Evaluating Patient-Centered Outcomes in Clinical Trials of Procedural Sedation, Part 2

Safety: Sedation Consortium on Endpoints and Procedures for Treatment, Education, and Research Recommendations

Denham S. Ward, MD, PhD,*† Mark R. Williams, MBBS, BSc,‡ John W. Berkenbosch, MD,§ Maala Bhatt, MD,¶ Douglas Carlson, MD,¶¶ Phillip Chappell, MD, MBA,** Randall M. Clark, MD,†† Isabelle Constant, MD, PhD,‡‡ Aaron Conway, RN, PhD,§§ Joseph Cravero, MD,¶¶¶ Albert Dahan, MD, PhD,¶¶¶¶ Franklin Dexter, MD, PhD,## Raymond Dionne, DDS, PhD,### Robert H. Dworkin, PhD,#### Tong J. Gan, MBA, MHS, FRCA,##### David Gozal, MD,§§§§ Steven Green, MD,¶¶¶¶¶ Michael G. Irwin, MBChB, MD,¶¶¶¶ Jennifer Lightdale, MD, MPH,¶¶¶¶ Ronald S. Litman, DO,§§§§§ Keira P. Mason, MD,¶¶¶¶ James Miner, MD,§§§§§¶¶¶¶Robert E. O’Connor, MD, MPH,¶¶¶¶¶ Pratik Pandharipande, MD,****** Richard R. Riker, MD,††††††††† Mark G. Roback, MD,¶¶¶¶¶ Daniel I. Sessler, MD,§§§§§ Anne Sexton, MD,¶¶¶¶¶¶¶¶ Joseph R. Tobin, MD,¶¶¶¶¶¶ Dennis C. Turk, PhD,#### Mark Weiss, MD,¶¶¶¶ Hannah Wunsch, MD, MSc,§§§§§§¶¶¶¶¶¶¶ and Anna Zhao-Wong, MD, PhD¶¶¶¶¶¶

The Sedation Consortium on Endpoints and Procedures for Treatment, Education, and Research (SCEPTER) has previously published recommendations on measuring efficacy in procedural sedation clinical trials.1 Subsequently, SCEPTER organized a meeting to identify core measures for the safety domain not addressed in the first meeting. SCEPTER was established by the Analgesic, Anesthetic, and Addiction Clinical Trial Translations, Innovations, Opportunities, and Networks (ACTTION) public–private partnership with the US Food and Drug Administration, a public–private partnership with the US Food and Drug Administration, which has received unrestricted grants to support the activities of ACTTION. Financial support for this project was provided by ACTTION, which has received research contracts, grants, or other revenue from the US Food and Drug Administration, multiple pharmaceutical and device companies, philanthropy, and other sources.

Reprints will not be available from the authors.

Address correspondence to Denham S. Ward, MD, PhD, Department of Anesthesiology and Perioperative Medicine, University of Rochester School of Medicine and Dentistry, 601 Elmwood Ave, Box 604, Rochester, NY 14642. Address e-mail to Denham_Ward@URMC.Rochester.edu.
The development of appropriately designed clinical trials has the potential to ensure that critical data are collected to measure important safety outcomes. For example, clinical trials of midazolam in the 1980s largely failed to identify the drug’s potential for ventilatory depression, particularly in patients who are elderly and debilitated. Clinical trials evaluated midazolam alone, but in common clinical use, it was combined with an opioid, resulting in >20 serious adverse events (SAEs) reported to the FDA soon after its release. Subsequently, the package insert for midazolam was revised, suggesting a 50% reduction in the recommended dose for patients who are elderly and debilitated. This episode demonstrates the difficulty in designing clinical trials to identify AEs that are rare and may occur only in subsets of at-risk patient populations or under unanticipated circumstances.

SCEPTER’s previous study used the Institute of Medicine’s (IOM’s) 6 aims in health care in Crossing the Quality Chasm as the domains that clinical trials should address. That report defined the safety domain as avoiding harm to patients from care intended to help them. SCEPTER’s concept for safety was “avoidance of physical or psychological harm” because procedural sedation is not by itself therapeutic but allows such procedures to be performed successfully without distressing pain or anxiety and, more importantly, without preventable harm. While safety is arguably the most important of the 6 IOM domains, its measurement in clinical trials presents complex problems and dilemmas, the foremost being the problem of cataloging rare AEs in clinical trials of a reasonable sample size.

The purpose of our recommendations is to provide some guidance in the design of clinical trials for procedural sedation that will provide information about the safety of new drugs, devices, or protocols and their efficacy. There are few agreed-on definitions for what constitutes an AE or which events actually represent potential risk to patients. It is understood that no reasonably sized set of clinical trials (ie, phase 1–3 trials) can provide definite answers for all possible AEs. However, it is hoped that these recommendations will help investigators design clinical trials that will provide useful data in the understanding of the safety profiles for new drugs, devices, or procedures used for procedural sedation.

**METHODS**

At the first SCEPTER meeting, the domains proposed by the IOM were utilized to recommend the evaluation of efficacy in procedural sedation clinical trials. At the second SCEPTER meeting held on November 17–19, 2016 in Washington, DC, recommendations were developed to evaluate safety. Individuals invited to the first meeting and others with specific expertise regarding safety were invited to this meeting (N = 56), and 46 participated. Participants received 2 review articles before the meeting; selected individuals were asked to facilitate meeting sections by providing brief presentations to stimulate discussion. Transcriptions of entire meeting recordings are available on the ACTTION website.

The meeting and this subsequent article were organized into 3 parts. First, a systematic review of the literature was undertaken to investigate safety measurements in procedural clinical trials. Second, a discussion of general clinical trial methodology, including standardization of reporting and statistical consideration, was undertaken with an emphasis on the problems of infrequent events. Third, we examined by organ system what measurements to make in clinical trials to detect and quantify adverse effects, events, and outcomes.

For the systematic review of safety studies, an independent professional librarian from the University of Rochester conducted PubMed searches (up to July 2016) for prospective studies comparing sedative agents/regimes for procedures using a comprehensive search strategy (Supplemental Digital Content, Appendix, http://links.lww.com/AA/C360). Only prospective randomized double-blind studies reported as full-text articles published in English were included. Retrospective and case studies were excluded, as were those involving general anesthesia or conducted with sedation in intensive care units. Studies using only electrophysiological measures to monitor sedation level (eg, bispectral index, auditory evoked potentials) were omitted. Abstracts identified through the literature were searched for relevance. References from relevant retrieved articles were also reviewed to identify any pertinent articles missed by the PubMed search. Articles were read in full, and physiologic parameters, including duration of change, representing AEs pertaining to sedation safety assessment were extracted. Defined side effects/AEs in each article were categorized by organ system and tabulated in a spreadsheet. General study characteristics, including author, year, and patient population, were also recorded.

The discussion by organ system attempted to distinguish among adverse effects, AEs, and adverse outcomes in procedural clinical trials. An adverse outcome was considered to be an event that actually causes patient harm or discomfort, for example, respiratory arrest requiring intubation. AEs are actual episodes during a procedural sedation that might cause harm or require an intervention, for example, moderate oxygen desaturation. Adverse effects refer to physiological effects of agents that may indicate an increased risk for an AE, for example, a moderate increase in Paco2.

The initial draft of this article, developed by the first author, was prepared using the meeting transcripts, notes, and materials presented at the meeting. The meeting’s general approach was to reach a consensus based on available data and the in-depth discussions of the meeting attendees, with further agreement being reached by iterative manuscript review and revision.

**RESULTS**

The literature search yielded 545 articles: 408 were excluded after initial review (by M.R.W.) of the title and/or abstract, and 137 were considered relevant and included. Among these articles, 6 assessed only premedication, not procedural sedation, and were therefore excluded. References of the remaining 131 articles were reviewed, identifying 2 more relevant articles. A total of 133 articles (Supplemental Digital Content, Appendix, http://links.lww.com/AA/C360) contains the
Supplemental Appendix Figure and Supplemental Appendix Table with the specific search strategy, flow chart, a summary table, and a bibliography of the articles reviewed) were included in our initial analysis of the safety measures used in procedural sedation literature. Only 6 of the 133 studies reviewed used a formal sedation safety measure to evaluate safety outcomes: the Quebec (3 adult and 1 pediatric article) and the World Society for Intravenous Anesthesia (SIVA) guidelines (2 pediatric articles).

The severity of recorded AEs was inconsistently defined throughout the studies assessed and included those that interfered with the patient’s daily functioning, was of significant intensity and interrupted the procedure, or required treatment. SAEs, as defined by the FDA, were mentioned in only 4 studies.

Monitoring patients’ vital signs was the most common method of collecting safety data, utilized in 132 of the 133 articles assessed. The remaining study reported no specific methods for collecting safety data. In 60 articles, vital signs and side effects were collected, but no further information was reported regarding specific physiological variables. The other 72 articles (48 adult, 23 pediatric, and 1 combined) recorded vital signs with defined physiological parameters and listed or defined AEs and side effects. The duration of an event necessary for classification as an AE was defined in 20 studies (14 adult, 6 pediatric). The duration predominantly involved a period that a given vital sign was outside a defined range, for example, oxygen saturations <90% for >30 seconds. The range was 10 seconds to 1 minute in the studies reviewed. One study recorded the cumulative time during which a vital sign was outside the specified range.

Respiratory and airway management items were recorded in all 72 articles, which reported the safety data collected. These items primarily included oxygen saturation, respiration rate, and apneic episodes, but also AEs such as laryngospasm, bronchospasm, and need for airway intervention. Seven studies specifically reported using end-tidal partial pressure of carbon dioxide (P_{ET}CO_{2}) as a monitoring modality; 5 specified P_{ET}CO_{2} as a safety outcome. Fifty reviewed studies (34 adult and 16 pediatric) recorded cardiovascular items. Vital signs were limited to blood pressure (BP), heart rate (HR), and arrhythmias. Gastrointestinal items recorded in 16 adult and 12 pediatric studies included nausea, vomiting, or retching. The use of antiemetics was reported in 3 adult and 1 pediatric study. Neurological items were captured in 25 studies (14 adult, 11 pediatric). One study prespecified a safety endpoint of “excessive sedation (observer’s assessment of alertness/sedation [OAA /S] ≤2);” however, the remainder of the articles reported neurological side effects. These focused on neurological symptoms and often included patient-reported side effects such as headache, vertigo, visual disturbances, unpleasant dreams, or prolonged amnesia. Observer-reported evaluations, including patient restlessness, confusion, agitation, delirium, or disinhibition, were used. Formal testing with the Mini-Mental State Examination was used in 1 study to evaluate any decline after sedation; 1 study tested a 3-object recall. Other items recorded in studies included symptoms such as dry eyes, dry mouth, increased secretions, rash, or pruritus. Prolonged emergency department stay, unplanned hospital admission, and use of reversal agents (naloxone and/or flumazenil) were sometimes noted as safety assessment components in the clinical trials.

In one of the largest data sets, Cravero et al. summarized data collected prospectively on 30,037 sedation/anesthesia encounters in pediatric patients (≤18 years of age) outside of the operating room. AEs, including inadequate sedation, were reported in 5.3% of the records. Desaturation <90% was reported in 1.6% of the records, vomiting during the procedure in 0.5%, and unexpected apnea in 0.24%. None of the records revealed any long-term morbidity, although there was 1 case of aspiration and 1 case of hypoxia leading to the need for cardiopulmonary resuscitation. As expected, respiratory events were more frequent for very young patients (≤1 year of age).

**DISCUSSION**

**General Methodological Recommendations**

Clinical trials primarily designed to collect safety data are different from designs when efficacy data are the primary focus. Efficacy trials will prespecify the outcomes and the specific measurements needed to show if that outcome is achieved. Preliminary data from preclinical and small phase 1 studies will not identify all possible AEs in advance. Broad-based measurements and event coding are required to ensure that possible AEs are not missed. The detection of rare events may require pooled data analysis from different trials along with designs that attempt to minimize type 2 errors. Still, incorrect data pooling from different studies can result in errors. Because there are unique methodological issues for measuring rare events, issues arise when data from clinical trials designed to test efficacy are combined to report on safety. Because meta-analysis of multiple studies is often used to gain insight into safety, data collected from a single study should be reported to facilitate subsequent pooling in a meta-analysis.

The reporting of safety data is important for regulatory activities. Medical Dictionary for Regulatory Activities (MedDRA) is the standard terminology for documenting drug safety (http://www.meddra.org). The use of MedDRA to code for AEs permits standardized data retrieval and comparison of studies, from phase 1 studies through the end of the product life cycle. MedDRA is used, but not required, as the AE terminology in FDA safety databases. MedDRA is structured in 5 layers, from the overall System Organ Class down 4 layers to “preferred terms” at the lowest level, with each level having synonym(s) and/or quasi-synonym(s) as subordinate lowest level terms. There are currently 21,920 preferred terms in version 19.0. For example, under the system organ class of “cardiac disorders,” a preferred term is “arrhythmia”; then lower level terms provide more specificity. However, MedDRA’s terms do not include descriptors that are important in clinical trial data, such as disease severity, demographic terms, or AE severity, although severity is often included as a supplementary rating. Subsets of MedDRA terms have been used as a basis for other AE terminologies, most notably by the National Cancer Institute in the Common Terminology Criteria for Adverse Events, with an added severity scale to assess toxicity in oncology therapeutic trials.
The FDA has also provided guidance for the collection of selective safety data for late-stage premarket and post-approval clinical investigations. There have been other efforts to standardize the terminology reporting AEs from procedural sedation. These reporting tools are designed for reporting AEs for quality improvement (QI) databases and are useful starting points for defining AEs in randomized controlled clinical trials for new drugs, protocols, or devices.

Clinical trials of procedural sedation are different from data collected for routine clinical QI databases. Clinical trials are generally tightly controlled, with careful selection of patients and sedation regimens, and the provider administering the sedation. A clinical trial with a tightly controlled protocol and experienced providers may show fewer AEs than would occur in actual clinical practice with less experienced providers, but in routine clinical practice without rigidly define protocols, some AEs may not be recorded. Rather than a tightly controlled clinical trial protocol, a pragmatically designed clinical trial allowing more variability in who provides the sedation, as well as other factors, including the use of adjunct sedative agents, may uncover more AEs as long as there are correctly defined measurements. Because a specific clinical intervention can modify the time course of a potential AE (eg, the duration of a desaturation episode terminated by a jaw thrust), all clinical trials must carefully prespecify both the criteria for recording an AE and the criteria for any intervention performed with the intent of treating or managing the event. This method has been espoused in what have become referred to as the “Quebec Guidelines.” Although subject to individual provider practice variations, intervention-based definitions have the benefit of providing a degree of clinical significance and relevance to sedation AEs and outcomes. For QI databases, the documentation needed for a clinical trial may be too laborious for consistent routine use.

An approach to AE documentation has also been proposed by a World SIVA task force. The proposed documentation includes (1) a description of the AE, (2) any interventions performed, (3) the outcome of the AE, and (4) the severity rating. This severity rating scale has 4 levels: minimal, minor, moderate, and sentinel. This format combines both interventions and events/thresholds. The SIVA reporting tool (www.AESedationReporting.com) is open access and is online for individual and institutional reporting and tracking. An independent but related classification and reporting effort has been undertaken by the International Committee for the Advancement of Procedural Sedation (www.ProceduralSedation.org). Their development of the Tracking and Reporting Outcomes of Procedural Sedation (TROOPS) tool and definitions is based on both the World SIVA tool and the Quebec Guidelines. The TROOPS tool is intended to be compliant with the MedDRA terminology. Data collected for QI of current sedation drugs, techniques, devices, and practices (including clinical setting, patient selection, experience level of the provider administering sedation, etc) can help clarify the important AEs and design clinical trials that will detect these AEs.

Efforts to classify and report AEs must also account for the differences between pediatric and adult procedural sedation. Clinical indications for sedation in children often differ from those in adults and may commonly involve sedation for nonpainful procedures only requiring a lack of movement (eg, magnetic resonance imaging). Importantly, the level of sedation must be carefully determined, particularly for nonpainful procedures in children, because ventilatory depression presumably increases with the depth of sedation, and without a stimulus, it may be difficult to distinguish sedation from general anesthesia. The American Society of Anesthesiologists’ statement on the continuum of depth of sedation provides useful definitions.

Organ System–Based Recommendations

The recommended measurements that should be made to detect adverse effects, events, and outcomes in procedural sedation clinical trials are summarized by organ system in the Table. These measurements need to be supplemented with the criteria and occurrence for any clinical interventions that take place.

**Pulmonary.** Ventilatory depression is the main safety consideration in procedural sedation; therefore, safety trials need to be designed to have appropriate measurements and patient populations to be able to measure the amount of ventilatory depression. Ventilatory depression is complex to measure and report. Many techniques have been proposed and utilized to study the effects of drugs on ventilatory control, but no standards have evolved. Measurements can be divided into those that simply monitor, in varying detail, ventilation, oxygen saturation, and airway CO2 and those that variably stimulate or challenge ventilation in a variety of ways (eg, hypercapnic ventilatory response). The former is suitable for large clinical studies while the latter requires specialized apparatus and monitoring devices, thus being suitable only for smaller phase 1 studies.

All studies investigating ventilatory AEs require the continuous monitoring and recording of oxygen saturation (SPO2) and the partial pressure of airway CO2. For SPO2, there is no consensus for what the threshold and time below the threshold is required to constitute an AE. When end-tidal CO2 is calculated from continuous airway CO2, it is important to recognize any artifacts (eg, a low value caused by shallow breathing or a misapplication of the sensor). For example, if the end-tidal values for several breaths are averaged together, as is commonly done, a few artifactual low values may result in an underestimation in the actual increase in CO2. Similarly, the length of any apneic periods needs to be confirmed by an inspection of the continuous tracing. The length of a ventilatory pause that constitutes an apnea should be defined as part of the protocol. For apnea, the number and length of episodes should be specified, with consideration of standards set forth by the sleep medicine community for central and peripheral apneas. A calculation of the apnea-hypopnea index for a new procedural sedation drug could be made; however, it is not known how an apnea-hypopnea index measured during sedation correlates with AEs during procedural sedation. Because clinical interventions may be required for safety, the data should precisely record the criteria used when an

---

*A “sentinel” adverse event was defined as “those critical enough to represent real or serious imminent risk of serious and major patient injury.”*
Table. Summary of Recommendations for Core and Supplemental Measurements for Procedural Sedation Clinical Trials

<table>
<thead>
<tr>
<th>Safety Subdomain</th>
<th>Core Measurements</th>
<th>Supplemental Measurements</th>
<th>Adverse Effects</th>
<th>Adverse Events</th>
<th>Adverse Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulmonary</td>
<td>Continuous SpO₂</td>
<td>HCVR, HVR Pcrit</td>
<td>Hypoventilation</td>
<td>Desaturation (number, threshold, and duration)</td>
<td>Mechanical ventilatory support including BIPAP or BMV</td>
</tr>
<tr>
<td></td>
<td>(use of supplemental O₂ must be fully documented)</td>
<td>Sedation-induced AHI</td>
<td>UAO</td>
<td>Hypercapnia (&gt;50 mm Hg)</td>
<td>Pulmonary aspiration</td>
</tr>
<tr>
<td></td>
<td>Continuous airway CO₂</td>
<td>(end-tidal CO₂)</td>
<td>Apneas (length, number, central, or peripheral)</td>
<td>Intervention to cause arousal or use of reversal agents</td>
<td>Hypoxia leading to organ injury</td>
</tr>
<tr>
<td></td>
<td>Need for manual airway support, eg, chin lift or insertion of airway support device</td>
<td></td>
<td>Depresion of HVR and CVR</td>
<td>Bronchospasm (with or without treatment)</td>
<td></td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>Continuous HR and rhythm BP (invasive or noninvasive) recorded at least every 5 min</td>
<td>NIRS Plasma lactate (long sedation)</td>
<td>Hypotension Bradycardia</td>
<td>Hypotension requiring intervention</td>
<td>Medical/ electrical cardioversion CPR</td>
</tr>
<tr>
<td></td>
<td>Required interventions (eg, anticholinergic or epinephrine for bradycardia or a fluid bolus for persistent hypotension)</td>
<td></td>
<td>Dysrhythmia (relative bradycardia or tachyarrhythmias)</td>
<td>Symptomatic bradycardia</td>
<td></td>
</tr>
<tr>
<td>Neurological</td>
<td>MMSE, 3 object recall pre- and post-sedation</td>
<td>Cardiovascular disease history Postprocedure follow-up for distressing neurological symptoms Delirium score</td>
<td>Changes in level of consciousness</td>
<td>Excessive sedation</td>
<td>Seizures</td>
</tr>
<tr>
<td></td>
<td>Postprocedure follow-up for distressing neurological symptoms</td>
<td>Elderly Dementia Infants, particularly premature infant</td>
<td></td>
<td>Use of sedation reversal agents</td>
<td>Myoclonus interfering with procedure</td>
</tr>
<tr>
<td></td>
<td>Delirium score</td>
<td></td>
<td></td>
<td>Prolonged recovery</td>
<td>Amnesia lasting &gt;24 h</td>
</tr>
<tr>
<td>Gastrointestinal</td>
<td>Qualitative nausea scale (eg, VRS in adults and BARF scale in children) Antiemetic use and timing, particularly if used as a prophylaxis</td>
<td>Non-smokers Young women Previous PONV and motion sickness history</td>
<td>Decrease in GI motility</td>
<td>Prolonged vomiting</td>
<td>Delirium lasting &gt;22 h</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Vomiting Nausea</td>
<td>Postprocedure cognitive decline</td>
</tr>
<tr>
<td>Other</td>
<td>Patient symptom or distress score</td>
<td></td>
<td>Pain on injection Myalgia</td>
<td>Provider dissatisfied Post-traumatic stress disorder</td>
<td></td>
</tr>
</tbody>
</table>

Adverse effects refer to physiological effects of agents that may indicate increased risks for adverse events or outcomes. These adverse effects are sometimes measured in laboratory phase 1 clinical trials with provocation. Adverse events refer to actual episodes during a procedural sedation that may require an intervention to prevent patient harm. Adverse outcomes refer to an adverse event that actually causes patient harm or discomfort. The list of adverse effects, events, and outcomes are examples and not intended to be exhaustive. Severe adverse outcomes are not listed (ie, death, cardiac arrest, respiratory arrest). See text for further discussion and references.

Abbreviations: AHI, apnea-hypopnea index; BARF, Baxter animated retching faces scale; BIPAP, bipressure airway support; BMV, bag mask ventilation; BP, blood pressure; CPR, cardiopulmonary resuscitation; HCVR, hypercapnic ventilatory response; HR, heart rate; HVR, hypoxic ventilatory response; MMSE, Mini-Mental Status Exam; NIRS, near-infrared spectroscopy; OSAS, obstructive sleep apnea syndrome; Pcrit, critical airway closing pressure; PONV, postoperative nausea and vomiting; QTc, corrected ECG QT interval; SpO₂, oxygen saturation; UAO, upper airway obstruction; VRS, 11-point (0–10) verbal rating scale.
intervention is made and a complete description of the specific intervention. When there is no intervention, reporting should provide data on the total time and on longest continuous time segment below (for saturation) or above (for end-tidal CO₂) a specified threshold.

Many studies have used a physiological stimulus to ventilation as a sensitive indicator of the depressive effects of drugs. The classic rebreathing test described by Read²³ has been used to obtain a change in the slope and intercept of the linear ventilatory response to hypercapnia, but there are more modern methods.²⁴ More recently, examination of the hypoxic drive to ventilation has been used as a provocative test. In both these tests, the experimental conditions are extremely important in determining the outcome. Sleep,²⁴ pain,²⁵,²⁶ and arousal state²⁷ have drug-specific interactions with both the hypercapnic and hypoxic ventilatory responses.²⁸ Other provocative tests have been applied. One test of considerable research interest is the airway pressure below which upper airway obstruction patency pressure that is provided by continuous positive airway pressure (CPAP). It has been found that anesthetics²⁸ and sedatives²⁹ increase the Pcrit but it is not known how this correlates with adverse ventilatory events.²²

While no recommendations can be made regarding the optimal provocation to elicit ventilatory vulnerability, it is recommended that the effects of new drugs for procedural sedation have their effects on the hypercapnic and hypoxic ventilatory response measured in normal subjects in phase 1 clinical trials.

Cardiovascular. Cardiovascular effects are also common and second in frequency only to ventilatory adverse effects. However, they tend to be relatively minor (eg, a nonhemodynamically significant bradycardia) and are less likely to require an intervention. Primary (ie, not secondary to hypoxia) cardiac arrest requiring cardiopulmonary resuscitation or an arrhythmia requiring cardioversion is extremely rare and would not be expected to occur in a clinical trial of any reasonable size.

Larger phase 2 and 3 studies should record HR continuously. BP should be recorded with a frequency consistent with the hemodynamic characteristics of the drug and the type of sedation, but at least every 5 minutes and preferably more frequently. Maximum and minimum HRs and BPs derived from continuous data are subject to artifacts; care needs to be taken to ensure accuracy. Other summary data can include time below or above preselected values (eg, area under the curve above or below a threshold). The results of preclinical studies should guide the need for specific invasive and/or provocative cardiovascular testing.

Gastrointestinal. While vomiting is an easily recognized AE, nausea is a patient-reported symptom and cannot be determined by an observer. Because of the impairments caused by the residual sedation and the sensation of nausea, a visual analog scale that requires the patient to physically demark the degree to which they sense nausea may be difficult for the patient to use. A verbal 11-point (0–10) numerical rating scale is a better instrument.³⁰ For children, the validated Baxter Animated Retching Faces scale³¹ may be utilized. Because nausea can be effectively treated, it would be inappropriate to deny a patient a rescue antiemetic.³² The time from the end of the procedure to the administration of an antiemetic and the nausea score when the patient asked for the antiemetic should be recorded. For some clinical trials, it may be appropriate to also record the duration of the nausea and the peak level. Consideration should be given to ensure that a sufficient number of high-risk³³ patients are included in clinical trials. If the protocol allows for prophylactic use of antiemetics, their use needs to be specified and controlled.

Neurological. Difficulty in achieving the desired level of sedation is perhaps the most common neurological adverse effect, but this would be reported as part of the assessment of efficacy in clinical trials.¹ Undersedation would most likely be manifest as patient dissatisfaction or inability to complete the procedure. Oversedation might be manifest by the appearance of effects on other organ system AEs (eg, ventilatory depression). Patient dissatisfaction with the sedation (eg, dysphoria or unpleasant dreams/recall, either intra- or postprocedure) would also be part of any efficacy clinical trial.¹ Similarly, prolonged recovery time is an important adverse effect that should be reported as part of measuring efficacy.

Rare, serious neurological adverse outcomes such as seizures should be classified as SAEs and would obviously be reported as part of any clinical trial. For general anesthesia, postoperative recall of intraoperative events is an important AE but may not be so troubling for procedural sedation if the proper patient preprocedure explanations and expectations are provided. Agitated or hypoactive delirium is an important AE after procedural sedation that needs to be explicitly elicited using a validated measure.³⁴

Other. Major adverse outcomes or events related to the liver or kidneys would be part of the toxicology assessments in phase 1 clinical trials. Because current agents used for procedural sedation can cause pain on injection, myalgia, and other patient-reported discomforts, patients should be asked about these and other possible symptoms as part of patient satisfaction surveys in efficacy clinical trials. The follow-up period should be sufficient, and the use of a symptom distress score can be useful.³⁵

CONCLUSIONS

Our definition for patient safety—“the avoidance of physical or psychological harm”—was also used by a previous consensus conference on pediatric procedural sedation,³⁶ but the definition does not translate directly to a design of procedural sedation clinical trials to test the safety of a new drug, protocol, or device. We distinguished among adverse effects, events, and outcomes. Adverse effects are the physiological consequences of the drugs or techniques used for the procedural sedation. The relationship between these adverse pharmacological effects and the occurrence of actual AEs in clinical practice that would require interventions to prevent actual harm is most often unknown (eg, the depression in the hypercapnic ventilatory response by opioids or bradycardia by dexmedetomidine).
Except for phase 1 trials, which are specifically designed to investigate possible adverse physiological effects of a new agent, premarking trials are primarily designed to ascertain efficacy and therefore might be underpowered to detect rare AEs and outcomes. However, there may be signals in this early phase 1 trial that would indicate the need to include specific at-risk patients in subsequent trials or to design pragmatic trials that utilize clinically relevant drug combinations. Our definition of adverse effects is related to safety biomarkers. The FDA has defined a safety biomarker as “measured before or after an exposure to a medical product … to indicate the likelihood, presence, or extent of toxicity as an adverse effect.”37 While the use of surrogate biomarkers may result in statistically more tractable measurements, the relationship between the surrogate biomarker (adverse effect) and actual harm (adverse outcome) is often unknown in procedural sedation.

We recommend that the criteria for an AE and the reason for any interventions, particularly when the intervention is not in response to a prespecified AE, should be carefully recorded and reported. An adverse outcome can be defined as an event that results in actual patient harm, which may be only temporary, or is an AE that requires an invasive intervention to prevent patient harm. These are not rigid definitions but rather should be considered suggestions. All clinical trials should clearly define criteria for classifying adverse effects, events, and outcomes, as well as the reasons for any clinical interventions. Classifications based on the World SIVA tool6 and/or the Quebec Guidelines13 provide excellent starting points. However, these tools are designed for QI purposes, and safety issues related to operator error or health system problems are not part of procedural clinical trials.

A major difficulty with designing procedural sedation clinical trials for the detection of AEs or outcomes is their rarity. From QI databases, overall AEs related to the sedation are approximately 5%,9,38,39 To have 80% power to show a relative risk reduction of 20% (eg, from 5% to 4%) would require a clinical trial of >10,000 patients. Obviously, specific AEs or outcomes are less frequent and would require even larger studies. Such large studies are not common in phase 3, and thus it is necessary to have reliable postmarketing surveillance and to consider the use of large simple trial designs for both phase 3 and 4 studies.40 For novel therapeutics approved between 2001 and 2010, such surveillance has resulted in a high incidence of postmarketing approval safety events.23 As more data become available from multiple studies, the use of meta-analysis may provide better estimates of the incidence of AEs, perhaps especially those that are relatively rare (eg, see recent Cochrane Library reviews on propofol45,49). Because the inclusion of studies into a meta-analysis is facilitated by commonality in the reporting of outcomes, we recommend continued attention to standardizing the recording and reporting of AEs.

A decade ago, Miner and Krauss42 emphasized that, “because the technique and goals of the operator affect the chance of an AE, evaluation of differences in complication rates between studies from different institutions has been difficult.” They stressed the need for clinical trials “with standardized sedation protocols, standardized AE reporting and standardized outcome measures to establish the true complication rate for each sedation level, drug, and procedure ….” Our recommendations differ from the previous study that emphasized QI data, in that we focus on the measurements that need to be made in prospective clinical trials. It is hoped that the recommendations listed in the Table and the previous SCEPTER recommendations on efficacy1 will help with the still-needed standardization of the design and reporting of procedural sedation clinical trials. Because a standardized or even a pragmatic trial design may result in detecting a lower rate of AEs than would be found in actual clinical practice, careful postmarketing surveillance will still be needed. It is hoped that improved preapproval clinical trials may help to reduce patient injury after regulatory approval.

**ACKNOWLEDGMENTS**

The authors thank Leah Crisafi, MD; Sharon Hertz, MD; Allison Lin, PharmD, PhD; Rigoberto Roca, MD; and Martha Van Cilef, MD, FDA Division of Anesthesia, Analgesia, and Addiction Products, Office of New Drugs; Bo Li, PhD, FDA Center for Drug Evaluation and Research, Office of Biostatistics; and Lydia I. Gilbert-McClain, MD, FCCP, FDA Division of Pulmonary, Allergy, and Rheumatology Products for their involvement in the meeting. The authors also thank Valerie Thompson and Andrea Speckin, Innovations Consulting Group LLC, Washington, DC, for their invaluable assistance in organizing the meeting and facilitating distribution of the manuscript among the authors.

From the *Department of Anesthesiology and Perioperative Medicine, University of Rochester School of Medicine and Dentistry, Rochester, New York; †Department of Anesthesiology, Tufts School of Medicine, Boston, Massachusetts; ‡Department of Anesthesiology and Perioperative Medicine, University of Rochester School of Medicine and Dentistry, Rochester, New York; ††Department of Pediatrics, University of Louisville School of Medicine, Norton Children’s Hospital, Louisville, Kentucky; †††Department of Pediatrics, Children’s Hospital of Eastern Ontario, University of Ottawa, Ottawa, Ontario, Canada; †‡University of Southern Illinois University School of Medicine, Springfield, Illinois; †§Department of Pediatrics, St John’s Children’s Hospital, Springfield, Illinois; †||Department of Anesthesiology, University of Colorado School of Medicine, Denver, Colorado; †|||Department of Anesthesiology and Critical Care Medicine, Boston Children’s Hospital and Harvard Medical School, Boston, Massachusetts; ‡‡Department of Anesthesiology, Leiden University Medical Center, Leiden, the Netherlands; ‡‡‡Division of Management Consulting, Department of Anesthesiology, The University of Iowa, Iowa City, Iowa; """Department of Pharmacology and Foundational Sciences, East Carolina University, Greenville, North Carolina; §§§Department of Anesthesiology, University of Rochester School of Medicine and Dentistry, Rochester, New York; §§§§Department of Anesthesiology, Stony Brook University, Stony Brook, New York; §§§§Division of Anesthesiology and Critical Care Medicine, Hadassah University Hospital, The Hebrew University of Jerusalem School of Medicine, Jerusalem, Israel; §§§||Department of Emergency Medicine, Loma Linda University, Loma Linda, California; §§§|||Department of Anesthesiology, University of Hong Kong, Hong Kong, China; §§§‡Department of Anesthesiology and Perioperative Medicine, University of Rochester School of Medicine and Dentistry, Rochester, New York; """"""""Department of Medicine, University of Pennsylvania Health System, Philadelphia, Pennsylvania; """"""""Department of Anesthesiology, John R. Oishei Children’s Hospital Jacobs School of Medicine and Biomedical Science, University at Buffalo, New York; §§§§§Department of Pediatrics, University of Massachusetts Memorial Children’s Medical Center, University of Massachusetts Medical School, Worcester, Massachusetts; §§§§§Department of Anesthesiology & Critical Care, The Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania; §§§§§§Department of Emergency, University of Minnesota Medical School, Minneapolis, Minnesota; §§§§§§Department of Emergency Medicine, Hennepin County Medical Center, Minneapolis, Minnesota; §§§§§§Department of Emergency Medicine, University of Virginia School of Medicine, Charlottesville, Virginia; """"""""Department of Anesthesiology, Vanderbilt University Medical School.
DISCLOSURES
Name: Denham S. Ward, MD, PhD.
Contribution: This author helped by organizing the meeting agenda, attending and presenting at the meeting, and writing the manuscript.
Name: Mark R. Williams, MBBS, BSc.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: John W. Berkenbosch, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Maala Bhatt, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Randall M. Clark, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Isabelle Constant, MD, PhD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Aaron Conway, RN, PhD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Harry L. Craver, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Albert Dahan, MD, PhD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Franklin Dexter, MD, PhD.
Contribution: This author helped by attending and revising the manuscript.
Name: Raymond Dionne, DDS, PhD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Robert H. Dworkin, PhD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Tong J. Gan, MBA, MHS, FRCA.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: David Gozal, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Steven Green, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Michael G. Irwin, MBchB, MD.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Suzanne Karan, MD.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Richard C. Prielipp, MD.
Contribution: This author helped by handling the manuscript.
Name: Richard D. Urman, MD, MBA, CPE.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Ronald E. O’Connor, MD, MPH.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Pratik Pandharipande, MD.
Contribution: This author helped by attending and revising the manuscript.
Name: Joseph R. Tobin, MD.
Contribution: This author helped by attending and revising the manuscript.
Name: Dennis C. Turk, PhD.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Rebecca S. Twerksy, MD, MPH.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Richard D. Urman, MD, MBA, CPE.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Anna Zhao-Wong, MD, PhD.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Hannah Wunsch, MD, MSc.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Mark Weiss, MD.
Contribution: This author helped by attending and revising the manuscript.
Name: Keira P. Mason, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: James Miner, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Maala Bhatt, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Joseph R. Tobin, MD.
Contribution: This author helped by attending and revising the manuscript.
Name: Robert E. O’Connor, MD, MPH.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Pratik Pandharipande, MD.
Contribution: This author helped by attending and revising the manuscript.
Name: Richard R. Riker, MD.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Mark G. Roback, MD.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Daniel I. Sessler, MD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Joseph R. Tobin, MD.
Contribution: This author helped by attending and revising the manuscript.
Name: Dennis C. Turk, PhD.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Rebecca S. Twerksy, MD, MPH.
Contribution: This author helped by attending the meeting and revising the manuscript.
Name: Richard D. Urman, MD, MBA, CPE.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
Name: Anna Zhao-Wong, MD, PhD.
Contribution: This author helped by attending and presenting at the meeting and revising the manuscript.
This manuscript was handled by: Richard C. Prielipp, MD.

REFERENCES