neurological functional outcome is a major limitation to the field. Moreover, it was thought that with a focused effort, an improved scoring tool could be developed in the near future. It is therefore recommended that organizations involved in resuscitation research prioritize the development of such a scale for assessment of neurological function to be used in future studies. This would have significant value to all researchers who seek to improve public health.
<table>
<thead>
<tr>
<th>Writing Group Member</th>
<th>Employment</th>
<th>Other Research Support</th>
<th>Speakers' Bureau/ Honoraria</th>
<th>Expert Witness</th>
<th>Ownership Interest</th>
<th>Consultant/Advisory Board</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lance B. Becker</td>
<td>University of Pennsylvania, Academic Medical Center</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Benjamin S. Abella</td>
<td>Hospital of the University of Pennsylvania</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Christophe Adele</td>
<td>Cochin Hospital Assistance Publique de Paris</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Tom P. Aufderheide</td>
<td>Medical College of Wisconsin Medical College/University</td>
<td>Resuscitation Outcomes Consortium Grant (NHLBI, money goes to the institution); Neurological Emergency Treatment Trials (NETT) Network (NINDS, money goes to the institution); Immediate Trial (NHLBI, money goes to the institution); ResQTrial (NHLBI, money goes to the institution); BeneChill Hypothermia Study*</td>
<td>Received software and equipment for Zoll Medical for Immediate Trial and Resuscitation Outcomes Consortium grants*</td>
<td>EMS Today 2008 Lecture*</td>
<td>None</td>
<td>None</td>
<td>Consultant for JoLife for a research project using their mechanical CPR device (money goes to the institution; consultancy discontinued July 2008); Medtronic (money goes to the institution); Secretary of the Board of Directors for Take Heart America Project*, volunteer; President and Chairman of the Board of Directors, Citizen CPR Foundation*, volunteer</td>
</tr>
<tr>
<td>Robert A. Berg</td>
<td>University of Pennsylvania, Children's Hospital of Philadelphia</td>
<td>NICHD grant as PI for CHOP site of the Collaborative Pediatric Critical Care Research Network†</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Clifton W. Callaway</td>
<td>University of Pittsburgh School of Medicine; UPMC Health Care System</td>
<td>American Heart Association Grant-in-Aid†, Cerebrovascular Effects of Thrombin Activation After Cardiac Arrest, 2007–2009, principal investigator (This project explores relationships between hypothermia, coagulation, and inflammation on brain recovery after cardiac arrest in a rat model); NHLBI U01HL077871; Pittsburgh Resuscitation Network, 2004–2016, principal investigator (This project is as a data collection site for a multicenter network to perform large clinical trials to improve outcomes after out-of-hospital cardiac arrest and life-threatening injury); Grant to the University of Pittsburgh for time as AHA Worksheet Expert Reviewer for 2010 Guidelines process (2009–2010)†</td>
<td>Loan of equipment from Medtronic, Inc, to support laboratory studies of hypothermia*</td>
<td>Hypothermia After Cardiac Arrest, 2009 Annual Meeting, Pennsylvania chapter, American College of Cardiology, October 3, 2009†; Keeping It Cool: How to Implement Hypothermia in Your Clinical Practice, Society for Critical Care Medicine, April 22–23, 2010†; Managing the Patients at Risk of Sudden Cardiac Death–Sudden Cardiac Arrest Association, Pittsburgh, PA; October 8, 2010†; Raising the Bar for Post-Arrest Care–Take Heart Austin, Texas; October 21, 2010†; Post-Cardiac Arrest Symposium, St. Mary's Hospital, Seoul, South Korea; December 18, 2010†</td>
<td>None</td>
<td>University of Pittsburgh licensed a patent to Medtronic ERS for technology related to ECG waveform analysis in ventricular fibrillation; coinventor (may receive royalties)†; Owns stock in Apple Computer*</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>University of Pittsburgh licensed a patent to Medtronic ERS for technology related to ECG waveform analysis in ventricular fibrillation; coinventor (may receive royalties)<em>; Owns stock in Apple Computer</em></td>
<td>(Continued)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
This table represents the relationships of writing group members that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all members of the writing group are required to complete and submit. A relationship is considered to be “significant” if (a) the person receives $10,000 or more during any 12-month period, or 5% or more of the person’s gross income; or (b) the person owns 5% or more of the voting stock or share of the entity, or owns $10,000 or more of the fair market value of the entity. A relationship is considered to be “modest” if it is less than “significant” under the preceding definition.

*Modest.
†Significant.
Reviewer Disclosures

<table>
<thead>
<tr>
<th>Reviewer</th>
<th>Employment</th>
<th>Research Grant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ben J. Bobrow</td>
<td>Arizona Department of Health Services</td>
<td>None</td>
</tr>
<tr>
<td>Salvador Cruz-Flores</td>
<td>Saint Louis University</td>
<td>None</td>
</tr>
<tr>
<td>Graham Nichol</td>
<td>University of Washington</td>
<td>Resuscitation Outcomes Consortium (NIH U01 HL077863-05)†; Evaluation of CPR Training Aid (Laerdal Foundation)‡; Randomized Trial of Cold Saline IV After Resuscitation (NIH)§; Resynchronization/Defibrillation for Advanced Heart Failure Trial RAFT Medtronic and Canadian Institutes for Health Research; Outcome and Cost Effectiveness of FDG PET in LV Dysfunction (PARR-2) Heart and Stroke Foundation of Canada; Novel Methods of Measuring Health Disparities (NIH 1RC2HL101759-01)†; Cascade Cardiac Resuscitation System of Care (Medtronic Foundation)‡</td>
</tr>
<tr>
<td>Tom Rea</td>
<td>University of Washington</td>
<td>None</td>
</tr>
</tbody>
</table>

This table represents the relationships of reviewers that may be perceived as actual or reasonably perceived conflicts of interest as reported on the Disclosure Questionnaire, which all reviewers are required to complete and submit. A relationship is considered to be “significant” if (a) the person receives $10,000 or more during any 12-month period, or 5% or more of the person’s gross income; or (b) the person owns 5% or more of the voting stock or share of the entity, or owns $10,000 or more of the fair market value of the entity. A relationship is considered to be “modest” if it is less than “significant” under the preceding definition. *Modest. †Significant.

References


Appendix

Primary Outcome for Resuscitation Science Studies Consensus Conference
Participants, Advisors, and Observers

May 5-6, 2008
Washington, DC

Benjamin S. Abella, MD, MPhil, University of Pennsylvania, Philadelphia, PA

Christophe Adrie, MD, Cochin Hospital, Paris, France

Dianne L. Atkins, MD, University of Iowa Hospital and Clinics, Iowa City, IA

Tom P. Aufderheide, MD, Medical College of Wisconsin, Milwaukee, WI

Eric P. Bastings, MD, US Food and Drug Administration, Silver Spring, MD*

Lance B. Becker, MD, University of Pennsylvania Health System, Center for Resuscitation Science, Philadelphia, PA

Wilhelm Behringer, MD, Vienna General Hospital, Medical University of Vienna, Vienna, Austria

Robert A. Berg, MD, University of Pennsylvania School of Medicine, Philadelphia, PA

Nicholas G. Bircher, MD, University of Pittsburgh Medical Center, Pittsburgh, PA

Clifton W. Callaway, MD, PhD, University of Pittsburgh, Pittsburgh, PA

Arthur L. Caplan, PhD, University of Pennsylvania–Center for Bioethics, Philadelphia, PA

Michael N. Diringer, MD, Washington University School of Medicine, St Louis, MO

Michael W. Donnino, MD, Beth Israel Deaconess Medical Center, Brookline, MA

Brian Eigel, PhD, American Heart Association, Dallas, TX
Michael T. Torbey, MD, MPH, Medical College of Wisconsin, Milwaukee, WI

Oscar H. Tovar, MD, US Food and Drug Administration, Rockville, MD*

Steve Trzeciak, MD, MPH, Cooper University Hospital, Camden, NJ

Terry L. Vanden Hoek, MD, The University of Chicago, Chicago, IL

Myron L. Weisfeldt, MD, Johns Hopkins Medical Institutions, Baltimore, MD

Catherine P. Wentz, MS, US Food and Drug Administration, Rockville, MD*

Jane G. Wigginton, MD, UT Southwestern Medical Center, Dallas, TX

Demetris Yannopoulos, MD, Johns Hopkins University, Baltimore, MD

G. Bryan Young, MD, University Hospital/University of Western Ontario, London, Ontario, Canada

Barbara A. Zimmerman, US Food and Drug Administration, Rockville, MD*

Bram D. Zuckerman, MD, US Food and Drug Administration, Rockville, MD*

*Conference attendees who were present as observers only. Attendance by observers is not to be construed as an endorsement of either the conference or this statement, nor as a contribution in any manner to the drafting of this statement.