**Supplental Appendix 1.**

**PADIS Clinical Practice Guideline Methods**

1. **Panel membership**

The American College of Critical Care Medicine (ACCM) chose the chair and vice-chair based on their experience with the PAD 2013 clinical practice guidelines consulted them in the choice of panel members based on their expertise in the five specific focus areas (1). The mix of experts included involvement in the PAD 2013 guidelines (50% of panel) and new members (50% of the panel), with the aim to represent the multidisciplinary nature of professions relevant to ICU practice (1). We sought to ensure a multi-national profile of panel members by engaging researchers and clinical experts in each of the focus topics from the North America, Europe and Australia (Table E1). Critical illness survivors were chosen by the PADIS leadership based on their experience in institutional ICU patient and family committees and/or prior guideline efforts. Subcommittees (i.e. sections) were assigned one of the five subtopic areas: pain (and analgesia), agitation (and sedation), delirium, immobility (i.e., rehabilitation and mobilization), and sleep disruption. An experienced medical librarian designed and conducted all literature searches. Each section independently worked to develop relevant questions and related outcomes; identified, reviewed and evaluated the literature; crafted statements and recommendations and drafted their section of the guideline document. The methodology team (GRADE experts) was chosen by ACCM and was composed of practicing adult intensivists with extensive experience with clinical practice guideline development using the GRADE process. With one methodologist assigned to each of the 5 sections, their mandate was to provide methodological support and expertise; specifically, they participated as GRADE educators, advisors and reviewers

1. **Membership communication**

A guideline steering group (i.e., chair, vice-chair, section leaders, methodologists, librarian, and SCCM staff) held monthly conference calls to develop and finalize methods, discuss and resolve barriers to completion, develop the agenda for live meetings, and structure the guideline publication. Web-based and telephone-based communication within the panel sections occurred at least monthly and was attended by the chair, vice chair or both, the critical illness survivor and the methodologist. Web-enabled training calls were held for all members on topics such as use of the referencing software (i.e., RefWorks), how to conduct risk of bias (ROB) evaluations for studies, and use of the GDT system. Additional PADIS CPG leadership calls were held to discuss methods, timelines, potential content overlap between sections and any current issues or barriers. Members were invited to attend live PADIS meetings at the annual SCCM congress [Phoenix, AZ (January, 2015); Orlando, FL (February 2016) and Honolulu, HI (January 2017)].

Additional discussions between members occurred at other critical care meetings [e.g. American Thoracic Society (San Francisco, CA, May 2016; Washington, DC, May 2017) and European Society of Intensive Care Medicine (Milan, Italy, October 2016)]. These live meetings allowed for face-to-face communication, methodological training and review of encountered challenges, methodological updates, and resolution of potentially overlapping topic content.

1. **Conflict of interest management**

All PADIS members completed and submitted a SCCM conflict of interest form at least annually to characterize potential financial conflicts among panel members and their families. Members attending a live meeting were asked to report whether they had any new conflict(s). If they indicated they did, the member filed an updated conflict of interest form with SCCM. Within sections, members with eitherintellectual or financial conflicts of interest refrained, where possible, from serving as the primary reviewer for a question or participating in discussions surrounding the evidence profile or the Evidence to Decision (EtoD) framework (i.e., recommendation development) process. Prior to the final voting, a new round of written conflict of interest forms were obtained from all members. The guideline Chair, in conjunction with section leaders, identified those members (that could include the section leader) who would be required to abstain from voting on certain questions (but could still be consulted to help interpret evidence). These included (but were not limited) to section members whose important contributions to the literature drove the evidence supporting that question or the section member receiving funding (e.g. consulting, research, or both) from an organization associated in any way to an intervention or product related to the recommendation.

1. **Clinical question development**

*Selection of topics*

Panel members in each of the five sections developed a list of topics they considered most relevant to their assigned area of critical care. Statements and recommendations from the PAD 2013 guidelines, the current literature, panel members’ current clinical practice, and research were considered. Topics were specific critical care focus areas (e.g., prevention of delirium, treatment of postoperative pain) and were listed based on the perceived importance to clinicians and patients.

Each group defined, discussed, and clarified the topics, which were then posted in an online survey. Members in each section ranked them with a score from 1 (very low importance) to 5 very high importance); ratings were tallied across sections. Section members, the critical illness survivors, the methodologist, and the chair or vice-chair then participated in a call to discuss the topic ranking. During this call, the critical illness survivor was educated about each topic and encouraged to provide input about how they would rank the topics from their perspective. After this discussion, which sometimes led to a re-prioritization of topics, the topics ranked in the top 50% for each group were considered as possible clinical questions.

*Question development*

Subsequently, each section developed a list of questions using the prioritized topics. Panel member preference and endorsement, critical illness survivor buy-in, and feasibility issues helped drive final question choice. Draft questions were formulated into either an actionable question [i.e., in a PICO (Population, Intervention, Comparison, Outcomes) format that had the potential to lead to a recommendation] or a descriptive question (that would lead to a non-actionable statement). The development of actionable questions (i.e., “do this or don’t do this as compared to XX”) over description questions was prioritized. Guideline leadership discussions expanded relevant topics to more than one group where applicable; for example, the agitation (sedation) section (versus the delirium section) considered the burden of delirium as an outcome when comparing intravenous sedative medication options. Once established, the question was considered immutable. We allowed minor revisions to a research question to improve clarity or address specific sub-questions, if suggested by a section group based on compelling arguments, endorsed by their methodologist and further approved by the chair and vice-chair. A posteriori question removal or the development of new questions after recommendation development had started was not permitted.

1. **Outcome prioritization**

For each question, section members developed a list of relevant outcomes. To focus on relevance to the patient, we used the GRADE approach for outcome prioritization. Outcome(s) that were critical for decision-making were selected. Each member rated each outcome on a scale of 1 to 9 (1= least important; 9 = critical to decision making). We calculated the mean scores then categorized each outcome as: “not important” (score = 1-3); “important” (4-6); or critical (7-9) and considered all critical and important outcomes, regardless of the anticipated availability (e.g., studies addressing the outcome) or anticipated quality (e.g., sparse data) of the relevant data. Each section’s critical illness survivor reviewed the outcome list during a routine section call; they were asked to specifically focus on an important outcome that might be missing and the harmony between the outcomes prioritized by the groups and their perceived value to patients. The final outcome ranking combined these decision-making steps for each question. Once each section finalized their clinical questions, the PADIS guideline leadership ensured that no duplication across sections existed. All guideline participants (including the critical illness survivors) agreed upon the final lists.

We established a pragmatic definition for each outcome *a priori*. Guideline members primarily relied on a literature review to define each outcome. When heterogeneous definitions for an outcome were identified, group members discussed the one that they would retain through discussion. For example, for an outcome like “light sedation”, that crossed more than one guideline group, and where multiple validated sedation scales exist, this process engendered considerable discussion, and fostered evidence gap identification. The most appropriate time frame to evaluate an outcome was also discussed. For example, the immobility section attributed less importance to longer vs. shorter mortality since causality with an ICU-based rehabilitation/mobility intervention would become harder to establish given additional events after discharge that may affect mortality. In contrast, the agitation (sedation) section favored longer-term mortality based on the increasing focus on longer-term, post-ICU cognitive and mortality outcomes, after sedative drug exposure in the critical care setting.

**6. Literature search and guideline reference database**

A University-based medical librarian facilitated the literature review by creating an electronic, Web-based, password-protected, database using RefWorks software (Bethesda, MD). The relevant evidence was identified with controlled vocabulary as appropriate for each database, combined with additional relevant keyword terms from each section, and keyword terms in CPG efforts from other organizations that were felt to be relevant to the scope of the PADIS CPG. The search strategy is outlined in Supplemental Table 5 (Supplemental Digital Content 7, http://links.lww.com/CCM/D765). Five electronic databases were included in all literature searches: PubMed (1990-) via PubMed.gov, EMBASE (1990-) via Embase.com, The Cochrane Library which includes the Cochrane Database of Systematic Reviews and the Central Register of Controlled Trials via cochranelibrary.com, CINAHL (1990-) via EBSCO, and Web of Science (1990-) via Thomson Reuters. The results were added to the RefWorks, where sections then reviewed for eligibility based on written inclusion and exclusion criteria. All publications based on original research in adult humans (≥ 18 years) published from 1990 to present were considered. This is the same start date used in the PAD 2013 guidelines (1).

There was no limitation on language. Non-English studies were considered if a section considered the study to be an important part of the evidence profile for one of their questions. In these instances, native language speakers translated only the methods and results section of the paper. We limited all our literature reviews to capturing studies examining the ‘critical care’ environment, defined for guideline purposes as any environment capable of providing mechanical ventilation and invasive hemodynamic monitoring that was not a post-operative recovery room, excluding studies primarily performed in centers where long-term mechanical ventilation and weaning might take place (e.g., a long term acute care hospital (LTACH). Evidence searches focused on the highest quality evidence available per outcome and per question in keeping with GRADE guidance. Case series were not included unless a large case series was the only source of literature for a particular question/particular outcome or an appropriate form of evidence for a specific outcome (e.g. safety). Any study with a sample size ≤ 10 was excluded.

Repeated literature search updates were conducted across all databases on a monthly basis by the medical librarian from Sept 1, 2014 until October 1, 2015 using the established search strategy. Updates were run by section (i.e., pain, agitation/sedation, delirium, rehabilitation/mobility, and sleep), de-duplicated, distributed via email, and integrated in the RefWorks group account where they were reviewed by members for inclusion. If a publication was considered to significantly impact one of the questions addressed by the group (i.e., potentially alter the recommendation) after October 2015 it was incorporated into the summary of evidence and evidence profile. Over 41,000 references were ultimately included in the study database.

**7. Reference Screening, data abstraction and risk of bias assessment**

Screening was performed for each question individually. Section members with a financial or intellectual conflict of interest avoided serving as reviewer for that question. An initial search of the titles and abstracts of all articles in each question folder was completed by a single reviewer. Full text screening was then performed in duplicate. Primary and secondary reviewers discussed all discordantly rated articles with the methodologist in an effort to clarify differences in interpretation; consensus was not mandated but agreement as to question inclusion occurred across all sections. Team leaders, other section members and the vice-chair and chair engaged in discussion as to relevance and content.

Once the articles addressing a specific question were finalized, a single reviewer extracted all relevant data and entered it into a pre-formatted data abstraction form. Team leaders, other section members and methodologists participated in discussion as to interpretation around specific data extraction issues. The secondary member for the question then reviewed the database that contained the data extracted by the primary reviewer for accuracy and completeness.

The primary reviewer for each question completed a risk of bias assessment for eligible RCTs using the Cochrane Risk of Bias tool [(2)](https://paperpile.com/c/p2mHuz/Ebaf8) risk of bias assessment was completed for each outcome and for each eligible RCT. Reviewers used the Newcastle-Ottawa scale [(3)](https://paperpile.com/c/p2mHuz/mgHII) to assess risk of bias for non-randomized trials. The section methodologist provided support for risk of bias assessment whenever necessary.

8**. Summarizing the evidence**

Across all groups, the methodologists worked closely and collaboratively with the primary question reviewer in producing evidence summaries. For actionable questions, we generated weighted pooled estimates using meta-analytic techniques when possible. We used random-effects model and the method of DerSimonian and Laird to pool the estimates of effect across eligible studies. We used fixed effect model as a sensitivity analysis; if there is discrepancy in the results we used the estimates generated by random-effects model except in the following situations: (a) a large trial with opposing results, or (b) when less than three studies were included in the analysis [(4)](https://paperpile.com/c/p2mHuz/u8gca). We used risk ratio (RR) or odds ratio (OR) with accompanying 95% confidence interval (CI) to summarize the treatment effect for dichotomous outcomes, and mean difference with 95% CI to summarize the treatment effect for continuous outcomes. When different scales are used to measure a continuous outcome, we used standardized mean difference (SMD) with 95% CI to summarize the treatment effect. We used Chi2 and I2 tests to assess for statistical heterogeneity. A p-value threshold of p < 0.1 or I2 value > 50%, respectively, were considered significant statistical heterogeneity. We assessed for presence of publication bias by inspecting funnel plots when 10 or more studies were included. Meta-analysis was conducted using the Cochrane Collaboration Review Manager version 5.3 software [(5)](https://paperpile.com/c/p2mHuz/bTVjA). When quantitative meta-analysis was not possible, the case for some actionable questions and all descriptive questions, evidence was synthesized qualitatively in narrative form.

**9. Certainty of evidence**

Across the five sections, the overall certainty of evidence (also known as quality of evidence or confidence in effect estimates) for each outcome was assessed following the GRADE approach and based on the following criteria: risk of bias, precision, consistency, directness of the evidence, risk of publication bias, presence of a dose-effect relationship, magnitude of effect, and assessment of the effect of plausible residual confounding or bias (Table E3). The certainty for each outcome was categorized as high, moderate, low or very low based on these assessments. Methodologists summarized the quality of evidence assessment in evidence profile format using GDT software (www.GRADEPRO.org). Each evidence profile was reviewed in multiple steps by the primary question reviewer, all section members, and the vice chair and chair before being considered final. The process by which descriptive questions are addressed is described in the methods paper.

**10. Evidence to Decision Framework and Recommendation Generation**

The Evidence to Decision (EtoD) framework was used to help organize discussion around each potential recommendation and ensure each of the following domains/components was considered [(6)](https://paperpile.com/c/p2mHuz/AK68G):

a. Priority of the problem. For the purpose of this guideline, all questions addressed important clinical problems.

b. Magnitude of the desirable anticipated effects. The larger the desirable outcomes, the more inclined we were to recommend the intervention.

c. Magnitude of the undesirable anticipated effects. The larger the undesirable consequences, the less inclined we were to recommend the intervention.

d. Certainty of evidence. The higher the quality of evidence the more likely we were to issue a strong recommendation.

e. Values and preferences. The more variability in values and preferences, or more uncertainty in values and preferences, the more likely we were to issue a conditional (i.e., weak) recommendation.

f. Balance between desirable and undesirable effects. When desirable consequences (desirable treatment effect, consistency in values and preferences) clearly outweigh the undesirable consequences (undesirable treatment effect, large variability in values and preferences) or vice-versa, we were more inclined to issue a strong recommendation. When the balance was less clear or undetermined, then we were more inclined to issue a conditional recommendation or a neutral recommendation.

g. Costs/resource allocation. The higher the costs of an intervention (that is, the more resources consumed) the less likely we were to issue a strong recommendation.

h. Acceptability to stakeholders. The more acceptable or feasible to stakeholders; the more likely we were to issue a strong recommendation.

i. Feasibility to implement. The more feasible to stakeholders; the more likely we were to issue a strong recommendation.

The primary reviewer for the question, with input from the methodologist, made preliminary judgments around these domains, and then they were discussed during a conference call that included section members, the chair and vice-chair of the guidelines, the methodologist and a critical illness survivor. Each call used Webex technology (Zoom, San Jose, CA) so that all members on the call could view the EtoD framework in real time as each domain was being discussed.

The strength of a recommendation for each actionable question, was defined as either strong or conditional, and either for or against an intervention, based on both the quality of evidence and the risks and benefits across all critical and important outcomes considered. A strong recommendation, either in favor or against an intervention, implied that the majority of task force members (where the input of the critical illness survivors was considered) believed that that the benefits of an intervention significantly outweighed the risks (or vice versa) and that the majority of patients and clinicians would and should pursue this course of action (or not). A conditional recommendation either in favor of or against an intervention implied that the benefits of the intervention likely outweighed the risks (or vice versa), but that task force members, and the critical illness survivor, were not confident as to the quality of the evidence or because the trade-offs between risks and benefits were closely balanced. Throughout the guidelines, for all strong recommendations, the phrase “we recommend…” was used, and for all conditional recommendations, “we suggest….” was used.

Through conference call(s) discussion, all non-conflicted section members came to a consensus on: i) the specific wording of each of their section’s recommendations and statements, ii) the defined strength of evidence for each recommendation (ie., high, moderate, low, or very low), iii) the strength of each recommendation (i.e. strong vs conditional), and iv) the direction of each recommendation (i.e. for or against) as it was being proposed in the GDT system.

**14. Good Practice Statements**

In situations where a large body of indirect evidence exists, without reasonable comparators, when evidence was difficult to summarize but shows clear benefit with no or very minimal harm, we considered issuing a good practice statement (7). These statements reflect a strong but ungraded recommendations, in which there is an unequivocal believe that the benefit of the intervention outweighs the risk but no available direct evidence that could be summarized or evaluated. A good practice statement had to meet all the following criteria: a) the statement was clear and actionable; b) the message was necessary for clinicians; c) the net benefit was large and unequivocal; d) the evidence was difficult to collect and summarize; e) the rationale was explicit; and f) it was not feasible to use GRADE methodology to assess the quality of evidence.

**15. Final Voting**

Once each group finalized its draft recommendations, a web-based survey system (Qualtrics, Provo, UT) was used to conduct and manage recommendation voting for that group. The survey was developed by the guideline chair and reviewed by the group leader and methodologist before distribution. Consistent with most practice guidelines, methodologists did not. Critical illness survivors did not vote given that the PADIS leadership felt they had neither the expertise nor clinical perspective to make a fully informed voting decision. Guideline voting results were anonymous except to the guideline leadership. Voting reminders were sent to guideline members until all members had voted. Voting did not occur for descriptive questions.

The poll for each recommendation included the question, the draft recommendation (including is strength and level of evidence), clarifying remarks about the recommendation (where pertinent), web links to the forest plots, evidence summary, and a brief rationale, a voting response choice (i.e. yes, no or abstain), and two text boxes (comments on the recommendation and comments on potential evidence gaps). As with other SCCM practice guidelines, at least 70% of guideline members had to vote in order for a decision on the recommendation to be made. A recommendation passed when 80% of the voting members (excluding those who abstained) voted to accept the recommendation [(8)](https://paperpile.com/c/p2mHuz/tasUR). Members with a conflict of interest for a particular question were required to abstain from voting for this question.

**References**

1. [Barr J, Fraser GL, Puntillo K, et al.: Clinical practice guidelines for the management of pain, agitation, and delirium in adult patients in the intensive care unit. *Crit Care Med* 2013; 41:263–306](http://paperpile.com/b/o91TtG/yymd2)

2. [Higgins JPT, Altman DG, Gøtzsche PC, et al.: The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. *BMJ* 2011; 343:d5928](http://paperpile.com/b/p2mHuz/Ebaf8)

3. [Sanderson S, Tatt ID, Higgins JPT: Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007; 36:666–676](http://paperpile.com/b/p2mHuz/mgHII)

4. [Guyatt GH, Norris SL, Schulman S, et al.: Methodology for the development of antithrombotic therapy and prevention of thrombosis guidelines: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest* 2012; 141:53S–70S](http://paperpile.com/b/p2mHuz/u8gca)

5. [RevMan 5 | Cochrane Community [Internet]. [cited 2017 Oct 2] Available from:](http://paperpile.com/b/p2mHuz/bTVjA) <http://community.cochrane.org/tools/review-production-tools/revman-5>

6. [Alonso-Coello P, Oxman AD, Moberg J, et al.: GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well informed healthcare choices. 2: Clinical practice guidelines. *BMJ* 2016; 353:i2089](http://paperpile.com/b/p2mHuz/AK68G)

7. Guyatt GH, Schunemann HJ, Djulbegovic B, Akl EA. Guideline panels should not GRADE good practice statements: *J Clin Epidem* 2015; 68:597-600

8. [Rhodes A, Evans LE, Alhazzani W, et al.: Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. *Crit Care Med* 2017; 45:486–552](http://paperpile.com/b/p2mHuz/tasUR)