Inconsistent relationship between depth of sedation and intensive care outcome: systematic review and meta-analysis

Aitken, L. M., Kydonaki, K., Blackwood, B., Trahair, L. G., Purssell, E., Sekhon, M., & Walsh, T. S. (2021). Inconsistent relationship between depth of sedation and intensive care outcome: systematic review and meta-analysis. *Thorax*. Advance online publication. https://doi.org/10.1136/thoraxjnl-2020-216098

Published in:
Thorax

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Title: Inconsistent relationship between depth of sedation and intensive care outcome: systematic review and meta-analysis

Authors:

Leanne M Aitken, PhD, Professor of Critical Care, School of Health Sciences, City, University of London, London, UK and School of Nursing and Midwifery, Griffith University, Brisbane, Australia. ORCID: 0000-0001-5722-9090

Leanne.aitken.1@city.ac.uk Phone: +44 (0) 20 7040 5968

Kalliopi Kydonaki, PhD, Lecturer, School of Health and Social Care, Edinburgh Napier University, Edinburgh, UK. ORCHID: 0000-0003-0607-7618

C.Kydonaki@napier.ac.uk

Bronagh Blackwood, PhD, Professor in Critical Care, Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, Northern Ireland, UK.

b.blackwood@qub.ac.uk

Laurence G Trahair, PhD, Adelaide Medical School, The University of Adelaide, Adelaide, Australia, ORCID: 0000-0002-3479-8490

trahair@internode.on.net

Edward Purssell, PhD, Senior Lecturer, School of Health Sciences, City, University of London, London, UK

Edward.Purssell@city.ac.uk

Mandeep Sekhon, PhD, Department of Population Health Sciences, School of Population Health & Environmental Sciences, Faculty of Life Sciences and Medicine, King's College London, London, UK

mandeep.sekhon@kcl.ac.uk

Timothy Simon Walsh, MD, Department of Anaesthesia, Critical Care & Pain Medicine and Usher Institute, University of Edinburgh, Edinburgh, UK ORCID: 0000-0002-3590-8540

Timothy.Walsh@ed.ac.uk

Corresponding Author: Leanne M Aitken, Professor of Critical Care, School of Health Sciences, City, University of London, Northampton Square, London, EC1V 0HB, UK Leanne.aitken.1@city.ac.uk Phone: +44 (0) 20 7040 5968

Authors' contributions to manuscript: Review concept and design – LMA, KK, BB, TSW; Study identification, critical appraisal and data extraction – LMA, KK, BB, LGT, MS; Statistical analysis – EP; Interpretation of analysis – all authors; drafting of manuscript – LMA; critical revision of the manuscript for important intellectual content – all authors.

Financial support: No specific funding was received for the conduct of this review.

Keywords: critical care, sedation, deep sedation, systematic review, patient outcome assessment, critical care outcomes

Word Count: 4008

ABSTRACT

Purpose: To determine the effect of depth of sedation on intensive care mortality, duration of mechanical ventilation, and other clinically important outcomes.

Methods: We searched MEDLINE, Embase, CENTRAL, CINAHL, PsycINFO from 2000 - 2020. Randomised controlled trials and cohort studies that examined the effect of sedation depth were included. Two reviewers independently screened, selected articles, extracted data and appraised quality. Data on study design, population, setting, patient characteristics, study interventions, depth of sedation and relevant outcomes were extracted. Quality was assessed using Critical Appraisal Skills Programme tools.

Results: We included data from 26 studies (n=7865 patients): 8 RCTs and 18 cohort studies. Heterogeneity of studies was substantial. There was no significant effect of lighter sedation on intensive care mortality. Lighter sedation did not affect duration of mechanical ventilation in RCTs (mean difference [MD]: -1.44 days [95% CI -3.79 to 0.91]) but did in cohort studies (MD: -1.54 days [95% CI -2.68 to -0.39]). No statistically significant benefit of lighter sedation was identified in RCTs. In cohort studies lighter sedation improved time to extubation, intensive care and hospital length of stay and Ventilator Associated Pneumonia. We found no significant effects for hospital mortality, delirium or adverse events.

Conclusion: Evidence of benefit from lighter sedation is limited, with inconsistency between observational and randomised studies. Positive effects were mainly limited to low quality evidence from observational studies, which could be attributable to bias and confounding factors.

KEY MESSAGES

What is the key question?

Does depth of sedation effect intensive care mortality and duration of mechanical ventilation, as well as secondary physiological, hospital mortality, resource use, adverse event and life impact outcomes?

What is the bottom line?

Evidence of the effect of sedation depth is limited, with inconsistency between observational and randomised studies. Positive effects from lighter sedation were mainly limited to low to very low quality evidence from observational studies.

Why read on?

Depth of sedation appears to have differential effect on various outcomes. We need to build on the current evidence to determine how to optimise patient outcomes, both within and beyond intensive care.

INTRODUCTION

Mechanically ventilated patients in intensive care receive sedation and analgesia to manage their discomfort. Although these medications are considered important for many patients, there is recognition that both the amount and type of sedation that patients receive are potentially related to patient outcomes (1). Various proposals and guidelines recommend alternative ways of administering sedation or using different sedative agents to improve outcomes from critical illness (1-3). Although interpretation of this literature is challenging due to inconsistent and problematic definitions, evidence suggests lighter sedation is probably beneficial (1). Despite this, recent reports show many ICU patients worldwide continue to be deeply sedated (4-6).

In a recent review of outcomes associated with sedation depth in the first 48 hours of mechanical ventilation across the Emergency Department (ED) and Intensive Care Unit (ICU) lighter sedation was associated with reduced mortality, mechanical ventilation and ICU stay days (7). Given many critically ill patients remain heavily sedated for longer than 48 hours, it would be useful to know if this relationship between sedation depth and patient outcomes extends across patients' entire ICU stay and relates to a range of patient outcomes or only the short term outcomes of mortality and duration of mechanical ventilation and ICU stay. The effect of lighter sedation on selected outcomes was also examined in the PADIS guidelines, however the included meta-analysis incorporated only studies where sedation depth was defined *a priori* (1), with inconsistent evidence identified. These reviews provide some insights into the evidence to guide sedation practice, but both reviews focused on specific subgroups of studies. We therefore considered a review of a wider range of relevant studies appropriate and important.

Objective

To systematically examine the effect of depth of sedation in ICU patients on patient outcomes that extend across the ICU stay and beyond. ICU mortality and duration of mechanical ventilation were co-primary outcomes selected because ICU mortality is patient-focused and duration of mechanical ventilation reflects sedation practice. Secondary outcomes from the five domains of the outcome taxonomy proposed by Dodd and colleagues (8) were selected and included hospital mortality, physiological outcomes (time to extubation, ventilator free days (Vfd) to day 28), resource use (ICU

and hospital length of stay), adverse events (incidence of delirium, self-extubation, reintubation and tracheostomy, ventilator associated pneumonia (VAP)) and life impact outcomes (memories, anxiety, depression and symptoms or diagnosis of PTSD); these latter outcomes mirror those identified as important to patients and family members in a research priority setting exercise (9).

METHODS

The protocol for this systematic review was registered on PROSPERO (CRD42018092554; www.crd.york.ac.uk/prospero/display_record.php?RecordID=92554). Additional detail is available in supplementary materials.

Search strategy

MEDLINE, Embase, Cochrane Register of Controlled Trials (CENTRAL), Cumulative Index to Nursing and Allied Health Literature (CINAHL) and PsycINFO were searched with the following strategy: (intensive care OR critical care OR critically ill) AND (sedat* OR midazolam OR propofol) AND (length of stay OR mortality OR outcome assessment OR physical function OR psychological OR cognitive OR memories).

We searched for publications reporting randomized controlled, quasi-experimental and before-after trials, and cohort studies (prospective and retrospective) published in English between January 2000 and February 2020.

Types of participants

We included studies in adult patients receiving invasive mechanical ventilation in ICU, including patients who commenced their ventilation in another location, e.g. ED, operating room. We excluded studies: (i) in patients receiving non-invasive ventilation and mechanically ventilated patients not admitted to ICU; (ii) where the intervention included different sedative agents. Studies testing the effect of different sedative agents were excluded because it is not possible to determine if any difference in outcome was due to effect of the different agent or different depth of sedation. We defined our exposure as deeper sedation at any time throughout the period of mechanical ventilation in the ICU. Our classification of depth of sedation as either ‘lighter’ or ‘deeper’ did not need to be (but could be) predefined by study authors, but was based on published information incorporating any objective measures of sedation depth including assessment using a validated sedation assessment

instrument, hourly or daily doses of sedatives. To clarify, studies that tested any intervention (e.g. goal or protocol directed sedation, no sedation), other than different sedative agents, were eligible for inclusion if one group of patients received lighter sedation than another group of patients in the study. Studies were not excluded on the basis of which sedative agent they used, and no attempt was made to control analgesic use, although it is recognised that many have a secondary sedative effect. Only the RASS and Riker Sedation-Agitation Scale (SAS) were accepted as validated instruments (10).

Study selection

Titles and abstracts were screened independently by two researchers, with full text of included studies reviewed by two authors to assess eligibility. Studies where separation of depth into 'lighter' and 'deeper' sedation could not be identified were excluded. Studies including >2 groups based on sedation depth were not included in the meta-analysis but were retained in the additional analyses. Sedation was defined as the use of pharmacological agents that have the primary purpose of calming or inducing sleep, and alternative agents such as analgesics were not included despite acknowledging that secondary effects of sedation are often present. We did not include different outcomes from the same patient cohort, reported in multiple papers, twice in any analysis but this relationship was noted.

Data Extraction

Two authors extracted data on study design, population and setting, patient characteristics, study interventions, measure of depth of sedation (methodology and results) and relevant outcomes.

Assessment of bias

The domains of bias for RCTs and cohort studies were assessed consistent with current guidance (11,12). Relevant confounding factors were not identified *a priori*, but were based on the study method and cohort and included demographic, clinical, and treatment variables with the potential to influence relevant outcomes. No studies were excluded on the basis of quality assessment.

Data Analysis

Two authors extracted data on study design, population and setting, patient characteristics, study interventions, measure of depth of sedation (methodology and results) and relevant outcomes. All studies that contained data suitable for inclusion in at least one meta-analysis were included in the

quantitative analysis. Continuous data were analysed as means and standard deviations. Where the median and inter-quartile range was reported, these were converted to mean and standard deviation using a standard method (13). Dichotomous data were analysed as risks and relative risks. Random effects meta-analyses were undertaken with the meta package (14) in R (15). This allowed for both within and between studies variance to be calculated, the latter being reflected in a statistical test of heterogeneity. Cohort studies and RCTs were analysed separately based on an *a priori* decision. The quality of evidence was rated using Grades of Recommendation, Assessment, Development and Evaluation (GRADE) (16). For outcomes where significant methodological differences occurred (e.g. different instruments or time points) results were combined descriptively.

Sensitivity analysis

Categorisation of patients into 'lighter' and 'deeper' sedation groups could be based on either a difference in RASS or SAS scores, a difference in average dose of sedation over time (hourly/daily) or a combination of both. Due to the potential differential effect of sedation amounts on patients' sedation levels, a *post hoc* decision was made to repeat meta-analyses incorporating only those studies where categorisation was based on RASS or SAS scores alone or in combination with sedation dose, i.e. to exclude studies where categorisation was based solely on sedation dose. Similarly, a post hoc analysis of cohort studies to examine the influence of the temporal nature of the design (i.e. prospective or retrospective) was conducted.

RESULTS

After removal of duplicates, 3390 articles were identified (Figure 1), with full text of 116 articles assessed. Ninety were excluded: 69 met exclusion criteria; and 21 because, although patients were in groups, levels of sedation did not differ between the groups.

Twenty-six articles reporting the results of 23 studies incorporating 8575 patients remained for descriptive synthesis with 17 articles (7027 patients) included in a meta-analysis for at least one outcome (17-33). The included papers reported results prospective (n=16; n=5534) and retrospective (n=2; n=2028) cohort studies and randomised controlled trials (n=8; n=1534) published between 2001 and 2020 conducted across Asia (n=2), Australia and New Zealand (n=3), Europe (n=10), Middle-east (n=1) and North (n=7) and South America (n=4) (Table S1). Depth of sedation was measured either

using sedation assessment instruments or average doses of sedatives or a combination of both (Table 1, Table S2). The level of sedation that constituted 'lighter' or 'deeper' sedation was inconsistent across studies.

Risk of bias was highly variable in the cohort studies. In the RCTs risk of bias was more consistent, with lack of blinding being the main source of bias. Blinding of participants and personnel was not possible and blinding of outcome assessors was rare (Figure S1, Table S3). There was infrequent incorporation of relevant confounding factors into analysis in cohort studies (Figure S2, Table S4). Included studies addressed both our primary outcomes, and secondary outcomes within the five domains of mortality, physiological outcomes, resource use, adverse events and life impact outcomes (8) (Table S5), with most outcomes assessed in meta-analyses (Table 2). Outcomes within the life impact domain could not be pooled, but a descriptive synthesis of results related to memory and psychological function is provided (Table 3). Studies not included in the meta-analyses are synthesised under Additional Analyses.

Primary Outcomes

ICU Mortality: When comparing lighter versus deeper sedation we found no difference in ICU mortality in either RCTs or cohort studies (Table 2, Figure 2).

Duration of mechanical ventilation: We found no difference in duration of mechanical ventilation in the RCTs comparing lighter versus deeper sedation, but identified reduced duration of mechanical ventilation with lighter sedation in cohort studies (MD -1.54 days [95% CI:-2.68 to -0.39], I^2=87%, 8 studies, 3304 participants) (Table 2, Figure 2).

Secondary Outcomes

***Hospital Mortality*:** Pooled data from 5 RCTs and 5 cohort studies showed no difference between lighter and deeper sedation on hospital mortality (Table 2, Figure S3).

Physiological outcomes: Pooled data from 4 RCTs and 2 cohort studies showed no difference between lighter and deeper sedation on 28-day Vfd (Table 2, Figure S4). There was no difference in time to extubation in a single RCT, but cohort studies reported reduced duration with lighter sedation (MD -3.77 days [95% CI:-5.49 to -2.06], I^2=98%, 2 studies, 2132 participants). Pooled data from 4

RCTs and 4 cohort studies showed no difference between lighter and deeper sedation on incidence of delirium (Table 2, Figure S4).

Resource Use: Pooled data from 6 RCTs showed no difference between lighter and deeper sedation on ICU LOS or hospital LOS, but a significant reduction favouring lighter sedation was identified in ICU and hospital LOS in cohort studies (8 and 6 studies respectively Table 2, Figure S5). Lighter sedation had no effect on frequency of tracheostomy (4 RCTs, 2 cohort studies; Table 2).

Adverse Events: We found no difference between lighter and deeper sedation on self-extubation (2 RCTs, 3 cohort studies) or reintubation (5 RCTs, 2 cohort studies) (Table 2, Figure S6). Lighter sedation had no effect on risk of VAP in 1 RCT, although data from 2 cohort studies showed a reduced risk with lighter sedation (RR 0.56 [95% CI:0.33 to 0.96], I^2=51%, 1906 participants) (Table 2, Figure S6).

Sensitivity analyses

Meta-analyses, incorporating only those studies where RASS or SAS data were available to categorise patients as lighter or deeper sedation, were repeated on outcomes where studies existed. Results were largely similar, although fewer significant differences were identified (Table S6).

Meta-analyses examining the influence of the temporal nature of the design in cohort studies, i.e. prospective or retrospective, were conducted. Results were largely similar to the overall results, although analysis of only the prospective studies substantially reduced the heterogeneity when examining ICU and hospital mortality and hospital length of stay but had no effect on heterogeneity in relation to other outcomes (Table S7).

Additional analyses

Nine studies met the inclusion criteria, but were excluded from all meta-analyses for reasons detailed in the Methods (34-42). The main reasons were single group cohort studies with multivariable regression analysis (35,39) or more than 2 groups of patients not able to be combined based on sedation depth (34,40), as well as variable time points and methods for outcome measurement. In addition, some studies (where the primary outcome has been incorporated in meta-analyses above) incorporated life impact outcomes as secondary measures, however differences in methods of

outcome assessment precluded a meta-analysis of life impact outcomes. A descriptive synthesis is provided here.

Mortality, physiological outcomes and adverse events: A positive relationship between deep sedation and increased mortality (35,39) and increased duration of MV (39,40) was reported in cohort studies, but depth of sedation was not associated with MV duration across different stages of implementation of a sedation protocol and education intervention (34). A relationship between deeper sedation and both delirium (39) and VAP (34) was identified.

Life Impact: Outcomes reflecting the impact of sedation depth on a person's life focused only on memories and psychological health measured in 10 studies using a variety of instruments at different times (Table 3). There was some evidence of a relationship between sedation depth and presence or type of memories that patients reported. In a cohort study of 128 Brazilian patients, those who received any sedation reported less real memories (21[24%] vs 29 [69%]), more illusionary memories (7[8%] vs 0) and more amnesia (16[19%] vs 4[10%]) than patients who received no sedation (40). In a cohort study of 313 Swedish patients increased time deeply sedated was associated with having no recall of ICU (odds ratio [OR]:1.60, 95% CI:1.35–1.91) (37). In further analysis of the same cohort, patients who spent more time awake were more likely to remember the endotracheal tube (OR:1.45, 95% CI:1.29–1.62) and be bothered by memories of stressful ICU experiences (OR:1.37, 95% CI:1.13–1.67), but sedation depth was not associated with nightmares during recovery (38). In contrast, in 289 patients in Canada and USA, patients with no recall of ICU received lower daily doses of midazolam (26.9 [SD 63.7] vs 82.5 [SD 314] mg), but delusional memories were not associated with higher sedative doses (OR:1.18, 95% CI:0.37–3.81) (41). No difference in frequency or type of memories was reported in 2 studies (27,36) or in studies exploring the relationship between psychological distress and sedation depth (18,24,30,40,42).

DISCUSSION

In this systematic review of data from 26 studies incorporating just under 8000 adult patients there was inconsistent and inadequate evidence of the relationship between sedation depth and patient outcomes. Moderate level evidence from RCTs was identified in relation to the primary outcomes of

ICU mortality and duration of mechanical ventilation, as well as secondary outcomes including hospital mortality, time to extubation, ventilator free days, ICU LOS, incidence of delirium and tracheostomies, however no benefit of lighter sedation was identified in any of these outcomes. Outcomes where benefit of lighter sedation was shown in cohort studies included duration of mechanical ventilation, time to extubation, ICU and hospital length of stay and VAP; the evidence was assessed as very low level for all these outcomes. Reasons for low levels of evidence were multi-factorial but included inconsistency and imprecision, frequently with very high levels of heterogeneity, likely occurring as a result of differences in the primary aim and design of included studies as well as variation in interventions used to achieve lighter sedation. The multi-dimensional nature of factors that influence each of the outcomes also likely influences the inconsistency in results. High levels of heterogeneity potentially occurred as a result of the different designs (RCTs as well as prospective and retrospective cohort studies), the intent of the project (e.g. primarily as a quality improvement project) and the level of sedation and intervention fidelity achieved. The heterogeneity shown in this review highlights the issue of sedation being a complex healthcare invention influenced by multiple factors including agent chose, patient characteristics, protocols and practices, contextual issues within ICUs and individual clinician values and beliefs. These issues increase the relevance of the possible uncertainty highlighted in our review.

There was little evidence of effect of sedation depth on life impact outcomes. There was no evidence that anxiety, depression or symptoms of post-traumatic stress were related to sedation depth (18,24,30,40,42). There was, however, inconsistent evidence of whether, and how, sedation depth might influence the presence and type of memories (18,27,36-38,40,41). The role of memories after critical illness, and the relationship with psychological health, is inconsistent, with some suggestion that intrusive, persecutory or delusional memories may be more harmful than real memories (43), with the possibility that more frightening memories might be associated with greater psychological trauma (44). No evidence of a relationship between sedation depth and delirium was identified in this review, however any potential relationship between sedation, delirium and memories requires further investigation (43).

Few of the included studies identified an *a priori* aim related to sedation depth. Instead, many studies examined the effect of interventions to improve sedation practice, or explored the relationship between sedation and outcomes. Labelling of groups as 'deeper' and 'lighter' sedation in this review may not be appropriate given that 'deeper' sedation in one study could be similar to 'lighter' sedation in another study or setting. For example, RASS -3 indicated moderate sedation in one study (40) and deep sedation in others (17,20), while one pre-post study achieved 'lighter' sedation with a median first RASS score of -4 post-intervention (17). No studies targeted RASS 0 to -1 (alert and calm to drowsy), with the exception of work from Scandinavia examining 'no sedation' (25,29,42). The diversity of clinical practice strategies to achieve lighter sedation also presented challenges. We aimed to summarise whether strategies, whatever their design or content, that targeted deeper sedation avoidance were effective in changing outcomes relative to the comparator.

Recently, a Peruvian multi-centre observational cohort study examining the relationship between benzodiazepine dose and mortality was published (45). In this study benzodiazepine dose was associated with a higher risk of mortality and a significant decrease in Vfd, although it should be noted that 98% of participants were deeply sedated at some point during the study and depth of sedation was assessed using either the Glasgow Coma Scale, Ramsay Sedation Scale or RASS. The primary results of the SPICE-III study comparing dexmedetomidine to usual sedation are also published (6). SPICE-III compared different sedatives and was therefore ineligible for this review. However, it is worth noting that although the dexmedetomidine group had a slightly higher proportion of patients with lighter RASS scores (56.6% vs 51.8%), no difference in outcomes was observed. In two French studies also not meeting our inclusion criteria, one multicentre study found no difference in Vfd or mortality with the introduction of an oversedation prevention strategy (46), while a single centre study found reduced duration of mechanical ventilation by stopping sedation immediately after ICU admission (47). The most recent relevant study published was the Danish NONSEDA study where a strategy of no sedation was compared to light sedation (25). In this high quality RCT with clear separation in sedation levels, a non-significant trend towards higher mortality in the non-sedated

group was identified, emphasising the need for a strong body of evidence to illuminate the effect of sedation depth on a range of patient outcomes.

The reasons for reporting the effects of sedation depth on clinical outcomes from cohort studies alongside those from RCTs deserves attention. Changing sedation practice frequently requires an integrated or bundled approach to sedation assessment and management to achieve cultural change of clinician behaviour (2,48). Cohort (before and after) studies are more amenable to achieve practice change than randomised studies. Once a shift in clinicians' sedation management behaviour has been learned, it can be difficult to apply earlier (usual care) practices when patients are randomised. The RCTs in this review all randomised at the patient level. So, although cohort studies provide lower quality evidence than RCTs, in the area of sedation practice they have provided a pragmatic method for studies designed to modify sedation depth. To improve the quality of evidence, we recommend cluster randomised trials to address the weakness of intervention contamination in patient level randomisation and improve the quality of evidence. We have also provided ratings of evidence using the GRADE criteria (16), although we note the limitations of this system in that it is based on subjective judgements and does not take into account the benefits of various study methodologies as outlined above.

There have been multiple calls in clinical guidelines and opinion papers for lighter sedation in ICU patients (1,2); these calls have been based on sub-sets of the available evidence (7) or individual studies (e.g. (28,49). In response to these calls, multiple strategies have been proposed to achieve lighter sedation including protocols (50), expert staffing patterns (51) and daily interruption of sedation (52). To date, systematic reviews have not identified consistently useful strategies (53,54), although reviews are ongoing (55).

This review represents the most comprehensive description of the current evidence related to sedation depth and patient outcomes. Despite the use of liberal inclusion criteria, and a wide range of outcomes examined, the certainty of evidence remains low and inconsistent. Additionally, the findings are limited by the variable nature of how 'lighter' and 'deeper' sedation were determined in the studies, the lack of control of analgesic agents and the frequent lack of determining this differentiation *a priori*

or indeed stating it as an aim. In some studies, the only measure of sedation depth was average dose of sedation, which may not reflect sedative effect on the individual patient. Ideally validated sedation scores such as RASS or SAS should be used to indicate the actual depth to which a patient is sedated. Yet, despite a sensitivity analysis of studies where the difference in sedation depth was based on RASS or SAS, the lack of consistency in effect on patient outcomes remained. The review only included studies that used sedation assessment scales validated for use in the ICU environment in international practice guidelines (10), and thus may have had the effect of biasing the meta-analysis. The review was also limited by including English language publications and published data only. The preponderance of cohort studies including those using two groups of patients before and after a behaviour change intervention, and the implicit limitations of them, represents a limitation of this body of evidence. There was also no examination of the effect of sedation depth on related activities such as early mobilisation or on infrequently measured adverse events such as thromboembolic events.

Based on the low certainty of evidence, there is an urgent need for systematic evaluation of the effect of sedation depth on patient-centred outcomes to provide direction for sedation management. Studies addressing this question should use a randomised controlled trial design, ideally with randomisation at cluster level to achieve cultural change in clinician behaviour. Studies should incorporate *a priori* identification of target 'light' sedation levels, based on individual patient need, and the effect on a range of patient-centred outcomes (56,57) should be assessed.

Despite inconsistency in results, all clinical benefits identified in this review were related to lighter sedation, and importantly this review did not identify any harm related to lighter sedation. In this context, strategies to embed lighter levels of patient sedation in critical care are warranted. The challenging and multi-dimensional nature of sedation practice has been identified (58), and additional evidence-based strategies are urgently needed to optimise sedation and related areas of care such as early mobilisation.

CONCLUSION

Despite a considerable body of evidence discussing the relationship between sedation depth and various outcomes, we identified low to very low quality evidence suggesting that lighter sedation may be beneficial in some patient outcomes. The inconsistency of this evidence is exacerbated by the variable risk of bias in included studies, the different evidence of impact between RCTs and cohort studies, the inconsistent evidence of benefit across different outcomes and the inconsistent methods used, preventing combining data in meta-analyses. Future studies using rigorous controlled trial designs measuring patient centred outcomes, with randomisation occurring at the cluster level, are needed to understand the benefits associated with lighter patient sedation across a range of patient outcomes.

ACKNOWLEDGMENTS

The authors would like to thank Dr Maggie Tarling for assistance with some of the data extraction.

References

1. Devlin JW, Skrobik Y, Gelinas C, et al. Clinical Practice Guidelines for the Prevention and Management of Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption in Adult Patients in the ICU. *Crit Care Med* 2018; 46: e825-e873.
2. Vincent JL, Shehabi Y, Walsh TS, et al. Comfort and patient-centred care without excessive sedation: the eCASH concept. *Intensive Care Med* 2016; 42: 962-971.
3. Balas MC, Burke WJ, Gannon D, et al. Implementing the awakening and breathing coordination, delirium monitoring/management, and early exercise/mobility bundle into everyday care: opportunities, challenges, and lessons learned for implementing the ICU Pain, Agitation, and Delirium Guidelines. *Crit Care Med* 2013; 41: S116-127.
4. Page VJ, Ely EW, Gates S, et al. Effect of intravenous haloperidol on the duration of delirium and coma in critically ill patients (Hope-ICU): a randomised, double-blind, placebo-controlled trial. *The lancet Respiratory medicine* 2013; 1: 515-523.
5. Hodgson C, Bellomo R, Berney S, et al. Early mobilization and recovery in mechanically ventilated patients in the ICU: a bi-national, multi-centre, prospective cohort study. *Crit Care* 2015; 19: 81.
6. Shehabi Y, Howe BD, Bellomo R, et al. Early Sedation with Dexmedetomidine in Critically Ill Patients. *N Engl J Med* 2019; 380: 2506-2517.
7. Stephens RJ, Dettmer MR, Roberts BW, et al. Practice Patterns and Outcomes Associated With Early Sedation Depth in Mechanically Ventilated Patients: A Systematic Review and Meta-Analysis. *Crit Care Med* 2018; 46: 471-479.
8. Dodd S, Clarke M, Becker L, et al. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. *J Clin Epidemiol* 2018; 96: 84-92.
9. Reay H, Arulkumaran N, Brett SJ. Priorities for Future Intensive Care Research in the UK: Results of a James Lind Alliance Priority Setting Partnership. *Journal of the Intensive Care Society* 2014; 15: 288-296.
10. 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.
11. Higgins JPT, Thomas J, Chandler J, et al. Cochrane handbook for systematic reviews of interventions version 6.1. Cochrane 2020.
12. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ* 2016; 355: i4919.
13. Wan X, Wang W, Liu J, et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014; 14: 135.
14. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 2010; 36: 1-48.
15. R Core Team. R: A language and environment for statistical computing. 2018 [cited 2019 10/05/2019].
16. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; 64: 401-406.
17. Balzer F, Weiss B, Kumpf O, et al. Early deep sedation is associated with decreased in-hospital and two-year follow-up survival. *Crit Care* 2015; 19: 197.
18. Bugedo G, Tobar E, Aguirre M, et al. The implementation of an analgesia-based sedation protocol reduced deep sedation and proved to be safe and feasible in patients on mechanical ventilation. *Rev Bras Ter Intensiva* 2013; 25: 188-196.
19. Dale CR, Bryson CL, Fan VS, et al. A greater analgesia, sedation, delirium order set quality score is associated with a decreased duration of mechanical ventilation in cardiovascular surgery patients. *Crit Care Med* 2013; 41: 2610-2617.

20. Faust AC, Rajan P, Sheperd LA, et al. Impact of an Analgesia-Based Sedation Protocol on Mechanically Ventilated Patients in a Medical Intensive Care Unit. *Anesth Analg* 2016; 123: 903-909.
21. Guttormson JL. "Releasing a lot of poisons from my mind": patients' delusional memories of intensive care. *Heart Lung* 2014; 43: 427-431.
22. Khan BA, Fadel WF, Tricker JL, et al. Effectiveness of implementing a wake up and breathe program on sedation and delirium in the ICU. *Crit Care Med* 2014; 42: e791-795.
23. Mehta S, Burry L, Cook D, et al. Daily sedation interruption in mechanically ventilated critically ill patients cared for with a sedation protocol: a randomized controlled trial. *JAMA* 2012; 308: 1985-1992.
24. Nassar Junior AP, Park M. Daily sedative interruption versus intermittent sedation in mechanically ventilated critically ill patients: a randomized trial. *Annals of intensive care* 2014; 4: 14.
25. Olsen HT, Nedergaard HK, Strom T, et al. Nonsedation or Light Sedation in Critically Ill, Mechanically Ventilated Patients. *N Engl J Med* 2020; 382: 1103-1111.
26. Quenot JP, Ladoire S, Devoucoux F, et al. Effect of a nurse-implemented sedation protocol on the incidence of ventilator-associated pneumonia. *Crit Care Med* 2007; 35: 2031-2036.
27. Samuelson KA, Lundberg D, Fridlund B. Light vs. heavy sedation during mechanical ventilation after oesophagectomy-a pilot experimental study focusing on memory. *Acta Anaesthesiol Scand* 2008; 52: 1116-1123.
28. Shehabi Y, Bellomo R, Reade MC, et al. Early goal-directed sedation versus standard sedation in mechanically ventilated critically ill patients: a pilot study*. *Crit Care Med* 2013; 41: 1983-1991.
29. Strom T, Martinussen T, Toft P. A protocol of no sedation for critically ill patients receiving mechanical ventilation: a randomised trial. *Lancet* 2010; 375: 475-480.
30. Treggiari MM, Romand JA, Yanez ND, et al. Randomized trial of light versus deep sedation on mental health after critical illness. *Crit Care Med* 2009; 37: 2527-2534.
31. Ren XL, Li JH, Peng C, et al. Effects of ABCDE Bundle on Hemodynamics in Patients on Mechanical Ventilation. *Med Sci Monit* 2017; 23: 4650-4656.
32. Sen S, Grgurich P, Tulolo A, et al. A Symptom-Triggered Benzodiazepine Protocol Utilizing SAS and CIWA-Ar Scoring for the Treatment of Alcohol Withdrawal Syndrome in the Critically Ill. *Ann Pharmacother* 2017; 51: 101-110.
33. Shehabi Y, Bellomo R, Reade MC, et al. Early intensive care sedation predicts long-term mortality in ventilated critically ill patients. *Am J Respir Crit Care Med* 2012; 186: 724-731.
34. Arabi Y, Haddad S, Hawes R, et al. Changing sedation practices in the intensive care unit--protocol implementation, multifaceted multidisciplinary approach and teamwork. *Middle East J Anaesthesiol* 2007; 19: 429-447.
35. Mendes CL, Vasconcelos LC, Tavares JS, et al. Ramsay and Richmond's scores are equivalent to assessment sedation level on critical patients. *Rev Bras Ter Intensiva* 2008; 20: 344-348.
36. Capuzzo M, Pinamonti A, Cingolani E, et al. Analgesia, sedation, and memory of intensive care. *Journal of Critical Care* 2001; 16: 83-89.
37. Samuelson K, Lundberg D, Fridlund B. Memory in relation to depth of sedation in adult mechanically ventilated intensive care patients. *Intensive Care Med* 2006; 32: 660-667.
38. Samuelson KAM, Lundberg D, Fridlund B. Stressful memories and psychological distress in adult mechanically ventilated intensive care patients - a 2-month follow-up study. *Acta Anaesthesiol Scand* 2007; 51: 671-678.
39. Shehabi Y, Bellomo R, Kadiman S, et al. Sedation Intensity in the First 48 Hours of Mechanical Ventilation and 180-Day Mortality: A Multinational Prospective Longitudinal Cohort Study. *Crit Care Med* 2018; 46: 850-859.
40. Costa JB, Marcon SS, Macedo CR, et al. Sedation and memories of patients subjected to mechanical ventilation in an intensive care unit. *Rev Bras Ter Intensiva* 2014; 26: 122-129.

41. Burry L, Cook D, Herridge M, et al. Recall of ICU Stay in Patients Managed With a Sedation Protocol or a Sedation Protocol With Daily Interruption. *Crit Care Med* 2015; 43: 2180-2190.
42. Strom T, Stylsvig M, Toft P. Long-term psychological effects of a no-sedation protocol in critically ill patients. *Crit Care* 2011; 15: R293.
43. Wade DM, Brewin CR, Howell DC, et al. Intrusive memories of hallucinations and delusions in traumatized intensive care patients: An interview study. *Br J Health Psychol* 2015; 20: 613-631.
44. Train S, Kydonaki K, Rattray J, et al. Frightening and Traumatic Memories Early after Intensive Care Discharge. *Am J Respir Crit Care Med* 2019; 199: 120-123.
45. Aragon RE, Proano A, Mongilardi N, et al. Sedation practices and clinical outcomes in mechanically ventilated patients in a prospective multicenter cohort. *Crit Care* 2019; 23: 130.
46. de Jonghe B, Aboab J, Aissaoui N, et al. Impact of oversedation prevention in ventilated critically ill patients: a randomized trial—the AWARE study. *Annals of intensive care* 2018; 8: 93.
47. Chanques G, Conseil M, Roger C, et al. Immediate interruption of sedation compared with usual sedation care in critically ill postoperative patients (SOS-Ventilation): a randomised, parallel-group clinical trial. *The lancet Respiratory medicine* 2017; 5: 795-805.
48. Pun BT, Balas MC, Barnes-Daly MA, et al. Caring for Critically Ill Patients with the ABCDEF Bundle: Results of the ICU Liberation Collaborative in Over 15,000 Adults. *Crit Care Med* 2019; 47: 3-14.
49. Tanaka LM, Azevedo LC, Park M, et al. Early sedation and clinical outcomes of mechanically ventilated patients: a prospective multicenter cohort study. *Crit Care* 2014; 18: R156.
50. Brook AD, Ahrens TS, Schaiff R, et al. Effect of a nursing-implemented sedation protocol on the duration of mechanical ventilation. *Crit Care Med* 1999; 27: 2609-2615.
51. Nassar AP, Jr., Zampieri FG, Salluh JI, et al. Organizational factors associated with target sedation on the first 48 h of mechanical ventilation: an analysis of checklist-ICU database. *Crit Care* 2019; 23: 34.
52. Kress JP, Pohlman AS, O'Connor MF, et al. Daily interruption of sedative infusions in critically ill patients undergoing mechanical ventilation. *N Engl J Med* 2000; 342: 1471-1477.
53. Burry L, Rose L, McCullagh IJ, et al. Daily sedation interruption versus no daily sedation interruption for critically ill adult patients requiring invasive mechanical ventilation. *The Cochrane database of systematic reviews* 2014: Cd009176.
54. Aitken LM, Bucknall T, Kent B, et al. Protocol-directed sedation versus non-protocol-directed sedation in mechanically ventilated intensive care adults and children. *The Cochrane database of systematic reviews* 2018; 11: Cd009771.
55. Moraes FdS, Marengo LL, Silva MT, et al. ABCDE and ABCDEF care bundles: A systematic review protocol of the implementation process in intensive care units. *Medicine (Baltimore)* 2019; 98: e14792-e14792.
56. Blackwood B, Ringrow S, Clarke M, et al. A Core Outcome Set for Critical Care Ventilation Trials. *Crit Care Med* 2019; 47: 1324-1331.
57. Needham DM, Sepulveda KA, Dinglas VD, et al. Core Outcome Measures for Clinical Research in Acute Respiratory Failure Survivors. An International Modified Delphi Consensus Study. *Am J Respir Crit Care Med* 2017; 196: 1122-1130.
58. Hetland B, Guttormson J, Tracy MF, et al. "Sedation is tricky": A qualitative content analysis of nurses' perceptions of sedation administration in mechanically ventilated intensive care unit patients. *Aust Crit Care* 2018; 31: 153-158.

Table 1 Criteria used in studies to separate 'deeper' vs 'lighter' sedation[1]

Study	Control / 'deeper' sedation	Intervention / 'lighter' sedation
Balzer 2015	Patients had >85% RASS scores ≤-3 First RASS -5[-5 to -4] Time to reach first RASS>-3:79[52-141]hrs	First RASS: -4[-5 to -1], p< 0.001 Time to reach first RASS >-3: 11 [5-20] hrs, p= 0.001
Bugedo 2013	Midazolam: 0.03[0.01-0.06] mg/kg/hr Fentanyl: 0.6[0.1-1.4] mcg/kg/hr Proportion: SAS 1–2:55%; SAS 3–4:37%	Midazolam: 0.01[0-0.03] mg/kg/hr, p<0.001 Fentanyl: 1.5[0.8-2.4] mcg/kg/hr, p<0.001 Proportion: SAS 1–2:44%; SAS 3–4:49%, p=0.001
Burry 2015	Midazolam: 97.0±200.8 mg/patient/day Fentanyl: 1.9±3.5 mg/patient/day	Midazolam: 64.7±245.8 mg/patient/day, p<0.0001 Fentanyl: 1.1±2.0 mg/patient/day, p<0.0001
Dale 2014	Hrly benzodiazepine dose: 0.23±0.018 mg Total benzodiazepine dose: 49.2±156.5 mg 24 hr weighted av. RASS: -1.30±0.026	Hrly benzodiazepine dose: 0.15±0.011 mg, p<0.01 Total benzodiazepine dose: 17.2±53.6, p<0.01 24 hr weighted av. RASS: -0.99±0.023, p <0.01
Faust 2016	RASS (median): -2.57[-3.23 to -1.40] % RASS scores -3 to -5 in first 24 hrs: 46.8±46.9%	RASS (median): -1.25[-2.3 to -0.40], p=0.001 % RASS scores -3 to -5 in first 24 hrs: 27.3±37.3%, p=0.006
Guttormson 2011	Classed as 'minimally arousable' based on Sedation Intensity Score	Classed as 'easily arousable' based on Sedation Intensity Score
Khan 2014	RASS scores – weekdays: median -4 RASS scores – weekends: median – 5	RASS – weekdays: increased by 0.88, p <0.0001 RASS – weekends: increased by 1.20, p <0.0001
Mehta 2012[2]	Midazolam: 102±326 mg/pt/day Fentanyl: 1780±4135 µg/pt/day	Midazolam: 102±326 mg/pt/day, p=0.04 Fentanyl: 1070±2066 µg/pt/day, p<0.001
Nassar Junior 2014	SAS scores: 3.2(2.6-3.7) Midazolam: 45(0,201) mg, Fentanyl: 1500(520-4215) mg	SAS scores: 3.6(3.4-4.0), p=0.035 Midazolam: 0(0.0-0.05) mg, p<0.001 Fentanyl: 300(100-1520) mg, p=0.004
Olsen 2020	Midazolam mg/kg/hr (day 2-28):0.000187 (0-0.003410) Propofol mg/kg/hr (day 1-2):0.84 (0.29-1.2); (day 3-28):0.0064 (0-0.034) Mean RASS: Day 1: -2.3; Day 7: -1.8	Midazolam mg/kg/hr (day 2-28):0(0-0.000005); NS Propofol mg/kg/hr (day 1-2): 0.22(0-0.054); Diff: -0.62 (-0.72; -0.53); (day 3-28): 0(0-0.013); Diff: -0.0063(-0.874; -0.0037) Mean RASS: Day 1: -1.3; Day 7: -0.8
Quenot 2007	Midazolam: 92±59 mg/pt/day Propofol: 2900±1400 mg /pt/day	Midazolam: 44±31 mg/pt/day, p=0.001 Propofol: 1840±750 mg/pt/day, p=0.01
Ren 2017	Sufentanil: 0.030±0.007 mg/kg/hr Midazolam: 0.029±0.007 mg/kg/hr	Sufentanil: 0.018±0.009 mg/kg/hr, p<0.0001 Midazolam: 0.017±0.009 mg/kg/hr, p<0.0001
Samuelson 2008	Target MAAS: 1 – 2 Actual MAAS: median 1.25(1.0)	Target MAAS: 3 – 4 Actual MAAS: median 3.0(0.0)
Sen 2017	Total benzodiazepine dose: 450±701 mg	Total benzodiazepine dose: 74±159 mg, p<0.01
Shehabi 2013	Dexmedetomidine: 20.58(20.58-20.58) µg Midazolam: 0.3(0.23-0.76) mg Propofol: 33.55(13.54-77.07) mg RASS assessments -2 to +1: 38%	Dexmedetomidine:36.55(16.38-13.23)µg, p<0.0001 Midazolam: 0.06(0.02-1) mg, p=0.036 Propofol: 9.89(2.41-22.51) mg, p=0.046 RASS assessments -2 to +1: 66%, p=0.01
Shehabi 2012	RASS -3 to -5 at 48 hours	RASS lighter than -3 to -5 at 48 hours
Strøm 2011	Propofol: 1.40(0.52-2.04) mg/kg/hr Midazolam: 0.01(0-0.04) mg/kg/hr	Propofol: 0(0-1.26) mg/kg/hr, p=0.013 Midazolam: 0(0-0) mg/kg/hr, p=0.003
Strøm 2010	Propofol: 0.77(0.15-1.65) mg/kg/hr Midazolam: 0.003(0-0.024) mg/kg/hr	Propofol: 0(0-0.52) mg/kg/hr, p=0.0001 Midazolam: 0(0-0) mg/kg/hr, p<0.0001
Treggiari 2009	Target Ramsay sedation score 3 – 4 Daily median Ramsay range: 3(2-4.5) to 4(3-5) Daily Midazolam range: 24.2±45.1 to 95.3±124.5 mg	Target Ramsay sedation score 1 – 2 Daily median Ramsay range: 1(1-2) to 3(1-3) Daily Midazolam range: 3.0±5.0 to 11.7±23.2 mg

1. It was not possible to create 2 categories of 'deeper' or 'lighter' sedation in 7 studies (34-39)
2. 'deeper' sedation group was the intervention (Daily Interruption of Sedation) group

Abbreviations: MAAS: Motor Activity Assessment Scale, RASS: Richmond Agitation Sedation Scale, SAS: Riker Sedation Agitation Scale.

Table 2: Summary of findings

Outcomes	Study Type	Number of studies (participants)	Values for clinical parameters in deep sedation groups for included studies [mean (range) of mean value reported in each study][1]	Effect Estimate & 95% CI (Risk ratio for events[2]; Mean Difference for duration[3])	I^2	Grade rating
Primary Outcomes						
Mortality						
ICU mortality (%)	RCT	4 (725)	28.8 (14.1 – 40.0)	0.82 [0.58 to 1.17][2]	30%	Moderate
	Cohort	3 (2474)	22.6 (2.2 – 38.9)	0.50 [0.13 to 1.86][2]	97%	Very low
Physiological outcomes						
Duration of mechanical ventilation (days)	RCT	2 (165)	6.6 (5.5 – 7.7)	-1.44 [-3.79 to 0.91][3]	20%	Moderate
	Cohort	8 (3304)	7.1 (1.2 – 10.7)	-1.54[-2.68 to -0.39][3]	87%	Very low
Secondary outcomes						
Mortality						
Hospital mortality (%)	RCT	5 (762)	29.8 (12.5 – 46.6)	0.93 [0.75 to 1.15][2]	0%	Moderate
	Cohort	5 (4636)	29.2 (13.7 – 44.7)	0.73 [0.41 to 1.30][2]	96%	Very low
Physiological outcomes						
Time to extubation (days)	RCT	1 (423)	8.0 (8.0 – 8.0)	-0.67 [-1.95 to 0.61][3]	0%	Moderate
	Cohort	2 (2132)	5.6 (3.7 – 7.4)	-3.77[-5.49 to -2.06][3]	98%	Very low
Ventilator free days to day 28 (days)	RCT	4 (910)	15.3 (9.6 – 20.1)	2.62 [-0.09 to 5.34][3]	31%	Moderate
	Cohort	2 (431)	17.0 (10.3 – 23.6)	0.65 [-0.65 to 1.95]*	0%	Low
Delirium (%)	RCT	4 (556)	30.1 (0 – 52.8)	1.04 [0.88 to 1.23][2]	0%	Moderate
	Cohort	4 (3953)	37.2 (10.7 – 55.3)	1.01 [0.63 to 1.62][2]	95%	Very low
Resource Use						
ICU length of stay (days)	RCT	6 (1462)	14.8 (6.3 – 28.0)	0.28 [-1.46 to 2.02][3]	32%	Moderate
	Cohort	8 (4537)	11.9 (3.7 – 23.7)	-4.30[-7.39 to -1.21][3]	97%	Very low
Hospital length of stay (days)	RCT	5 (762)	27.5 (16.6 – 58.6)	-0.69 [-6.96 to 5.58][3]	80%	Very low
	Cohort	6 (4917)	19.9 (12.3 – 30.7)	-4.21[-7.22 to -1.19][3]	88%	Very low
Tracheostomy (%)	RCT	4 (725)	15.4 (3.3 – 29.3)	1.07 [0.81 to 1.43][2]	0%	Moderate
	Cohort	2 (431)	12.3 (7.7 – 16.9)	0.59 [0.31 to 1.12][2]	0%	Very low
Adverse Events						
Self-extubation (%)	RCT	2 (189)	3.2 (3.1 – 3.3)	1.31 [0.30 to 5.82][2]	0%	Moderate
	Cohort	3 (854)	6.4 (3.1 – 9.0)	1.32 [0.84 to 2.09][2]	0%	Low

Re-intubation (%)	RCT	5 (1348)	7.2 (1.6 – 13.3)	1.45 [0.78 to 2.71][2]	30%	Low
	Cohort	2 (362)	4.2 (1.5 – 6.9)	1.07 [0.43 to 2.65][2]	0%	Very low
VAP (%)	RCT	1 (113)	12.1 (12.1 – 12.1)	0.90 [0.32 to 2.52][2]	0%	Low
	Cohort	2 (1906)	10.8 (6.5 – 15.0)	0.56 [0.33 to 0.96][2]	51%	Very low

Abbreviations: ICU: Intensive Care Unit, RCT: Randomized Control Trial, VAP: Ventilator-Associated Pneumonia

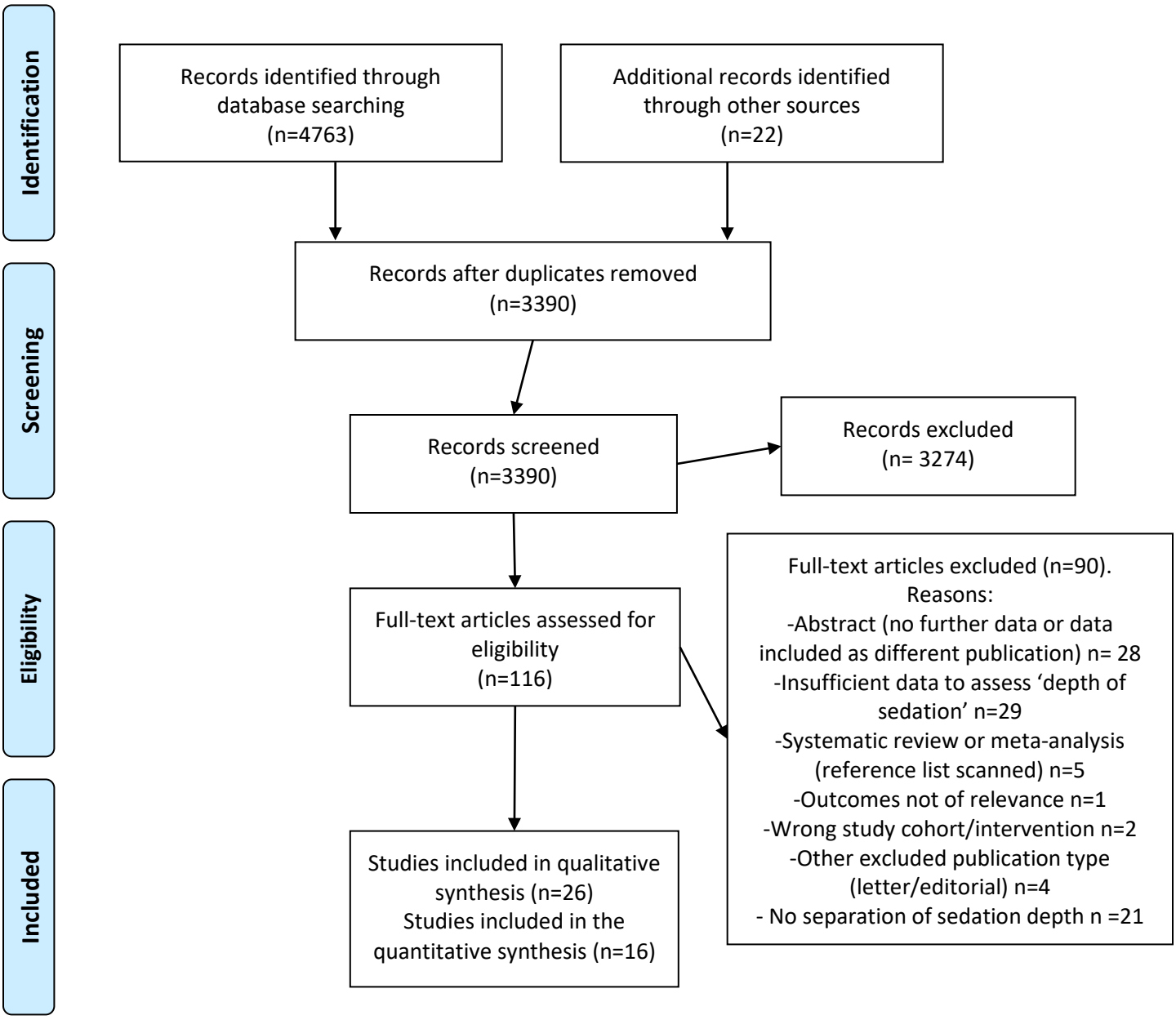
Grade Assessment: RCTs started at high quality; Cohort studies started at low quality (due to risk of bias); Reasons for downgrade included risk of bias (RCTs), imprecision, inconsistency, indirectness.

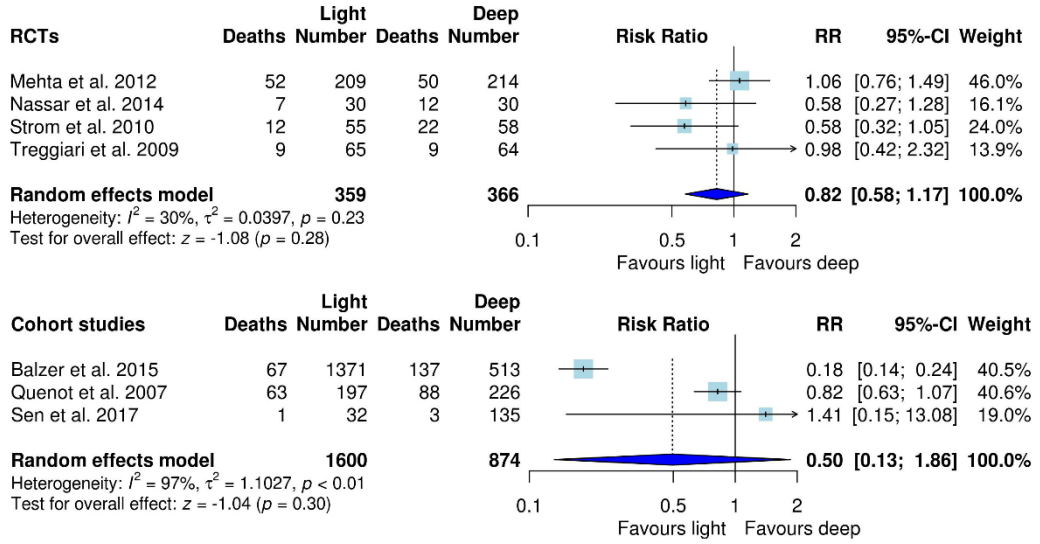
Notes: 1 – Values for clinical parameters in study populations in deep sedation groups for included studies [mean (range) of mean deep sedation group value reported in each study]; 2 – Risk ratio for events; 3 – Mean Difference for duration.

Table 3. Life impact outcomes

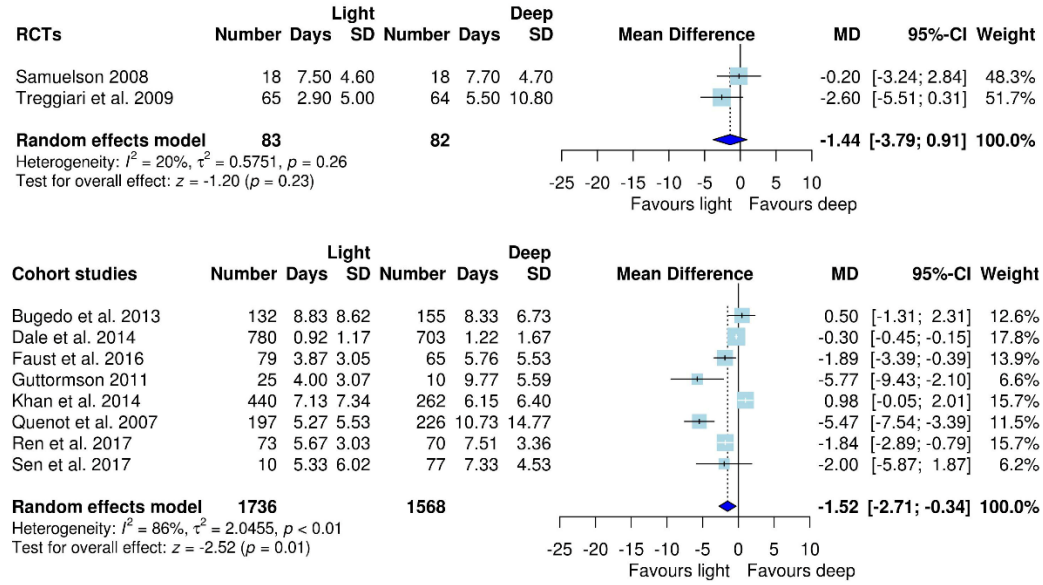
Study	Time point	Outcome measured	Results
Bugedo 2013	1 year post discharge	Screening for memories via telephone interview. Post-Traumatic Stress Syndrome – 10 (PTSS-10) Scale	No difference in incidence of nightmares (n= 22[55%] vs 15[43%], p=0.294), severe anxiety or panic (n= 16[40%] vs 12[34%], p=0.610) or pain (n= 12[30%] vs 13[37%], p=0.513, feelings of suffocation or PTSS-10 (28[19-3(sic)] vs 26[17-38], p=0.840) questionnaire scores between the deep and light sedation groups.
Burry 2015 (sub-study of Mehta 2012)	28 days post ICU discharge	ICU Memory Tool	Patients who reported 'not remembering the ICU' had less sedation (average daily midazolam dose 26.9 [SD 63.7] vs 82.5 [SD 314] mg), but no difference in SAS scores (3.34 [SD 0.70] vs 3.27 [0.65]). In a multivariate model, total midazolam (OR 1.182, 95% CI 0.37 – 3.81) and fentanyl (OR 2.27, 95% CI 0.64 – 8.14) exposure above the mean (deeper sedation) were not associated with increased risk of delusional memories.
Capuzzo 2001	6 months post hospital discharge	Memories explored through semi-structured interviewed, then retrospectively categorised.	No significant difference in recall of factual (A (No morphine/minimal sedatives): n=16[36%]; B (Morphine only): 29[34%]; C (Morphine and sedatives): 4[18%]), sensation (A: n=4[9%]; B: 13[15%]; C: 3[14%]) or emotional (A: n=4[9%]; B: 6[7%]; C: 4[18%]) memories of ICU between the groups.
Costa 2014	Approximately 3 days post ICU discharge	Locally adapted ICU Memory Tool Not specified how anxiety, depression or PTSD were measured	No difference in the incidence of anxiety, depression or PTSD across mild-moderate, deep or not sedation groups. Patients who received any level of sedation reported less real memories (21[24%] vs 29 [69%]), more real and illusory memories (42[49%] vs 9[21%]), more illusory memories (7[8%] vs 0) and more amnesia (16[19%] vs 4[10%]) than patients who received no sedated (p<0.001).
Nassar Junior 2014	6 months post ICU discharge	Impact of Events Scale (IES)	No difference in the level of psychological stress on the IES (22[8-31] vs 16[4-34], p=0.750) between intermittent sedation or daily interruption of sedation groups.
Samuelson 2006	3 – 5 days post ICU discharge	ICU Memory Tool	Deep sedation was associated with amnesia (OR 1.60, 95% CI 1.35 – 1.91) and delusional memories (OR 1.76, 95% CI 1.14 – 2.72) on multivariate analysis.
Samuelson 2007	3 – 5 days post ICU discharge	ICU Memory Tool Locally adapted ICU Stressful Experiences Questionnaire	Patients with memory of ETT had higher proportion of MAAS 3 (awake) than those with no memory (0.56[0.42] vs 0.18[0.42], p<0.0001) - this relationship was confirmed on multivariate analysis (OR 1.45, 95% CI 1.29 – 1.62). Similarly, patients with a higher proportion of MAAS 3 were more likely to be bothered by memories of stressful experiences of ICU (OR 1.37, 95% CI 1.13 – 1.67).
Samuelson 2008	3 – 5 days post ICU	ICU Memory Tool Locally adapted ICU	No difference in memories of ICU (n=15[88%] vs 17[94%], p=0.60), presence of delusional

	discharge & 2 months	Stressful Experiences Questionnaire Impact of Events Scale - Revised	memories in ICU (n=1[6%] vs 6[33%], p=0.09), or memories of pain (n=4[23%] vs 9[50%], p=0.20) between the groups.
Strom 2011	2 years post randomisation	ICU Memory Tool SF-36, Beck Depression Index (BDI), Impact of Events Scale (IES), State Anxiety Inventory, PTSD Symptoms (PTSS) - 10	No difference in psychological problems post-discharge (n=2[15%] vs 6[46%], p=0.20), PTSS-10 score >35 (n=1[8%] vs 0[0%], p=0.14) or any of the other psychological health outcomes between the no sedation and sedation groups.
Treggiari 2009	At discharge and 4 weeks post ICU discharge	PTSD Checklist (PCL) Impact of Events Scale – Revised (IES-R) Hospital Anxiety and Depression Scale (HADS)	No difference in PTSD questionnaire score (discharge: 57±30 vs 52±33, p=0.39; 4 wk follow-up: 56±29 vs 46±29, p=0.07), PTSD symptom clusters, anxiety or depression (discharge: 6.5±4.7 vs 5.3±3.4, p=0.13; 4 wk follow-up: 3.1±3.7 vs 3.4±3.7, p=0.72) scores or cases at either discharge or 4 week follow-up between the groups.


Identification
Screening
Eligibility
Included
Records identified through database searching (n=4763)
Additional records identified through other sources (n=22)
Records after duplicates removed (n=3390)
Records screened (n=3390)
Records excluded (n= 3274)
Full-text articles assessed for eligibility (n=116)
Full-text articles excluded (n=90). Reasons:
-Abstract (no further data or data included as different publication) n= 28
-Insufficient data to assess 'depth of sedation' n=29
-Systematic review or meta-analysis (reference list scanned) n=5
-Outcomes not of relevance n=1
-Wrong study cohort/intervention n=2
-Other excluded publication type (letter/editorial) n=4
- No separation of sedation depth n =21
Studies included in qualitative synthesis (n=26)
Studies included in the quantitative synthesis (n=16)



a)



b)

Figure 2: Forest plots for primary outcome: a) ICU mortality; b) Duration of mechanical ventilation

Note: data converted from median/IRQ to mean/SD[12] for duration of MV in the following studies: Bugedo et al 2013; Dale et al 2014; Guttormson et al 2011; Quenot et al 2007; Sen et al 2017.

DETAILED DESCRIPTION OF METHODS

Protocol Registration

The protocol for this systematic review was registered on PROSPERO (CRD42018092554; www.crd.york.ac.uk/prospero/display_record.php?RecordID=92554).

Information sources, search strategy and eligibility criteria

Databases searched: MEDLINE, Embase, Cochrane Register of Controlled Trials (CENTRAL), Cumulative Index to Nursing and Allied Health Literature (CINAHL) and PsycINFO

Search strategy: (intensive care OR critical care OR critically ill) AND (sedat* OR midazolam OR propofol) AND (length of stay OR mortality OR outcome assessment OR physical function OR psychological OR cognitive OR memories)

We searched for publications reporting randomized controlled, quasi-experimental and before-after trials, and cohort studies (both prospective and retrospective) published between January 2000 and February 2020 and published in English. Review articles, correspondences, editorials and non-human studies were excluded but reference lists were scanned for relevant publications. Studies published prior to 2000 were not included given the significant changes in critical care since that time, and therefore the potential lack of relevance to current care.

We included studies in adult patients (usually ≥18 years, although if a jurisdiction categorised adults as ≥16 years we included that study) receiving invasive mechanical ventilation in ICU; including patients who commenced their ventilation in another location, e.g. ED, operating room. We excluded studies in patients receiving non-invasive ventilation and mechanically ventilated patients not admitted to ICU. We excluded studies where the intervention included different sedative agents as it was not possible to determine if any effect on outcome was due to the different agent or different depth. We defined our exposure or intervention as deeper sedation at any time throughout the period of mechanical ventilation in the ICU. Our classification of depth of sedation as either ‘lighter’ or ‘deeper’ was based on published information and incorporated both sedation assessment and average dose of sedatives. There was no predefined level of ‘deeper’ sedation, only that one group of patients received deeper sedation than the other group. The primary study authors did not necessarily label the groups as ‘lighter’ or ‘deeper’ sedation – we made that judgement during data extraction. Sedation

depth could be measured through any objective measures of sedation depth including, but not limited to, assessment using a validated sedation assessment instrument, hourly or daily doses of sedatives. Measures of total doses of sedatives in isolation were not sufficient, given total dose could be affected by length of stay. Further, where there was inconsistency between measures, e.g. no separation in hourly dose, but a separation in total dose, preference was given to measures of sedative state (e.g. sedation assessment) or sedation administered in discreet periods (e.g. hours or days) rather than total dose. Only the Richmond Agitation-Sedation Scale (RASS) and Riker Sedation-Agitation Scale (SAS) were accepted as validated instruments (1).

Study selection and data extraction

Titles and abstracts were screened for inclusion independently by two researchers to determine relevance, with full text of included studies then reviewed by two authors to assess eligibility. The reference lists of eligible articles were checked to identify additional publications of interest. Discrepancy between researchers at any stage was resolved through discussion and consultation with a third reviewer where necessary to achieve consensus. Studies where separation of depth of sedation into 'lighter sedation' and 'deeper sedation' could not be identified were excluded. Where studies included >2 groups based on depth of sedation they were not able to be included in the meta-analysis but were retained in the additional analyses. Where two or more papers reported different outcomes from the same cohort of patients, this relationship was indicated in the study description and data were not included twice in any analysis.

For eligible articles, two authors extracted data on study design, population and setting, patient characteristics (e.g. age, gender, severity of illness score), study interventions, measure of depth of sedation (methodology and results) and all relevant outcomes. Data extraction was recorded on standardised forms. Quality was assessed using the relevant Critical Appraisal Skills Programme (CASP) data extraction and quality assessment forms and completed forms were compared for any discrepancies and discussed to achieve consensus.

Assessment of bias

The domains of bias for RCTs were assessed using an adapted form of the Critical Appraisal Skills Programme (CASP) Checklist – Randomised controlled trials (2) and included: 1) random sequence

generation, 2) allocation concealment, 3) blinding of participants, outcome assessors and others, 4) incomplete outcome data, and 5) selective reporting. For cohort studies, an adapted form of the CASP Checklist – Cohort studies (3) was used to assess the domains of bias: 1) selection of cohort, 2) ascertainment of exposure, 3) assessment of outcome, and 4) adequacy of follow-up. Relevant confounding factors were not identified *a priori*, but were based on the study method and cohort and included demographic, clinical, and treatment variables with the potential to influence relevant outcomes. No studies were excluded on the basis of quality assessment.

Data Analysis

All studies that contained data suitable for combination in a meta-analysis for at least one of the pre-determined primary or secondary outcomes were included in the quantitative analysis. Data reported as median and inter-quartile range were converted to mean and standard deviation using the method devised by Wan et al (4). Random effects meta-analyses were undertaken with the meta package (5) in R (6). This allows for both within and between studies variance to be calculated, the latter being reflected in a statistical test of heterogeneity, and the I^2 that shows the percentage of the variation in the result that is due to heterogeneity rather than sampling error. Cohort studies and RCTs were analysed separately based on an *a priori* decision that they formed distinct types of evidence. The quality of evidence was rated using Grades of Recommendation, assessment, Development and Evaluation (GRADE) for all outcomes (7). For outcomes where significant methodological differences occurred (for example use of different instruments to measure an outcome or different time points) results were combined descriptively.

Changes from the protocol: During the review we identified that both duration of mechanical ventilation and mortality were key clinical outcomes, therefore we have presented them as co-primary outcomes in contrast to the protocol where ICU mortality was the sole primary outcome.

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.

2. Critical Appraisal Skills Programme (CASP). CASP Checklist - Randomised Controlled Trials. [serial online] 2018. Available from: www.casp-uk.net
3. Critical Appraisal Skills Programme (CASP). CASP Checklist - Cohort Study. [serial online] 2018. Available from: www.casp-uk.net
4. Wan X, Wang W, Liu J, et al. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med Res Methodol* 2014; 14: 135.
5. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* 2010; 36: 1-48.
6. R Core Team. R: A language and environment for statistical computing. 2018 [cited 2019 10/05/2019].
7. Balshem H, Helfand M, Schunemann HJ, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011; 64: 401-406.

Supplementary Table 1: Summary of study characteristics

Study; Location	Design, Primary Aim of Study	Setting & Dates	Intervention	Patients
Included in meta-analyses				
Balzer 2015; Germany	Retrospective, observational cohort to determine the effect of early deep sedation	4 ICUs: 2 surgical, 1 cardiac, 1 mixed, 1 medical	**Deeply sedated** vs **not deeply sedated**, based on RASS scores in first 48 hours of admission.	1884 MV pts in ICU ≥ 48 hrs
Bugedo 2013; Chile	Before/after, prospective cohort - effect of analgesia based, goal directed, nurse delivered sedation protocol	13 ICUs- details not reported	**Before:** SAS assessed twice a day, usual care. **After:** sedation protocol - defined doses of sedatives & analgesics, daily sedation goal	287 pts with expected MV>48 hrs
Dale 2014; USA	Before/after prospective cohort to determine effect of sedation protocol (assessment, DIS, SBT) on delirium & MV.	Single mixed trauma/surgical ICU	**Sedation protocol** with regular assessment and documentation of sedation and analgesia, DIS and daily SBTs.	1483 MV pts
Faust 2016; USA	Retrospective, before/after cohort to evaluate impact of analgesia-based sedation protocol	Single medical ICU	**Pre:** sedation goal, regular RASS, daily SAT, primarily propofol, **Post:** primarily fentanyl, other care similar	144 MV pts
Guttormson 2011; USA	Prospective, observational cohort to evaluate relationship between sedation and memories	Single mixed medical/surgical ICU	Patients' **pattern of sedation**, incorporating **Sedation Intensity Score** and **MASS**, determined retrospectively.	35 MV pts in ICU >24 hrs
Khan 2014; USA	Before/after prospective cohort study to evaluate the effects of a 'wake up and breath' program	Single mixed medical/surgical ICU	**Pre:** Physician directed all sedation and analgesia. **Post:** DIS, twice daily RASS assessment.	702 ICU pts
Mehta 2012; Canada and USA	RCT to compare protocolised sedation with protocolised sedation + DIS.	16 ICUs- various medical/surgical/trauma	**Protocolised sedation:** Sedative protocol with target RASS -3 – 0 or SAS 3 – 4. **Protocolised sedation + DIS:** Protocol with DIS	423 pts expected to require ≥48 hrs MV
Nassar Junior 2014; Brazil	RCT to compare effect of DIS and intermittent sedation on ventilator-free days and safety.	Single mixed ICU	**Intermittent:** Physician directed sedation/ analgesia, 6/24 SAS, DIS & SBT **DIS:** As above to achieve SAS 3 – 4, 8/24 SAS, DIS daily	60 pts expected to require >24 hrs MV
Olsen 2020; Denmark, Norway & Sweden	RCT to assess effect of a plan of no sedation compared to a plan of light sedation on mortality	11 mixed medical/surgical ICUs	**Control (sedation)**: continuous sedative infusions, RASS goal -2 to -3 **Intervention (no sedation):** no sedatives, bolus doses of morphine, sedation if required	710 pts expected to require >24 hrs MV
Quenot 2007; France	Before/after prospective cohort to determine if nurse-implemented protocol to achieve target sedation score reduced VAP	Single medical ICU	**Control:** physician managed sedatives/ analgesics **Protocol:** Protocol guided, nurse adjusted sedatives, target sedation score.	423 pts with >48 hrs MV
Ren 2017; China	Before/after prospective cohort to investigate effects	Type of ICU not reported	**Pre-ABCDE:** Physician directed sedation/ analgesia **Post-ABCDE:** daily SBT, delirium	143 pts with ≥48 hrs MV

	of ABCDE bundle on hemodynamic status.		monitoring/management, exercise.	
Samuelson 2008; Sweden	RCT to assess protocol feasibility, and examine patients' stressful memories of light vs heavy sedation	Single mixed medical/surgical/trauma ICU	Patients randomised to either **light sedation** (MAAS 3 – 4) or **heavy sedation** (MAAS 1 – 2).	36 post-operative MV pts
Sen 2017; USA	Before/after prospective cohort – evaluate a symptom-triggered benzodiazepine protocol for treatment of AW syndrome	2 medical ICUs	**Control:** Fixed dosing benzodiazepines on AW scale. **Protocol:** Combined symptom-triggered & fixed dosing on AW scale & SAS.	167 ICU pts requiring treatment of alcohol withdrawal >48 hrs
Shehabi 2013; Australia and New Zealand	RCT - pilot study to assess feasibility & safety of early goal-directed sedation	6 mixed medical/surgical ICU	**Standard:** Physician directed sedation, midazolam, propofol, opioids, 4/24 RASS; **Early goal directed:** dexmedetomidine to achieve RASS -2 to 1, opioids, 4/24 RASS	37 pts expected to be sedated & MV ≥24 hrs
Shehabi 2012; Australia and New Zealand	Prospective cohort: assess relationship between early sedation depth & time to extubation, delirium, mortality.	25 ICUs-specific details not reported	Usual care including RASS assessments categorised as: **Light** (-2 to +1), **Deep** (-3 to -5), **Agitation** (+2 to +4)	251 pts expected to be sedated & MV ≥24 hrs
Strøm 2010; Denmark	RCT to determine if no-sedation versus sedation with DIS reduced duration of MV.	Single closed mixed medical/surgical ICU	**Sedation with DIS** or **no sedation** with bolus morphine only.	140 pts MV >24 hrs enrolled; 113 in analysis
Treggiari 2009; Switzerland	RCT - determine effect of light vs deep sedation on mental health after critical illness.	Single mixed medical/surgical ICU	**Light:** awake and cooperative; **deep:** awakening on stimulation	129 pts expected to need >12 hrs MV
Not included in meta-analyses[#]				
Arabi 2007;[#] Saudi Arabia	Prospective, 4-arm before/after cohort to evaluate 1) education & 2) protocol directed sedation	Single closed, mixed medical/ surgical ICU	**Education:** Lectures/in services/bed side teaching **Protocol:** Goal-directed sedation protocol with regular assessment	207 MV pts in ICU ≥ 24 hrs
Burry 2015* (sub-study of Mehta, 2012); USA, Canada	Prospective cohort to describe the psychological outcomes after protocolised sedation +/- DIS.	16 mixed medical/surgical ICUs	Protocolised sedation: Sedative protocol with target RASS -3 – 0 or SAS 3 – 4. Protocolised sedation + DIS: Protocol with DIS	289 pts MV >48 hrs with analgesia and/or sedation
Capuzzo 2001;[#] Italy	Prospective, observational cohort to investigate relationship between ICU memories and analgesics/sedatives	2 mixed medical/surgical ICUs	Patients retrospectively grouped as receiving no morphine, only morphine, or morphine and other sedatives.	152 pts in ICU>24 hrs
Costa 2014;[#] Brazil	Prospective, observational cohort to investigate the relationship between sedation and ICU memories.	Type of ICU not reported	Patients retrospectively grouped based on **mild-moderate**, **deep** or **no sedation** in ICU.	128 MV pts in ICU >24 hrs
Mendes 2008;[#] Brazil	Prospective, observational cohort to compare RASS and Ramsay score and to relate these to mortality.	Single ICU-type not reported	No alterations to usual care. Sedation and agitation was assessed daily in all patients	45 pts with MV >48 hrs

			with both RASS and Ramsay score.	
Samuelson 2006;* Sweden	Prospective cohort to investigate relationship between ICU memories and depth of sedation	2 mixed medical/surgical ICUs	Patients received usual care. ICU memories assessed post-ICU, proportion of MAAS scores in categories of 0 – 2, 3, 4 – 6.	313 MV pts in ICU >24 hrs; 250 pts in this analysis
Samuelson 2007* (sub-study of Samuelson 2006); Sweden	Prospective cohort to investigate relationship between stressful experiences and depth of sedation	2 mixed medical/surgical ICUs	Interventions described above. Stressful experiences assessed - ICU Stressful Experiences Questionnaire (local adaptation)	313 MV pts in ICU >24 hrs; 206 pts in this analysis
Shehabi 2018;# New Zealand, Australia, Malaysia, Singapore	Prospective cohort to quantify relationship between early sedation depth and 180 day survival, time to extubation, delirium.	42 ICUs-specific details not reported	Usual care including RASS assessment	703 pts expected to be sedation & MV ≥24 hrs
Strøm 2011* (Sub-study of Strom 2010); Denmark	RCT to determine if no-sedation versus sedation with DIS affected long-term psychological outcomes.	Single closed mixed medical/ surgical ICU	Patients randomised to either sedation with DIS or no sedation with bolus morphine only.	26 pts MV >24 hrs

Abbreviations: DIS: Daily Interruption of Sedation, ICU: Intensive Care Unit, MAAS: Motor Activity Assessment Scale, MV: Mechanical Ventilation, RASS: Richmond Agitation Sedation Scale, RCT: Randomised Controlled Trial, SAS: Riker Sedation Agitation Scale, SAT: Spontaneous Awakening Trial, SBT: Spontaneous Breathing Trial.
Notes: Study excluded from meta-analysis due to: # unable combine participants to form 2 groups (lighter and deeper sedation) based on sedation depth – either 1, 3 or 4 groups were presented; * variable methods of outcome to assess psychological health.

Supplementary table 2 Measurement of depth of sedation

Study	Evidence of differing depth of sedation
Arabi 2007 *Saudi Arabia*	Less sedation in education period compared to no education period (daily propofol: Group (G)1: 236±65; G2 264±106; G3: 171±84; G4 105±33 mcg, p=0.10 all groups; p=0.03 G1&G2 vs G3&G4). More patients in ideal SAS range on nights 3 & 4 and day 4 in all groups compared to group 1 (Day 4 SAS 3-4: (G1: 15%; G2 30%; G3: 45%; G4 50%, p=0.012.
Balzer 2015 *Germany*	Patients were divided into two groups, based on sedation depth (RASS) (deeply sedated - >85% of documented RASS scores during the study period ≤-3). In addition, deeply sedated patients had a lower first RASS (-5[-5 to -4] vs -4[-5 to -1], p< 0.001) and took longer to reach a first RASS >-3 (79 [52-141] vs 11 [5-20] hrs, p= 0.001).
Bugedo 2013 *Chile*	Midazolam decreased in the intervention period compared with the control period (average rate - control: 0.03[0.01-0.06] vs intervention: 0.01[0-0.03], p<0.001); fentanyl dose increased in intervention period (average rate 0.6[0.1-1.4] vs 1.5[0.8-2.4], p<0.001). The proportion of SAS scores in the deep sedation range (SAS 1 – 2) was lower (55% vs 44%) while the proportion of SAS scores in the ideal range (SAS 3 – 4) was higher (37% vs 49%, p=0.001) in the intervention period.
Burry L 2015 *USA & Canada*	Higher doses of midazolam (97.0±200.8 vs 64.7±245.8 mg/patient/day, p<0.0001) and fentanyl (1.9±3.5 vs 1.1±2.0 mg/patient/day, p<0.0001) in the protocolised sedation with DIS group compared with protocolised sedation alone.
Capuzzo 2001 *Italy*	Patients were divided into 3 groups based on types of sedatives received. Patients receiving morphine plus other sedatives were more deeply sedated than patients receiving either no morphine plus maximum 2 doses of benzodiazepines or only morphine (no morphine: A, only morphine: B: 14±9 mg/d (morphine), morphine and other sedatives: C: 15±9 mg/d (morphine), 176±757 mg/d (propofol)).
Costa 2014 *Brazil*	Based on RASS scores over at least 24 hours in ICU, patients were retrospectively classified as either mild-moderate (A: RASS -2 and -3), deep (B: RASS -4 and -5) or no sedation (C). Patients in the deep sedation vs mild-moderate sedation group received total midazolam 2933±4724 vs 482±720 mg, p = 0.078 and fentanyl 33.7±56.5 vs 7.3±13.2 mg, p = 0.112.
Dale 2014 *USA*	Hourly (0.15±0.011 vs 0.23±0.018 mg, p<0.01) and total benzodiazepine doses (17.2±53.6 vs 49.2±156.5 mg, p<0.01) were lower in the sedation-reducing intervention period, compared with the observation period. 24 hour weighted average RASS was higher (lighter) (-0.99±0.023 vs -1.30±0.026) in the intervention period.
Faust 2016 *USA*	Following the implementation of the sedation protocol, median RASS scores were higher (-1.25[-2.3 to -0.40] vs -2.57[-3.23 to -1.40], p=0.001), compared to the control period, and percent of RASS scores between -3 and -5 in the first 24 hrs was lower (27.3±37.3% vs 46.8±46.9%, p=0.006), indicating lighter sedation.
Guttormson 2011 *USA*	All patients had a Sedation Intensity Score calculated based on total dose of all sedatives administered. A mathematical model was used to identify factors associated with sedation intensity and patients were divided into two groups based on increasing or decreasing sedation intensity. MAAS was also used to assess arousability.
Khan 2014 *USA*	Compared with the pre-implementation phase, in the post-implementation phase RASS scores on weekdays (median: -4, increase of 0.88 post-protocol, p<0.0001) and weekends (median: -5, increase of 1.2 post-protocol, p<0.0001) were higher, reflecting lighter sedation in this group.
Mehta 2012 *Canada and USA*	Patients in the protocolised sedation with DIS group received higher infusion and lower bolus doses of midazolam (daily doses: 102±326 vs 82±287 mg, p=0.04; infusions: 101±325 vs 82±287 mg, p=0.03; bolus: 0.49±2.65 vs 0.99±5.9 mg, p<0.001) and fentanyl (daily doses: 1070±2066 vs 1780±4135 µg, p<0.001; infusions: 984±2002 vs 1664±4070 µg; bolus: 86±169 vs 116±215 µg, p<0.001) than patients in the protocolised sedation group alone, reflecting deeper sedation in the protocolised sedation + DIS group.
Mendes 2008 *Brazil*	Depth of sedation was assessed in all patients with RASS and Ramsay sedation score with mean sedation levels calculated for each patient (in whole cohort, mean dose of midazolam: 1.7 mg/kg/day (range 0.12-10.89); mean dose of fentanyl: 15.73 µg/kg/day (range 4.00-68.84).
Nassar Junior 2014 *Brazil*	Compared with patients in the DIS group, patients in the intermittent sedation group had higher SAS scores (3.6[3.4-4.0] vs 3.2[2.6-3.7], p=0.035) and less midazolam (0[0.0-0.05] vs 45[0,201] mg, p<0.001) and fentanyl (300[100-1520] vs 1500[520-4215] µg, p=0.004) use, reflecting less sedation in this group.
Olsen 2020	Patients in the light sedation group were more heavily sedated (RASS day 1: -2.3; day 7: -1.8) and received more sedation [Midazolam mg/kg/hr (day 2 – 28): 0.000187 (0 – 0.003410); Propofol mg/kg/hr (day 1 – 2): 0.84 (0.29-1.2); Propofol mg/kg/hr (day 3 – 28): 0.0064 (0 – 0.034)] than patients in the non-sedation group [RASS day 1: -1.3; day 7: -0.8; Midazolam mg/kg/hr (day 2 – 28): 0 (0 – 0.000005), NS; Propofol mg/kg/hr (day 1 – 2): 0.22 (0- 0.054), Diff: -0.62 (-0.72; -0.53); Propofol mg/kg/hr (day 3 – 28): 0 (0 – 0.013), Diff: -0.0063(-0.874; -0.0037)].

Quenot 2007 *France*	In the intervention group, compared with the control group, daily doses of midazolam (44±31 vs 92±59 mg, p=0.001) and propofol (1840±750 vs 2900±1400 mg, p=0.01) were lower, reflecting less sedation in this group.
Ren 2017 *China*	Doses of sufentanil (0.018±0.009 vs 0.030±0.007 mg/kg/hr, p<0.0001) and midazolam (0.017±0.009 vs 0.029±0.007 mg/kg/hr, p<0.0001) were lower after implementation of the ABCDE bundle, compared to before, reflecting less sedation in this group.
Samuelson 2006 *Sweden*	Depth of sedation was reported as proportion of MAAS between 0-2 (deep sedation), 3 (awake, calm and cooperative) and 4-6 (agitated).
Samuelson 2007 *Sweden*	Depth of sedation was reported as proportion of MAAS 3 (awake, calm and cooperative). Patients were categorised as those who experienced each of the outcomes of: (A) bothered vs not bothered by at least one ICU experience, (B) bothered vs not bothered by nightmares and (C) with memories vs no memories of ETT.
Samuelson 2008 *Sweden*	Patients divided into two groups according to sedation depth- light sedation (MAAS 3 – 4: median 3.0[0.0]), heavy sedation: (MAAS 1 – 2: median 1.25[1.0]).
Sen 2017 *USA*	Compared with the control (pre-protocol group), patients in the protocol group received less total benzodiazepine dose (450±701 vs 74±159 mg, p<0.01), reflecting less sedation in this group.
Shehabi 2018 *Aus, NZ, Malaysia, Singapore*	RASS measurements taken frequently in all patients and 'sedation index' calculated – this is measure of intensity of sedation on a continuous scale for each patient over the first 48 hours. Multivariate analysis to determine the impact of sedation index on various outcomes.
Shehabi 2013 *Australia and New Zealand*	Patients in the early goal directed sedation group received more dexmedetomidine (36.55[16.38-13.23] vs 20.58[20.58-20.58] µg, p<0.0001), but less midazolam (0.06[0.02-1] vs 0.3[0.23-0.76] mg, p=0.036) and propofol (9.89[2.41-22.51] vs 33.55[13.54-77.07] mg, p=0.046) and had more RASS assessments in the light sedation range (66% vs 38%, p=0.01) compared with the control group, reflecting less deeply sedated patients in this group.
Shehabi 2012 *Australia and New Zealand*	RASS measurements taken frequently in all patients and data divided into light (RASS -2 to +1) or deep sedation (RASS -3 to -5), or agitation (RASS +2 to +4); (n=1642[61.9%], 942[35.5%], 72[2.7%], assessments respectively). Cumulative dose of midazolam and fentanyl was also analysed.
Strøm 2011 *Denmark*	Compared with the DIS group, patients in the no sedation group received less propofol (0[0-1.26] vs 1.40[0.52-2.04] mg/kg/hr, p=0.013) and midazolam (0[0-0] vs 0.01[0-0.04] mg/kg/hr, p=0.003), reflecting less sedation in this group.
Strøm 2010 *Denmark*	Compared with the DIS group, patients in the no sedation group received less propofol (0[0-0.52] vs 0.77[0.15-1.65] mg/kg/hr, p=0.0001) and midazolam (0[0-0] vs 0.003[0-0.024] mg/kg/hr, p<0.0001), reflecting less sedation in this group.
Treggiari 2009 *Switzerland*	Patients divided into two groups according to sedation depth- light sedation: Ramsay sedation score 1 – 2, heavy sedation: Ramsay sedation score 3 – 4. Daily median Ramsay ranged from: light sedation: 1(1-2) to 3(1-3); deep sedation: 3(2-4.5) to 4(3-5). Daily midazolam ranged from: light sedation: 3.0 ±5.0 to 11.7±23.2 mg; deep sedation: 24.2 ±45.1 to 95.3±124.5 mg.

Abbreviations: DIS: Daily Interruption of Sedation, ETT: Endotracheal Tube, ICU: Intensive Care Unit, MAAS: Motor Activity Assessment Scale, RASS: Richmond Agitation Sedation Scale, SAS: Riker Sedation Agitation Scale.

Supplementary Table 3: Risk of bias assessment of randomised controlled trials

Study	Risk of selection bias r/t random allocation generation	Risk of selection bias r/t allocation concealment	Risk of performance bias r/t blinding of participants & personnel	Risk of detection bias r/t blinding of outcome assessment	Risk of attrition bias r/t incomplete outcome data	Risk of reporting bias r/t selective reporting
Mehta et al. 2012 *Canada and USA*	Low Random allocation to groups	Low Automated telephone system used to randomise patients	High Not blinded	High – time to extubation, ICU LOS Hospital LOS, delirium reintubation, tracheostomies – not blinded Low - ICU mortality, hospital mortality – objective outcomes	Low Outcome data appears complete, good follow-up	Low All data appear reported
Nassar Junior & Park. 2014 *Brazil*	Low Random allocation to groups	Low Allocation concealment was achieved with sealed envelopes	High Intervention was DIS vs no-DIS so patients and personnel not blinded	High – VFD to day 28 ICU LOS, hospital LOS delirium, self-extubation, reintubation, tracheostomies – not blinded Low - ICU mortality, hospital mortality (objective outcomes); psychological distress (patient self-report)	Low Outcome data complete for all randomised subjects	Low All data appear reported
Olsen et al. 2020 Denmark, Norway & Sweden	Low Random allocation to groups	Low Computer generated assignment sequence with variable block sizes	High Intervention was no sedation vs light sedation, personnel not blinded	Low – mortality; acute kidney injury High – number of days free from coma or delirium; number of ICU-free days, number of ventilator-free days	Low Information provided for all randomised participants; 99% follow-up	Low All data appear reported
Samuelson et al. 2008 *Sweden*	Low Random allocation to groups	Low Allocation concealment was achieved with sealed envelopes	High Patients blinded (as sedated). Nurses not blinded as administering different sedation	High – MV duration, delirium, reinbutation, tracheotomies – not blinded Low – memories, stressful experiences, psychological distress – patient self-report	Low Data appears complete-negligible drop-outs and good follow-up	Low Reported data are relevant, qualitative interviews and questionnaires likely to yield far more data than what is reported

Shehabi et al. 2013 *Australia and New Zealand* *EGDS*	Unclear Method of selecting patients unclear- block randomisation was used	Low Allocation concealment was achieved with sealed envelopes	High Patients effectively blinded as sedated. Not possible to blind personnel as intervention had to be delivered	High – VFD to day 28, ICU LOS, hospital LOS, delirium – not blinded Low – hospital mortality – objective measure	Low Outcome data are complete, all randomized patients included	Low All outcomes are reported
Strøm et al. 2011 *Denmark*	Low Random allocation to groups, block randomisation	Low Allocation concealment was achieved with sealed envelopes	High Patients and personnel were not blinded	Low – interviewer blinded to intervention patient received and all outcomes patient self-report	High Similar numbers in each group, however drop out of ~40%	Low Reported data are relevant, interviews and questionnaires likely to yield far more data than what is reported
Strøm et al. 2010 *Denmark*	Low Random allocation to groups, block randomisation	Low Allocation concealment was achieved with sealed envelopes	High Patients and personnel were not blinded	High - Ventilator free days to day 28, ICU LOS hospital LOS, VAP – not blinded Low – ICU mortality hospital mortality – objective measures	Low Outcome data are complete, good follow-up, ITT analysis	Low All outcomes are reported
Treggiari et al. 2009 *Switzerland*	Low Random allocation to groups, computer generated sequence of random numbers	Unclear Allocation concealment was used, but method not reported	High Patients not blinded as intervention is sedation/no sedation, personnel not blinded as they have to administer sedation	High - MV duration ICU LOS, hospital LOS self-extubation, reintubation, tracheostomies Low – anxiety, depression, psychological distress, post-traumatic stress (patient self-report and outcome assessor was blinded to group allocation); ICU mortality, hospital mortality (objective measures)	Low Outcome data are complete, good follow-up	Low All outcomes are reported

Supplementary Table 4: Risk of bias assessment of cohort studies

Study	Risk of selection bias r/t method of recruitment	Risk of bias r/t measurement of depth of sedation	Risk of detection bias r/t outcome assessments	Risk of bias r/t important confounding factors	Risk of attrition bias r/t length & completeness of follow-up	Risk of potential bias due to other sources
Arabi et al. 2007 *Saudi Arabia*	Low Consecutive recruitment	Low SAS and daily doses of analgesics and sedatives for each patient	High – MV duration, ICU LOS – not blinded Low – VAP – monitored independently	High – baseline differences examined, but not integrated into analyses	Unclear Only final patient numbers reported- unclear if outcome data is complete; follow-up limited to hospital	High Education given before and throughout study period (i.e. in the pre & post period) despite it being part of the 'intervention'
Balzer et al. 2015 *Germany*	Low All patients admitted to study ICU over relevant timeframe were included retrospectively, then groups were determined based on sedation depth	Low RASS - measurements converted to single continuous variables by calculating the ratio of RASS measurements ≤−3 and the total number of RASS measurements.	High - time to extubation, ICU LOS hospital LOS, delirium – not blinded Low - ICU mortality, hospital mortality – objective measures	Unclear - Cox regression adjusted for relevant factors, unclear if these factors were identified *a priori*	Low Retrospective study - all patients included in the analysis. Good follow-up to 2 years	Unclear Only daily RASS measurements, with the timing not standardised
Bugedo et al. 2013 *Chile*	Unclear Method of recruitment not reported	Low SAS and mean doses of sedatives	High - MV duration, VFD to day 28, ICU LOS, hospital LOS, self-extubation, reintubation, tracheostomies – not blinded Low – memories, post-traumatic stress – patient self-report using validated instrument	High - baseline demographic and clinical characteristics not incorporated into multivariable analysis	High In hospital outcome data appears complete. Follow-up at 1 year is only 52%	N/A
Burry et al. 2015 USA & Canada	Unclear Unclear whether consecutive patients were recruited	Low SAS and daily doses of midazolam and fentanyl	Low – memories – patients interviewed by research personnel using validated instrument	Low - multivariable analysis with factors identified *a priori*	High 35% follow-up at 90 days	N/A
Capuzzo et al. 2001 *Italy*	Low Consecutive recruitment	High Average dose of morphine and propofol ; no information about sedation assessment	High - memories measured using local developed questions	Unclear – not clear what factors were incorporated into multivariable analysis or	Low Outcome data appears complete. 6 month follow up ~75%	N/A

				how these were determined		
Costa et al. 2014 *Brazil*	Low Consecutive recruitment	High Appears either RASS or Ramsay scale was used – no information about conversion to RASS	High – memories assessed using translated and locally adapted version of ICU Memory tool without validity testing Unclear – anxiety, depression, PTSD – no information provided regarding measurement	High – no multivariable analysis conducted	High 3 month follow-up - 46%	N/A
Dale et al. 2014 *USA*	Low Consecutive recruitment	Low RASS and hourly and daily doses of sedatives	High – MV duration, ICU LOS, hospital LOS, delirium – not blinded Low – VAP (assessed by infection prevention team), hospital mortality (objective measure)	Low - factors identified *a priori* and incorporated into multivariable analyses	Low Outcome data are available for all eligible admissions in the study period	N/A
Faust et al. 2016 *USA*	Low All eligible admissions in study period included	Low RASS and doses of sedatives	High – MV duration, VFD to day 28, ICU LOS, self-extubation, reintubation, tracheostomies Low – hospital mortality – objective measure	Unclear – relevant factors incorporated into multivariable analyses, unclear if these factors were identified *a priori*	Low Outcome data are available for all eligible admissions in the study period	N/A
Guttormson. 2011 *USA*	High Recruitment methods not specifically stated, but seems like individual patients were identified and approached	High Motor Activity Assessment Scale (MAAS) and Sedation Intensity Score not previously validated	High – MV duration, ICU LOS – not blinded Low - memories – self-report using validated instrument	Unclear Multivariable analysis incorporating relevant factors was conducted but it is unclear if these factors were identified *a priori*	High Follow-up at final interview only ~35%	Unclear Inclusion criteria modified during study to increase recruitment.
Khan et al. 2014 *USA*	Low Consecutive recruitment	Low RASS	High – MV duration, hospital LOS, delirium-not blinded Low - hospital mortality – objective measure	Unclear – relevant factors incorporated into multivariable analyses, unclear if these factors were identified *a priori*	Low Follow-up data appears complete	N/A
Mendes et al. 2008 *Brazil*	Unclear Recruitment method not reported	Low Both RASS or Ramsay scales assessed for each	High – MV duration, ICU LOS – not blinded	High – no multivariable analyses	Low Outcome data are available for all eligible	N/A

		patient; daily dose of sedatives	Low – ICU mortality – objective measure		admissions in the study period	
Quenot et al. 2007 *France*	Low Consecutive recruitment	High Cambridge score and daily doses of midazolam and propofol	High – MV duration, ICU LOS, hospital LOS, VAP – not blinded Low – ICU mortality, hospital mortality, self-extubation	Unclear - relevant variables entered into Cox proportional hazards model, unclear if these variables were identified *a priori*	Low Follow-up data appears complete, hospital only	N/A
Ren et al. 2017 *China*	Unclear Recruitment method not reported	High Total dose and average dose/hour of sufentanil and midazolam, no information about sedation assessment	High – MV duration, ICU LOS, delirium – not blinded	High - no multivariable analyses	Low Outcome data are available for all eligible admissions in the study period	N/A
Samuelson et al. 2006 *Sweden*	Low Consecutive recruitment	High Motor Activity Assessment Scale (MAAS)	Low - memories assessed through self-report using validated instruments	Unclear - theoretically important variables (ICU LOS, severity of illness, age) included in multivariable analysis, unclear if variables were identified *a priori*, delirium does not appear to have been included in analysis	Low 80% follow-up for interview.	N/A
Samuelson et al. 2007 *Sweden*	Low Consecutive recruitment	High Motor Activity Assessment Scale (MAAS)	Low – memories & stressful experiences assessed through self-report using validated instruments	Unclear - theoretically important variables (ICU LOS, severity of illness, age) included in multivariable analysis, unclear if variables were identified a priori, delirium does not appear to have been included in analysis	Low 80% follow-up for interview.	N/A
Sen et al. 2017 *USA*	Low Pre/post study, consecutive recruitment	Low SAS and total benzodiazepine exposure	High – MV duration, ICU LOS, hospital LOS – not blinded Low – ICU mortality – objective assessment	Unclear – not clear if factors incorporated into multivariable analyses were identified *a priori*	Low Data are complete for all subjects randomised in the study. Hospital only follow-up	High Study groups mismatched in number, baseline characteristics and comparative analysis did not account for differences

Shehabi et al. 2018 *Australia, New Zealand, Malaysia, Singapore*	Unclear Method of selecting patients - "over 3 mths each ICU recruited up to a maximum of 20 patients"	Unclear RASS converted to sedation intensity score – score not previously validated	High - time to extubation, delirium – not blinded Low - 180 day mortality – objective measure	Unclear - relevant factors incorporated into multivariable analyses, but not clear if these factors were identified *a priori*	Low Outcomes are within hospital and complete, except for 180 day mortality which is complete	N/A
Shehabi et al. 2012 *Australia and New Zealand* *AJRCCM*	Unclear Method of selecting patients - "over 3 mths each ICU recruited up to a maximum of 20 patients"	Low RASS	High - VFD to day 28, ICU LOS, hospital LOS, delirium – not blinded Low - hospital & 180 day mortality – objective measure	Low Relevant factors identified *a priori* and incorporated into multivariable analyses	Low Outcomes are within hospital and complete, except for 180 day mortality which is complete	Unclear Delirium assessment was only conducted daily with no standardisation of timing

RASS or SAS were the only sedation instruments considered acceptable; * MV outcome assessments including MV duration, ventilator free days to day 28, time to extubation

Critical Appraisal criteria

Selection bias related to method of recruitment – low risk of bias was considered to exist when consecutive patients or all eligible patients over a specified time frame were included in the study

Bias related to measurement of depth of sedation – low risk of bias was considered to exist when sedation assessments were conducted using either RASS or SAS; this may or may not have been supported by average doses of sedative medications provided for each patient.

Detection bias related to outcome assessments – low risk of bias was considered to exist when outcome assessments were performed by personnel blinded to group allocation, unless the outcome was objective (e.g. mortality).

Related to important confounding factors – low risk of bias was considered to exist when multivariable analysis incorporating relevant factors identified a priori was conducted; factors to include age, illness severity, depth of sedation plus delirium for post-hospital psychological, quality of life or memory outcomes

Attrition bias related to length and completeness of follow-up – low risk of bias was considered to exist when complete hospital data were provided and/or >70% of participants retained in study depending on what outcomes were analysed in the study.

Supplementary Table 5 Results identified in studies included in systematic review

Study	Evidence of differing depth of sedation	Outcomes					
		Ventilation Outcomes	**Length of stay (ICU/Hospital)**	**Mortality (ICU/Hospital)**	**Neurological (Coma, Delirium)**	**Psychological**	**Adverse Events and Other Outcomes**
Arabi Y, et al. 2007 *Saudi Arabia*	Less sedation in education period compared to no education period (daily propofol: Group (G)1: 236±65; G2 264±106; G3: 1710±84; G4 105±33 mcg, p=0.10 all groups; p=0.03 G1&G2 vs G3&G4]. More patients in ideal SAS range on nights 3 & 4 and day 4 in all groups compared to group 1 (Day 4 SAS 3-4: (G1: 15%; G2 30%; G3: 45%; G4 50%, p=0.012.	**No difference in duration of MV** between the groups (G1: 12±2; G2 11±1; G3: 10±1; G4 8±1 days, p=0.21)	**No difference in either ICU** (G1: 13±2; G2 13±1; G3: 12±1; G4 10±1 days, p=0.42) **or hospital LOS** (G1: 50±7; G2 55±8; G3: 41±7; G4 40±6 days, p=0.34) between the groups.	**No difference in either ICU** (G1: n=10[20%]; G2 9[18%]; G3: 12[23%]; G4 7[13%], p=0.64) **or hospital mortality** (G1: n=12[24%]; G2 12[24%]; G3: 19[36%]; G4 12[23%], p=0.35) between the groups.	N/A	N/A	**Incidence of VAP reduced with both protocol and education periods** (G1: n=14[28%]; G2 15[29%]; G3: 6[11%]; G4 6[11%], p=0.002). **No difference in incidence of tracheostomy** (G1: n=11[22%]; G2 15[29%]; G3 12[23%]; 8[15%], P=0.23) between the groups.
Balzer F, et al. 2015 *Germany*	Patients were divided into two groups, based on sedation depth (RASS) (deeply sedated - >85% of documented RASS scores during the study period ≤-3). In addition, deeply sedated patients had a lower first RASS (-5[-5 to -4] vs -4[-5 to -1], p< 0.001) and took longer to reach a first RASS >-3 (79 [52-141] vs 11 [5-20] hrs, p= 0.001).	**Time to extubation was longer in the more deeply sedated patients** (75[37-156] vs 17[8-33] hrs, p<0.001).	**ICU** (21[12-38] vs 8[5-16] days, p<0.001) and **hospital LOS** 28[16-48] vs 18[12-33] days, p<0.001) was **longer in the more deeply sedated patients** compared to those not deeply sedated.	**ICU** (n=137[27%] vs 67[5%], p<0.001), **hospital** (n=175[34%] vs 131[10%], p<0.001), **and 2 year** (n=222[62%] vs 307[32%], p<0.001), **mortality was higher in the more deeply sedated patients**.	**Incidence of delirium** was higher in the deeply sedated group (42[8%] vs 445[33%], p<0.001) in unmatched analysis, but this **difference was not present in matched analysis** (deeply sedated – 215[42%] vs 213[42%], p=0.899).	N/A	**Proportion of patients receiving haemodialysis was higher in the more deeply sedated patients.** The same outcomes (time to extubation, ICU & hospital LOS and mortality) were examined in a sub-group of 1020 patients matched on APACHE II and type of admission, with similar differences identified.
Bugedo G, et al. 2013 *Chile*	Midazolam decreased in the intervention period compared with the control period	**No difference in duration of MV** (8[4-13] vs 7[4-15.5] days,	**No difference in either ICU** (10[6-15] vs 11[6-18]	**No difference in 28 day** (n=57[37%] vs 45[34%], p=0.636)	N/A	**No difference in incidence of nightmares** (n=	**No difference in incidence self-extubations**

	(average rate - control: 0.03[0.01-0.06] vs intervention: 0.01[0-0.03], p<0.001); fentanyl dose increased in intervention period (average rate 0.6[0.1-1.4] vs 1.5[0.8-2.4], p<0.001). The proportion of SAS scores in the deep sedation range (SAS 1 – 2) was lower (55% vs 44%) while the proportion of SAS scores in the ideal range (SAS 3 – 4) was higher (37% vs 49%, p=0.001) in the intervention period.	p=0.934) **or ventilator free days** to day 28 (8[0-23] vs 12[0-24] days, p=0.430) between the groups.	days, p=0.457) or **hospital** (18[10-33] vs 18[10-31] days, p=0.795) **LOS** between the groups.	**or 1 year** (77[50%] vs 65[49%], p=0.941) **mortality** between the groups.		22[55%] vs 15[43%], p=0.294), **severe anxiety or panic** (n= 16[40%] vs 12[34%], p=0.610) or **pain** (n= 12[30%] vs 13[37%], p=0.5130, **feelings of suffocation or PTSS-10** (28[19-3] vs 26[17-38], p=0.840) questionnaire scores between the groups.	(n=14[9.0%] vs 12[9.1%], p=0.98), **reintubation within 48 hours** (n=8/116[6.9%] vs 7/102[6.9%], p=0.98), **tracheostomy** (n=12[7.7%] vs 6[4.5%], p=0.27) or **central catheter or nasogastric tube displacement** between the groups.
Burry L, et al. 2015 *USA & Canada*	Higher doses of midazolam (dose/patient/d: 97.0±200.8 vs 64.7±245.8 mg, p<0.0001) and fentanyl (dose/patient/d: 1.9±3.5 vs 1.1±2.0 mg, p<0.0001) in the protocolised sedation with DIS group compared with protocolised sedation alone.	**No difference in time to successful extubation** between the groups (6[4-13] vs 6[3-11] days, p=0.16).	**No difference in either ICU** (10[7-20] vs 9[5-16] **days**, P=0.22) **or hospital** (20.5[13-47] vs 22[13-44] **days**, p=0.59) **LOS** between the groups.	**No difference in hospital mortality** between the groups (n=15[10.3%] vs 11[7.7%], p=0.44).	**No difference in incidence of delirium** (n=82[56.3%] vs 80[55.9%], p=0.97) **or coma** (n=37[25.3%] vs 37[25.9%], p=0.92) between the groups	**Patients who reported 'not remembering the ICU' had less sedation** (average daily midazolam dose), but no difference in SAS scores. In a multivariate model, **total midazolam and fentanyl exposure above the mean (deeper sedation) was associated with increased risk of delusional memories.**	**No difference in incidence of tracheostomy** (n=38[26.8%] vs 32[22.5], p=0.49) **or central line, endotracheal tube, gastric tube or urinary catheter removal** between the groups.
Capuzzo M, et al. 2001 *Italy*	Patients were divided into 3 groups based on types of sedatives received. Patients receiving morphine plus other sedatives were more deeply sedated than patients receiving either no morphine	**Duration of MV was longer in the more deeply sedated patients** (A: 1.2±1.8; B: 1.5±2.0; C: 7.2±11.1 days, p<0.001).	**ICU** (A: 3.2±2.8; B: 3.3±3.2; C: 15.3±26.5 days, p<0.001) and **hospital LOS** (A: 11.6±8.2; B: 13.5±10.5; C:	N/A	N/A	**No difference in recall of factual** (A: n=16[36%]; B: 29[34%]; C: 4[18%], p=NS), **sensation** (A: n=4[9%]; B: 13[15%]; C: 3[14%],	N/A

	plus maximum 2 doses of benzodiazepines or only morphine (no morphine: A, only morphine: B: 14±9 mg/d (morphine), morphine and other sedatives: C: 15±9 mg/d (morphine), 176±757 mg/d (propofol)).		32.5±32.9 days, p<0.001) **was longer in the more deeply sedated patients.**			p=NS) or **emotional** (A: n=4[9%]; B: 6[7%]; C: 4[18%], p=NS) **memories** of ICU between the groups.	
Costa JB, et al. 2014 *Brazil*	Based on RASS scores over at least 24 hours in ICU, patients were retrospectively classified as either mild-moderate (A: RASS -2 and -3), deep (B: RASS -4 and -5) or no sedation (C).	The **number of patients requiring MV for >2 days was highest in the deeply sedated group** (A: n=4/12[33.3%], B: 66/74[89.2%], C: 2/42[4.8%], p<0.001).	**Number of patients with ICU LOS >7 days was highest in the deeply sedated patients** (A: n=3[25.0%]) B: 30/74[40.5%], C: 3/42[2.4%], p<0.001). **No difference in hospital LOS** between the groups (A: n=4/12[33.4%], B: 32/74[43.2%], C: 18/42[42.9%], p=0.950).	N/A	N/A	**No difference in the incidence depression** (A: n=1[8.3%], B: 10[13.5%], C: 5[11.9%], p=0.458) in the more deeply sedated patients.	N/A
Dale CR, et al. 2014 *USA*	Hourly (0.15±0.011 vs 0.23±0.018 mg, p<0.01) and total benzodiazepine doses (17.2±53.6 vs 49.2±156.5 mg, p<0.01) were lower in the sedation-reducing intervention period, compared with the observation period. 24 hour weighted average RASS was higher (lighter) (-0.99±0.023 vs -1.30±0.026) in the intervention period.	**Duration of MV was longer in the more deeply sedated patients** (20[7-16] vs 16[6-44] days, p=0.01).	**ICU** (3[1-7] vs 3[1-6] days, p=0.03) **and hospital** (11[5-21] vs 10[4-18] days, p=0.02) **LOS was longer in the more deeply sedated patients.**	**No difference in hospital mortality** (n=96[13.8%] vs 107[13.7%], p=1.00) between the groups; **1 day increase in the median number of ventilator free survival days at 28 days** (25[17-26] vs 26[20-26], p<0.01).	If all ICU stay is used: **Incidence** (n=176[22.6%] vs 75[10.7%], p<0.01), **and total number of days** (n=455[21.2%] vs 172[25.1%], p<0.01) **of positive CAM-ICU scores was higher in the lightly sedated group.** If only periods in which the CAM-ICU score was measured are used: delirium	N/A	**No difference in incidence of VAP** (n=46[6.5%] vs 36[4.6%], p=0.08) between the groups.

					incidence decreased by 3.9% (21.2 vs 25.1, p<0.01) in lightly sedated group.		
Faust AC, et al. 2016 *USA*	Following the implementation of the sedation protocol, median RASS scores were higher (-1.25[-2.3 to -0.40] vs -2.57[-3.23 to -1.40], p=0.001), compared to the control period, and percent of RASS scores between -3 and -5 in the first 24 hrs was lower (27.3±37.3% vs 46.8±46.9%], p=0.006), indicating lighter sedation.	**Duration of MV was longer** (138.3±132.6 vs 92.9±73.3 hrs, p=0.01) **in the more deeply sedated patients**, but there was **no difference in 28 day ventilator free days** (23.6±4.9 vs 24.1±3.1 days, p=0.47) between the groups.	**ICU LOS was longer** (211.5±164.3 vs 160.7±125.7 hrs, p=0.038) **in the more deeply sedated patients.**	**No difference in hospital mortality** (n=22[33.8%] vs 24[30%], p=0.72) between the groups.	N/A	N/A	**No difference in incidence of self-extubation** (n=5[5.9%] vs 2[3.0%], p=0.46), **reintubation within 24 hours** (n=2[40%] vs 1[50%], p=1.00) or **tracheostomy** (8[10.1%] vs 11[16.9%], p=0.32) between the groups.
Guttormson JL. 2011 *USA*	All patients had a Sedation Intensity Score calculated based on total dose of all sedatives administered. A mathematical model was used to identify factors associated with sedation intensity and patients were divided into two groups based on increasing or decreasing sedation intensity. MAAS was also used to assess arousability.	**Number of ventilator days was higher** (10.8[6.0-12.5] vs 4.1[2-5.9] days, p=0.006) with deeper sedation.	**ICU LOS** (14.5[10.8-29.4] vs 5.1[3.7-9.1], p=0.001) **was longer with deeper sedation.**	N/A	N/A	**No difference in satisfaction of ICU experience, memories of frightening experiences, negative feelings, awareness or factual memories** between the groups. **Incidence of delusional memories was higher with deeper sedation.**	N/A
Khan BA, et al. 2014 *USA*	Compared with the pre-implementation phase, in the post-implementation phase RASS scores on weekdays (median: -4, increase of 0.88 post-protocol, p<0.0001) and weekends (median: -5, increase of 1.2 post-protocol, p<0.0001) were higher,	**Duration of MV was shorter** (median: 4 vs 5 days, p<0.01) in the more deeply sedated patients	**No difference in hospital LOS** (median: 14 vs 14 days, p=0.56) between the groups.	**No difference in hospital mortality** (19.5% vs 19.6%, p=0.97) between the groups.	**Prevalence of delirium** (n=94[66.7%] vs 167[55.3%], p=0.02) and **acute brain dysfunction** (coma + delirium) (n=238[90.8%] vs 374[85%], p=0.02] were h**igher in the**	N/A	N/A

	reflecting lighter sedation in this group.				**more deeply sedated patients,** with **no difference in incidence of delirium** (n=14[23.0%] vs 33[19.6%], p=0.58), **or coma** (n=205[78.2%] vs 323[73.4%], p=0.15), between the groups.		
Mehta S, et al. 2012 *Canada and USA*	Patients in the protocolised sedation with DIS group received higher infusion and lower bolus doses of midazolam (daily doses: 102±326 vs 82±287 mg, p=0.04; infusions: 101±325 vs 82±287 mg, p=0.03; bolus: 0.49±2.65 vs 0.99±5.9 mg, p<0.001) and fentanyl (daily doses: 1070±2066 vs 1780±4135 µg, p<0.001; infusions: 984±2002 vs 1664±4070 µg; bolus: 86±169 vs 116±215 µg, p<0.001) than patients in the protocolised sedation group alone, reflecting deeper sedation in the protocolised sedation + DIS group.	**No difference in days to successful extubation** (7[3-12] vs 7[4-13] days, p=0.52) between the groups.	**No difference in ICU** (10[6-20] vs 10[5-17] days, p=0.36) **or hospital** (20[10-48] vs 20[10-36] days, p=0.42) **LOS** between the groups.	**No difference in ICU** (n=52[24.9%] vs 50[23.4%], p=0.720 **or hospital** (n=63[30.1%] vs 63[29.6%], p=0.89) **mortality** between the groups.	**No difference in incidence of delirium** (n=113[54.1%] vs 113[53.3%], p=0.83) between the groups.	N/A	**No difference in incidence of ARDS, number of patients requiring vasopressors/inotropes/renal replacement or NMB, gastric tube, ETT, urinary catheter, central venous or arterial catheter removal, use of physical restraints, reintubation in 48 hours** (n=16[7.7%] vs 12[5.6%], p=0.39) or **tracheostomy** (n=54[26.3%] vs 49[23.2%], p=0.46) between the groups.
Mendes CL, et al. 2008 *Brazil*	Depth of sedation was assessed in all patients with RASS and Ramsay sedation score (in whole cohort, mean dose of midazolam: 1.7 mg/kg/day (range 0.12-10.89); mean dose of fentanyl: 15.73 µg/kg/day (range 4.00-68.84).	N/A	N/A	**There was a positive correlation between sedation dose and ICU mortality** (AUC for RASS -4 to -5: 0.803), **and between adequate sedation and survival** (AUC for	N/A	N/A	N/A

				RASS 0 to -3: 0.819).			
Nassar Junior AP, Park M. 2014 *Brazil*	Compared with patients in the DIS group, patients in the intermittent sedation group had higher SAS scores (3.6[3.4-4.0] vs 3.2[2.6-3.7], p=0.035) and less midazolam (0[0.0-0.05] vs 45[0,201] mg, p<0.001) and fentanyl (300[100-1520] vs 1500[520-4215], p=0.004) use, reflecting less sedation in this group.	**No difference in 28-day ventilator free days** (25[21-27] vs 24[0-26] days, p=0.160) between the groups.	**No difference in either ICU** (11[6-16] vs 8[5-19] days, p=0.595) or **hospital** (22[13-38] vs 15[9-28] days, p=0.099) **LOS** between the groups.	**No difference in either ICU** (n=7 [23%] vs 12[40%], p=0.165) or **hospital** (n=9[30%] vs 13[43.3%] days, p=0.284) **mortality** between the groups.	**No difference in the incidence of delirium** (n=12[40%] vs 9[30%], p=0.472) between the groups.	**No difference in the level of psychological stress** (22[8-31] vs 16[4-34], p=0.750) at 6 months between the groups.	**No difference in the incidence of reintubation** (n=1[3%] vs 4[13%], p=0.161), **self-extubation** (n=2[7%] vs 1[3%], p=0.514), **accidental removal of catheters or tracheostomy** (n=1[3%] vs 1[3%], p=1.00) between the groups.
Olsen et al. 2020 *Denmark, Norway & Sweden*	Patients in the light sedation group were more heavily sedated compared with patients in the non-sedation group (RASS day 1: -2.3 vs -1.3; day 7: -1.8 vs -0.8) and received more sedation [Midazolam mg/kg/hr (day 2 – 28): 0.000187 (0 – 0.003410) vs 0 (0 – 0.000005), NS; Propofol mg/kg/hr (day 1 – 2): 0.84 (0.29-1.2) vs 0.22 (0-0.054), Diff: -0.62 (-0.72; -0.53); Propofol mg/kg/hr (day 3 – 28): 0.0064 (0 – 0.034) vs 0 (0 – 0.013), Diff: -0.0063(-0.874; -0.0037).	No difference in ventilator free days to 28) [19 (0 – 25) vs 20 (0 – 26) days, 95% CI: 1 (-3 to 3) days].	No difference in ICU LOS (censored at day 28) [14 (0 – 23) vs 13 (0 – 23) days, 95% CI: -1 (-7 to 4) days].	N/A	Patients in the light sedation group had fewer delirium or coma free days to day 28 [26 (22 - 28) vs 27 (21 – 28), 95% CI: 1 (0 – 2) days]	N/A	**No difference in reintubations within 1 hr** [1(0.3%) vs 4 (1.1%)] **or 24 hrs of self-extubation** [14(4%) vs 31(8.9%)]. **No difference in self-removal of:** - **central lines** [3(0.9%) vs 3(0.9%) - **peripheral lines** [10(2.8%) vs 9(2.6%)] - **other equipment** [32(9.1%) vs 53(15.2%) **No difference in 90 day all cause mortality** [130 (37.0%) vs 148 (42.4%), 95% CI: 5.4% (-2.2 to 12.2, p = 0.65] **Light sedation patients had more major thromboembolic events** [10(2.8%) vs 1(0.3%), 95% CI: -2.5 (-4.8 to -0.7).

Quenot J, et al. 2007 *France*	In the intervention group, compared with the control group, daily doses of midazolam (44±31 vs 92±59 mg, p=0.001) and propofol (1840±750 vs 2900±1400, p=0.01) were lower, reflecting less sedation in this group.	**Duration of MV** (4.2[2.1-9.5] vs 8[2.2-22] days, p=0.001) **and time from end of sedation to extubation** (33[12-75] vs 65[36-123] hrs, p=0.01) **were longer in the more deeply sedated patients.**	**ICU** (5[2.5-13] vs 11[2.5-27] days, p=0.004) and **hospital** (17[5-22] vs 21[5-33] days, p=0.003) **LOS was longer in the more deeply sedated patients.**	**No difference in ICU** (n=63[31%] vs 88[39%], p=0.19) **or hospital** (n=75[38%] vs 101[45%], p=0.22) **mortality** between the groups.	N/A	N/A	**Incidence of VAP** (n=12[6%] vs 34[15%], p=0.005) and **extubation failure was higher in the more deeply sedated patients, but there were was no difference in self-extubations** (n=21[10.7%] vs 16[7%], p=0.09) between the groups.
Ren XL, et al. 2017 *China*	Doses of sufentanil (0.018±0.009 vs 0.030±0.007 mg/kg/hr, p<0.0001) and midazolam (0.017±0.009 vs 0.029±0.007 mg/kg/hr, p<0.0001) were lower after implementation of the ABCDE bundle, compared to before, reflecting less sedation in this group.	**Duration of MV was longer** (5.67±3.03 vs 7.51±3.36 days, p=0.001) **in the more deeply sedated patients.**	**ICU stay was longer** (7.47±2.53 vs 9.76±3.75 days, p<0.001) **in the more deeply sedated patients.**	**28 day survival was lower** (72.9% vs 87.7%, p=0.026) **in the more deeply sedated patients.**	**Incidence of delirium was higher** (41.4% vs 17.8%, p=0.002) **in the more deeply sedated patients.**	N/A	N/A
Samuelson K, et al. 2006 *Sweden*	Depth of sedation was reported as proportion of MAAS between 0-2 (deep sedation), 3 (awake, calm and cooperative) and 4-6 (agitated).	**There was a positive correlation between sedation depth and days of MV** (MAAS 0-2: r=0.36, p<0.0001; MAAS 3: r=-0.38, p<0.0001, MAAS 4-6: r=0.29, p<0.0001).	**There was a positive correlation between sedation depth and ICU LOS** (MAAS 0-2: r=0.17, p=0.009; MAAS 3: r=-0.23, p<0.0001, MAAS 4-6: r=0.35, p<0.0001).	N/A	N/A	MAAS score for total ICU stay (proportion) in patients with memories vs no memories: MAAS 0-2: 0.25[0.26] vs 0.50[0.43]; MAAS 3: 0.70[0.32] vs 0.37[0.43]; MAAS 4-6: 0.0[0.10] vs 0.0[0.13]. **Deep sedation was associated with amnesia, paranoid and delusional memories on multivariate analysis.**	N/A

Samuelson K, et al. 2007 *Sweden*	Depth of sedation was reported as proportion of MAAS 3 (awake, calm and cooperative). Patients were categorised as those who experienced each of the outcomes of: (A) bothered vs not bothered by at least one ICU experience, (B) bothered vs not bothered by nightmares and (C) with memories vs no memories of ETT.	**Duration of MV was longer in the more deeply sedated patients in relation to each outcome:** (A: 1.69[4.47] vs 0.71[1.30] days, p=0.002; B: 4.80[7.17] vs 1.04[2.78] days, p<0.0001; C: 1.55[4.30] vs 1.28[4.03] days, p=0.081 (NS))	**ICU LOS was longer in the more deeply sedated patients in relation to each outcome**: (A: 4.09[5.95] vs 2.14[3.04] days, p=0.001; B: 7.08[10.7] vs 3.00[3.74] days, p<0.0001; C: not reported)	N/A	N/A	**Patients with memory of ETT had higher proportion of MAAS 3 than those with no memory**(0.56[0.42] vs 0.18[0.42], p<0.0001)	**No difference in incidence of reintubation** (A: 8.3% vs 5.4%, p=0.742; B: not reported; C: 8.6% vs 6.7%, p=0.609) or tracheostomy (A: 8.3% vs 2.7%, p=0.317; B: not reported; C: 6.0% vs 8.9%, p=0.797) between the groups.
Samuelson K, et al. 2008 *Sweden*	Patients divided into two groups according to sedation depth- light sedation (MAAS 3 – 4: median 3.0[0.0]), heavy sedation: (MAAS 1 – 2: median 1.25[1.0].	**No difference in the duration of MV** between the groups (7.5±4.6 vs 7.7±4.7 hrs, p=0.89).	**No difference in ICU LOS** between the groups (18[7.8], vs 24[53] hrs, p=0.08).	N/A	**No difference in incidence of delirium** between the groups (n=0[0%] vs 0[0%], p=1.00).	**No difference in memories of ICU** (n=15[88%] vs 17[94%], p=0.60), **presence of delusional memories in ICU** (n=1[6%] vs 6[33%], p=0.09), or **memories of pain** (n=4[23%] vs 9[50%], p=0.20) between the groups.	**No difference in number of patients reintubated** between the groups (n=1[6%] vs 2[11%], p=1.00).
Sen S, et al. 2017 *USA*	Compared with the control (pre-protocol group), patients in the protocol group received less total benzodiazepine dose (450±701 vs 74±159 mg, p<0.01), reflecting less sedation in this group.	**No difference in duration of MV** between the groups (8[4-10] vs 5[2-9] days, p=0.12].	**ICU** (7[4-11] vs 4[2-7] days, p=0.02) **and hospital** (13[9-18] vs 9[6-13] days, p=0.01) **LOS were longer in the patients who received more sedation.**	**No difference in ICU mortality** between the groups (n=3[2.2%] vs 1[3.1%], p=0.58).	N/A	N/A	**No difference in number of patients with pneumonia** (n=51[37.8%] vs 8[25.0%], p=0.22) **or requiring brain imaging** between the groups. **Duration of treatment for alcohol withdrawal was longer in the patients who received more sedation.**

Shehabi Y, et al. 2018 *Australia, New Zealand, Malaysia, Singapore*	RASS measurements taken frequently in all patients and 'sedation index' calculated. Multivariate analysis to determine the impact of sedation depth on various outcomes.	**Deep sedation was associated with longer time to extubation** (HR 0.80[0.73-0.87], p<0.001).	N/A	**Deep sedation was associated with higher mortality at 180 days** (HR 1.29[1.15-1.46], p<0.001).	**Deep sedation was associated with increased risk of delirium** (HR 1.25[1.10-1.43], p=0.001).	N/A	N/A
Shehabi Y, et al. 2013 *Australia and New Zealand*	Patients in the early goal directed sedation group received more dexmedetomidine (36.55[16.38-13.23] vs 20.58[20.58-20.58] µg, p<0.0001), but less midazolam (0.06[0.02-1] vs 0.3[0.23-0.76] mg, p=0.036) and propofol (9.89[2.41-22.51] vs 33.55[13.54-77.07] mg, p=0.046) and had more RASS assessments in the light sedation range (66% vs 38%, p=0.01) compared with the control group, reflecting less deeply sedated patients in this group.	**No difference in ventilator free days to day 28** (21.3±9.2 vs 20.1±10.1 days, p=0.72) between the groups.	**No difference in ICU** (5.5[4.1-10.4] vs 7.0[2.5-9.4] days, p=0.44) or **hospital** (16.1[9.3-33.3] vs 17[4.0-29.0] days, p=0.49) **LOS** between the groups.	**No difference in either hospital** (n=3[14.3%] vs 2[12.5%], p=1.0) or **90 day** (n=5[23.8%] vs 2[12.5%], p=0.38) **mortality** between the groups.	**No difference in incidence of delirium** (n=8[38%] vs 6[38%], p=0.97) between the groups.	N/A	**No difference in number of patients mobilised, requiring NMB, physically restrained or extubated within 7 days** between the groups.
Shehabi Y, et al. 2012 *Australia and New Zealand*	RASS measurements taken frequently in all patients and data divided into light (RASS -2 to +1) or deep sedation (RASS -3 to -5), or agitation (RASS +2 to +4); (n=1642[61.9%], 942[35.5%], 72[2.7%], assessments respectively). Cumulative dose of midazolam and fentanyl was also analysed.	**Time to extubation was longer** (7.7[6.0-8.6] vs 2.4[1.9-4.0] days) **in patients deeply sedated early in ICU.** In **multivariable analysis deep sedation was independently associated with time to extubation** (HR 0.90, 95% CI 0.87-0.94, p<0.001).	N/A	In **multivariable analysis deep sedation was independently associated with 180 day mortality** (HR 1.08[1.01-1.16], p=0.027).	**Risk of delirium** (RR 1.7[CI 1.00-3.02], p=0.046) **was higher in the more deeply sedated patients,** but time to delirium was not associated with early deep sedation.	N/A	N/A
Strøm T, Stylsvig M, Toft P. 2011	Compared with the DIS group, patients in the no sedation group received less propofol	**No difference in ventilator free days to day 28**	N/A	N/A	N/A	**No difference in psychological problems post-**	N/A

Denmark	(0[0-1.26] vs 1.40[0.52-2.04] mg/kg/hr, p=0.013) and midazolam (0[0-0] vs 0.01[0-0.04] mg/kg/hr, p=0.003), reflecting less sedation in this group.	(23.2[19.0-25.4] vs 16.1[3.9-22.7] days, p=0.12) between the groups.				**discharge** (n=2[15%] vs 6[46%], p=0.20) **or PTSS-10 score >35** (n=1[8%] vs 0[0%], p=0.14) between the groups.	
Strøm T, Martinussen T, Toft P. 2010 *Denmark*	Compared with the DIS group, patients in the no sedation group received less propofol (0[0-0.52] vs 0.77[0.15-1.65] mg/kg/hr, p=0.0001) and midazolam (0[0-0] vs 0.003[0-0.024] mg/kg/hr, p<0.0001), reflecting less sedation in this group.	**Ventilator free days to day 28 was lower** (6.9[0-20.5] vs 18.0[0-24.1] days, p=0.019) **in the patients who received more sedation.**	**ICU** (22.8[11.7-NR] vs 13.1[5.7-NR] days, p=0.032) and **hospital** (58[33-85] vs 34[17-65] days, p=0.004) **LOS were longer in the patients who received more sedation.**	**No difference in ICU** (n=22[38%] vs 12[22%], p=0.06) or **hospital** (n=27[47%] vs 20[36%], p=0.27) **mortality** between the groups.	N/A	N/A	**No difference in incidence of VAP** (n=7[12%] vs 6[11%], p=0.85), **or patients requiring tracheostomy** (n=17[29%] vs 16[29%], p=0.98) between the groups.
Treggiari MM, et al. 2009 *Switzerland*	Patients divided into two groups according to sedation depth- light sedation: Ramsay sedation score 1 – 2, heavy sedation: Ramsay sedation score 3 – 4. Doses of midazolam and opioids were higher in the heavy sedation group.	**Days of MV were higher** (5.5±10.8 vs 2.9±5.0 days, p=0.02), **and ventilator free days to day 28 lower** (26.6±NR vs 27.6±NR, p=0.03) **in the more deeply sedated patients.**	**ICU LOS** (5.5[2-99] vs 4.0[1-129]* days, p=0.03) **was longer in the more deeply sedated patients, but there was no difference in hospital LOS** (20[13-38] vs 16[12.5-32.5] days, p=0.47) between the groups. *(median[range])	**No difference in either ICU** (n=9[14%] vs 9[14%], p>0.99) or **hospital** (n=11[17%] vs 12[18%], p=0.65) **mortality** between the groups.	N/A	**No difference in PTSD questionnaire score** (discharge: 57±30 vs 52±33, p=0.39; 4 wk follow-up: 56±29 vs 46±29, p=0.07), **PTSD symptom clusters, anxiety or depression** (discharge: 6.5±4.7 vs 5.3±3.4, p=0.13; 4 wk follow-up: 3.1±3.7 vs 3.4±3.7, p=0.72) scores at either discharge or 4 week follow-up between the groups.	**No difference in rate of agitation, use of restraints, self-extubation** (n=2[5%] vs 2[3%], p=0.68), **extubation failure, tracheotomy** (n=4[6%] vs 3[5%], p=0.73) **or the incidence of organ dysfunction** between the groups.

Abbreviations: ARDS: Acute Respiratory Distress Syndrome, CAM-ICU: Confusion Assessment Method for the ICU, D%IS: Daily Interruption of Sedation, ETT: Endotracheal Tube, ICU: Intensive Care Unit, LOS: Length of Stay, MAAS: Motor Activity Assessment Scale, MV: Mechanical Ventilation, NMB: Neuromuscular Blockade, PTSD: Post-Traumatic Stress Disorder, RASS: Richmond Agitation Sedation Scale, SAS: Riker Sedation Agitation Scale, SAT: Spontaneous Awakening Trial, SBT: Spontaneous Breathing Trial, VAP: Ventilator-Associated Pneumonia.

Table S6: Summary of findings – sensitivity analysis (using studies where difference in depth of sedation was demonstrated in RASS or SAS)

Outcomes	Study Type	Number of studies (participants)	Effect on Outcome & 95% CI	I^2
Primary Outcomes				
Mortality				
ICU mortality (%)	RCT	2 (189)	0.74 (0.42, 1.32)	0%
	Cohort	1 (1884)	0.18 (0.14, 0.24)	-
Physiological outcomes				
Duration of mechanical ventilation (days)	RCT	2 (165)	-1.44 (-3.79, 0.91)	20%
	Cohort	5 (2651)	-0.58 (-1.76, 0.59)	79%
Secondary outcomes				
Mortality				
Hospital mortality (%)	RCT	3 (226)	0.87 (0.54, 1.40)	0%
	Cohort	4 (4213)	0.70 (0.33, 1.49)	96%
Physiological outcomes				
Time to extubation (days)	RCT	Nil		
	Cohort	2 (2132)	-3.77 (-5.49, -2.06)	96%
Ventilator free days to day 28 (days)	RCT	3 (797)	4.13 (-2.18, 10.43)	41%
	Cohort	2 (431)	0.65 (-0.65, 1.95)	0%
Delirium (%)	RCT	3 (133)	1.19 (0.70, 2.04)	0%
	Cohort	3 (3810)	1.25 (0.75, 2.07)	96%
Resource Use				
ICU length of stay (days)	RCT	4 (926)	0.96 (-1.71, 3.63)	0%
	Cohort	5 (3833)	-4.88 (-10.59, 0.83)	98%
Hospital length of stay (days)	RCT	3 (226)	0.98 (-6.59, 8.56)	68%
	Cohort	4 (4356)	-4.01 (-8.91, 0.89)	93%
Tracheostomy (%)	RCT	2 (189)	0.79 (0.22, 2.85)	0%
	Cohort	2 (431)	0.59 (0.31, 1.12)	0%
Adverse Events				
Self-extubation (%)	RCT	2 (189)	1.31 (0.30, 5.82)	0%
	Cohort	2 (431)	1.14 (0.58, 2.22)	0%
Re-intubation (%)	RCT	4 (925)	1.17 (0.38, 3.57)	43%
	Cohort	2 (362)	1.07 (0.43, 2.65)	0%
VAP (%)	RCT	Nil		
	Cohort	1 (1483)	0.71 (0.46, 1.08)	-

Abbreviations: ICU: Intensive Care Unit, RCT: Randomized Control Trial, VAP: Ventilator-Associated Pneumonia; RASS: Richmond Agitation-Sedation Scale; SAS: Riker Sedation-Agitation Scale

Supplementary Table 7: Summary of findings – sensitivity analysis (examining influence of retrospective and prospective design)

Outcomes	Study Type	Number of studies (participants)	Effect on Outcome & 95% CI	I^2
ICU mortality (%)	Prospective	2 (590)	0.83 [0.64 to 1.07]	0%
Duration of MV (days)	Prospective	7 (3160)	-1.49 [-2.81 to -0.17]	87%
Hospital mortality (%)	Prospective	3 (2608)	0.93 [0.80 to 1.09]	0%
	Retrospective	2 (2028)	0.49 [0.16 to 1.54]	95%
Delirium (%)	Prospective	3 (2069)	1.09 [0.58 to 2.05]	94%
ICU LOS (days)	Prospective	6 (2538)	-2.30 [-4.15 to -0.45]	90%
	Retrospective	2 (2028)	-8.07 [-19.72 to 3.58]	99%
Hospital LOS (days)	Prospective	5 (3062)	-2.78 [-4.37 to -1.20]	43%

MV: mechanical ventilation; LOS: length of stay

Supplementary figure legends

Supplementary figure 1: Risk of bias summary – RCTs

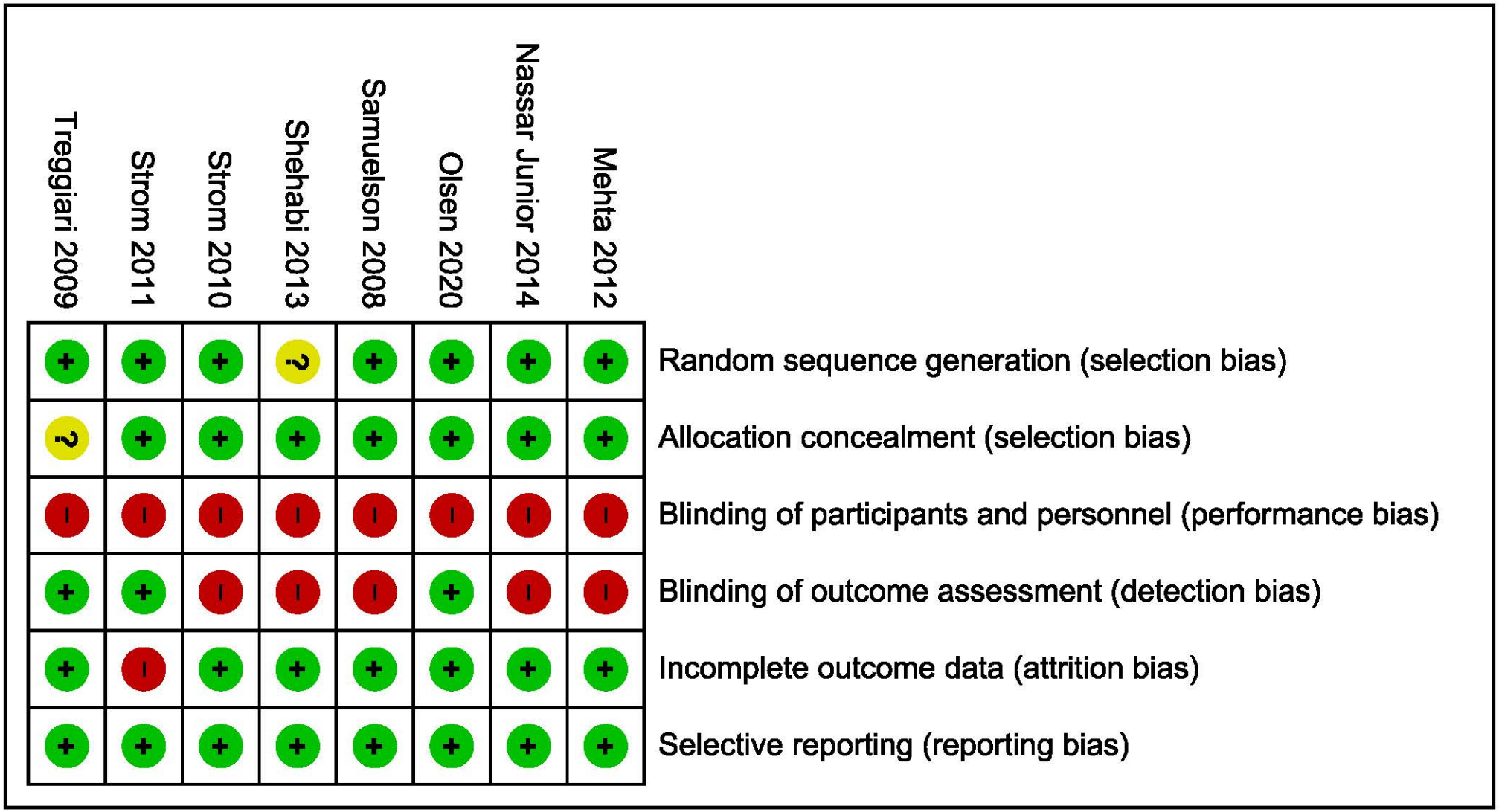
Supplementary figure 2: Risk of bias summary – cohort studies

Supplementary Figure 3: Forest plots for secondary outcomes – mortality domain: hospital mortality

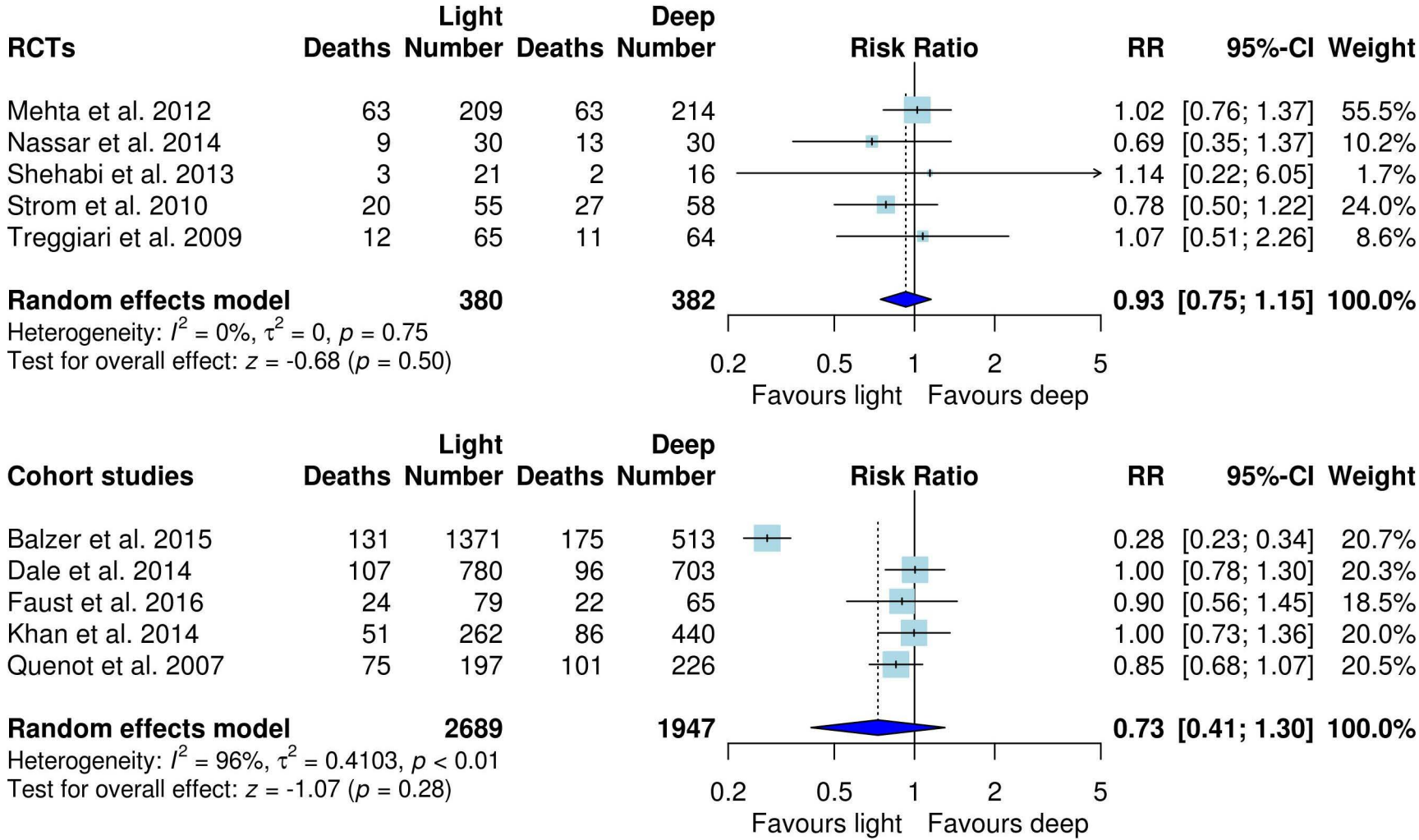
Supplementary Figure 4: Forest plots for secondary outcomes – physiological domain: a) time to extubation; b) ventilator free days to day 28; c) delirium

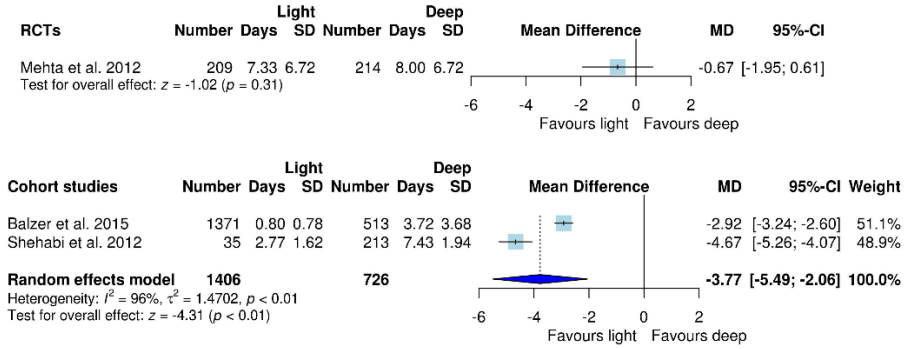
Supplementary Figure 5: Forest plots – secondary outcomes – resource use domain: a) ICU LOS; b) hospital LOS; c) tracheostomies

Supplementary Figure 6: Forest plots – secondary outcomes – adverse events domain: a) self-extubation; b) re-intubation; c) Ventilator associated pneumonia

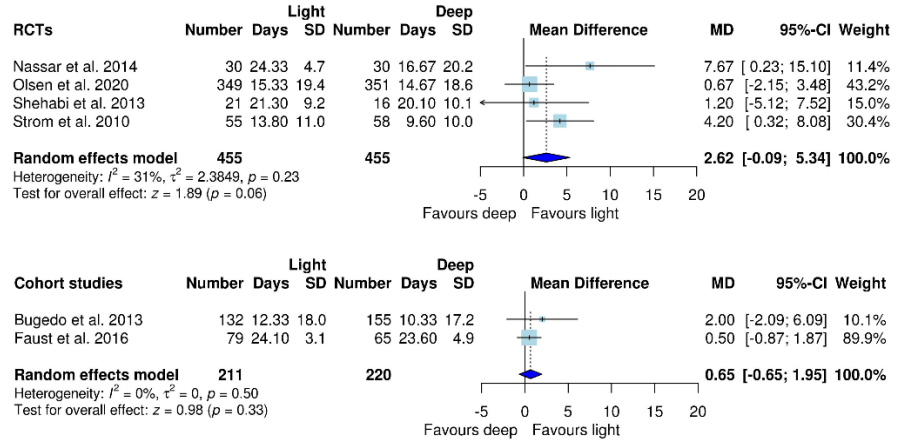

Treggiari 2009
Strom 2011
Strom 2010
Shehabi 2013
Samuelson 2008
Olsen 2020
Nassar Junior 2014
Mehta 2012
Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Incomplete outcome data (attrition bias)
Selective reporting (reporting bias)

Method of recruitment (selection bias)
Measurement of depth of sedation bias
Outcome assessment (detection bias)
Important confounding factors (bias)
Length & Completeness of follow-up (attrition bias)
Other bias
Arabi 2007
Balzer 2015
Bugedo 2013
Burry 2015
Capuzzo 2001
Costa 2014
Dale 2014
Faust 2016
Guttormson 2011
Khan 2014
Mendes 2008
Quenot 2007
Ren 2017
Samuelson 2006
Samuelson 2007
Sen 2017
Shehabi 2012
Shehabi 2018

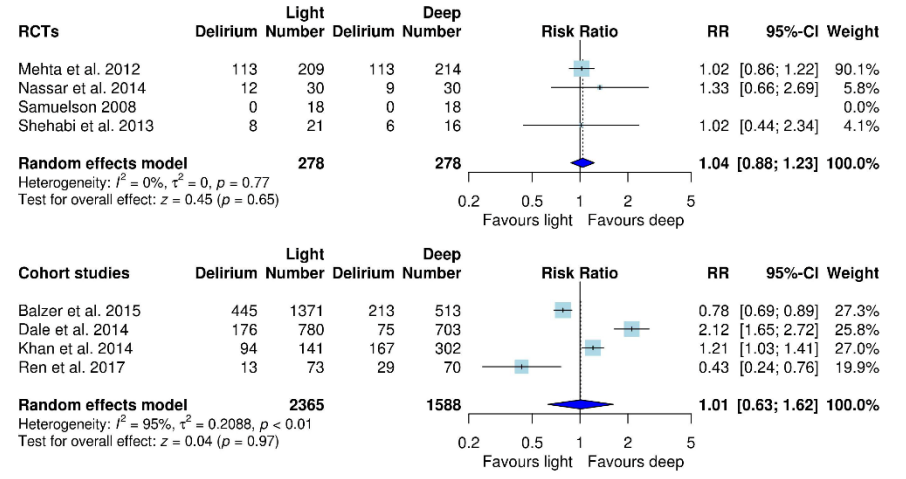

RCTs
Light Deaths
Light Number
Deep Deaths
Deep Number
Risk Ratio
RR
95%-CI
Weight
Mehta et al. 2012 63 209 63 214 1.02 [0.76; 1.37] 55.5%
Nassar et al. 2014 9 30 13 30 0.69 [0.35; 1.37] 10.2%
Shehabi et al. 2013 3 21 2 16 1.14 [0.22; 6.05] 1.7%
Strom et al. 2010 20 55 27 58 0.78 [0.50; 1.22] 24.0%
Treggiari et al. 2009 12 65 11 64 1.07 [0.51; 2.26] 8.6%
Random effects model 380 382 0.93 [0.75; 1.15] 100.0%
Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.75
Test for overall effect: z = -0.68 (p = 0.50)
0.2
0.5
1
2
5
Favours light
Favours deep
Cohort studies
Light Deaths
Light Number
Deep Deaths
Deep Number
Risk Ratio
RR
95%-CI
Weight
Balzer et al. 2015 131 1371 175 513 0.28 [0.23; 0.34] 20.7%
Dale et al. 2014 107 780 96 703 1.00 [0.78; 1.30] 20.3%
Faust et al. 2016 24 79 22 65 0.90 [0.56; 1.45] 18.5%
Khan et al. 2014 51 262 86 440 1.00 [0.73; 1.36] 20.0%
Quenot et al. 2007 75 197 101 226 0.85 [0.68; 1.07] 20.5%
Random effects model 2689 1947 0.73 [0.41; 1.30] 100.0%
Heterogeneity: I^2 = 96%, τ^2 = 0.4103, p < 0.01
Test for overall effect: z = -1.07 (p = 0.28)
0.2
0.5
1
2
5
Favours light
Favours deep


RCTs
Light
Number Days SD
Deep
Number Days SD
Mean Difference
MD 95%-CI
Mehta et al. 2012 209 7.33 6.72 214 8.00 6.72 -0.67 [-1.95; 0.61]
Test for overall effect: z = -1.02 (p = 0.31)
-6 -4 -2 0 2
Favours light Favours deep
Cohort studies
Number Days SD Number Days SD
Mean Difference
MD 95%-CI Weight
Balzer et al. 2015 1371 0.80 0.78 513 3.72 3.68 -2.92 [-3.24; -2.60] 51.1%
Shehabi et al. 2012 35 2.77 1.62 213 7.43 1.94 -4.67 [-5.26; -4.07] 48.9%
Random effects model 1406 726 -3.77 [-5.49; -2.06] 100.0%
Heterogeneity: I² = 96%, τ² = 1.4702, p < 0.01
Test for overall effect: z = -4.31 (p < 0.01)
-6 -4 -2 0 2
Favours light Favours deep

a)

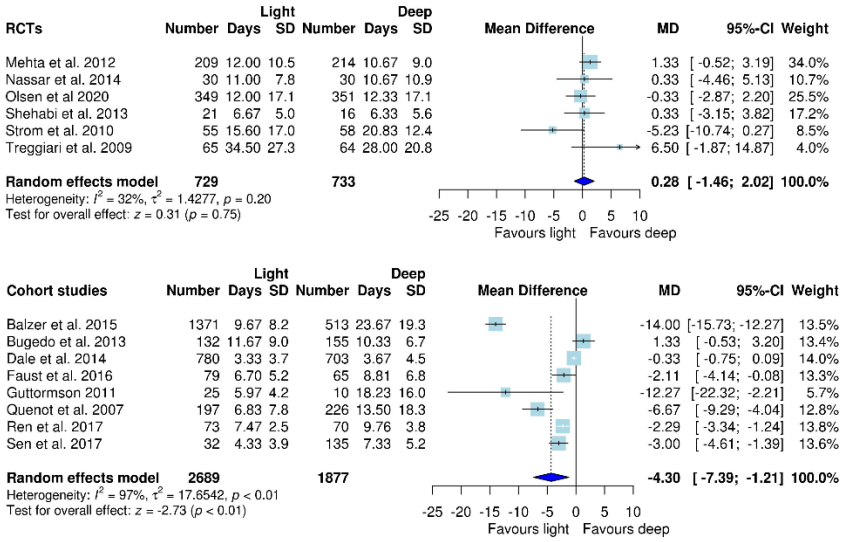

RCTs
Light
Number Days SD
Deep
Number Days SD
Mean Difference
MD 95%-CI Weight
Nassar et al. 2014 30 24.33 4.7 30 16.67 20.2 7.67 [0.23; 15.10] 11.4%
Olsen et al. 2020 349 15.33 19.4 351 14.67 18.6 0.67 [-2.15; 3.48] 43.2%
Shehabi et al. 2013 21 21.30 9.2 16 20.10 10.1 1.20 [-5.12; 7.52] 15.0%
Strom et al. 2010 55 13.80 11.0 58 9.60 10.0 4.20 [0.32; 8.08] 30.4%
Random effects model 455 455 2.62 [-0.09; 5.34] 100.0%
Heterogeneity: I² = 31%, τ² = 2.3849, p = 0.23
Test for overall effect: z = 1.89 (p = 0.06)
-5 0 5 10 15 20
Favours deep Favours light
Cohort studies
Number Days SD Number Days SD
Mean Difference
MD 95%-CI Weight
Bugedo et al. 2013 132 12.33 18.0 155 10.33 17.2 2.00 [-2.09; 6.09] 10.1%
Faust et al. 2016 79 24.10 3.1 65 23.60 4.9 0.50 [-0.87; 1.87] 89.9%
Random effects model 211 220 0.65 [-0.65; 1.95] 100.0%
Heterogeneity: I² = 0%, τ² = 0, p = 0.50
Test for overall effect: z = 0.98 (p = 0.33)
-5 0 5 10 15 20
Favours deep Favours light

b)

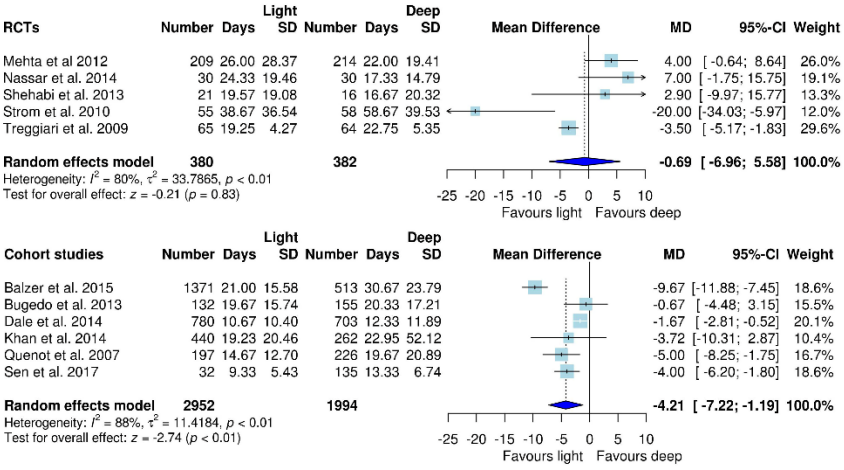

RCTs
Light
Delirium Number
Deep
Delirium Number
Risk Ratio
RR 95%-CI Weight
Mehta et al. 2012 113 209 113 214 1.02 [0.86; 1.22] 90.1%
Nassar et al. 2014 12 30 9 30 1.33 [0.66; 2.69] 5.8%
Samuelson 2008 0 18 0 18 0.0%
Shehabi et al. 2013 8 21 6 16 1.02 [0.44; 2.34] 4.1%
Random effects model 278 278 1.04 [0.88; 1.23] 100.0%
Heterogeneity: I² = 0%, τ² = 0, p = 0.77
Test for overall effect: z = 0.45 (p = 0.65)
0.2 0.5 1 2 5
Favours light Favours deep
Cohort studies
Delirium Number Delirium Number
Risk Ratio
RR 95%-CI Weight
Balzer et al. 2015 445 1371 213 513 0.78 [0.69; 0.89] 27.3%
Dale et al. 2014 176 780 75 703 2.12 [1.65; 2.72] 25.8%
Khan et al. 2014 94 141 167 302 1.21 [1.03; 1.41] 27.0%
Ren et al. 2017 13 73 29 70 0.43 [0.24; 0.76] 19.9%
Random effects model 2365 1588 1.01 [0.63; 1.62] 100.0%
Heterogeneity: I² = 95%, τ² = 0.2088, p < 0.01
Test for overall effect: z = 0.04 (p = 0.97)
0.2 0.5 1 2 5
Favours light Favours deep

c)

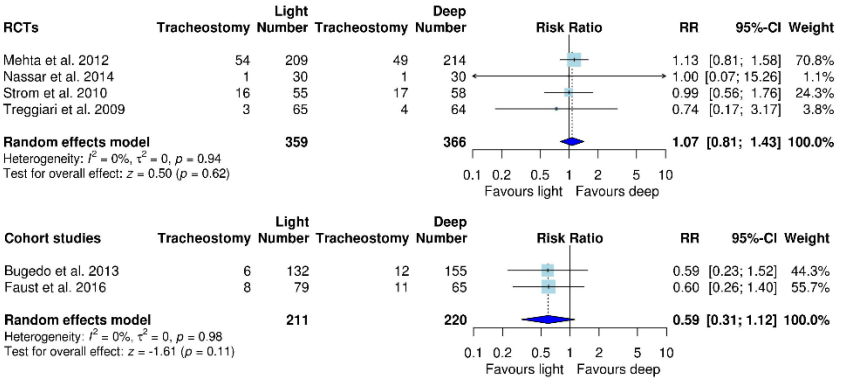
Supplementary Figure 4: Forest plots for secondary outcomes – physiological domain: a) time to extubation; b) ventilator free days to day 28 (VFD28); c) delirium
Note: data converted from median/IRQ to mean/SD[12] – time to extubation in the following studies: Balzer et al 2015; Mehta et al 2012; Shehabi et al 2012; VFD28 in the following studies: Bugedo et al 2013; Nassar Jr et al 2014.



a)



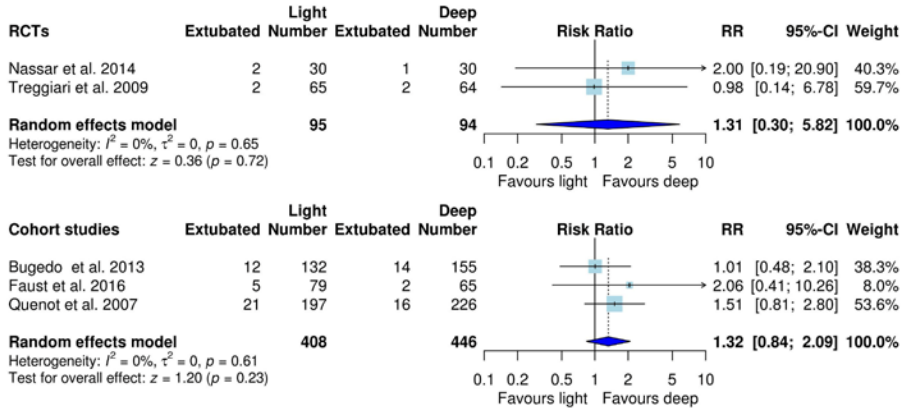
b)



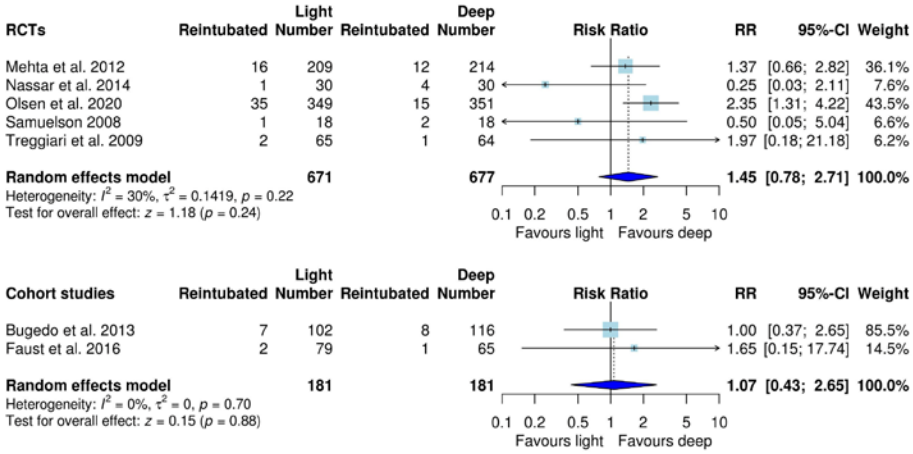
c)

Supplementary Figure 5: Forest plots – secondary outcomes – resource use domain: a) ICU LOS; b) hospital LOS; c) tracheostomies

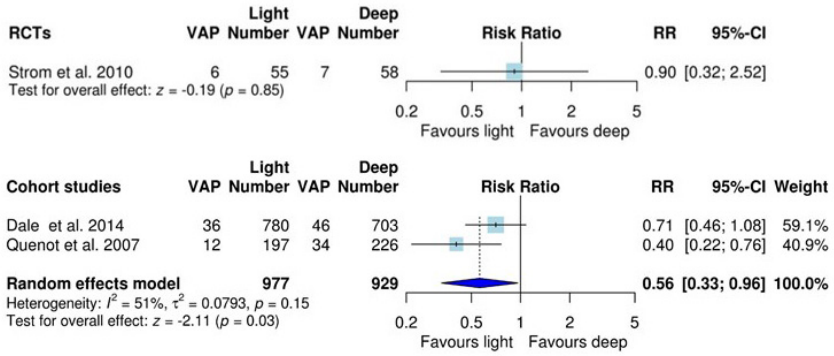
Note: data converted from median/IRQ to mean/SD[12] – ICU LOS& hospital LOS in the following studies: Balzer et al 2015; Bugedo et al 2013; Dale et al 2014; Guttormson et al 2011 (ICU LOS only); Mehta et al 2012; Nassar Jr et al 2014; Quenot et al 2007; Sen et al 2017; Shehabi et al 2013; Strom et al 2010; Treggiari et al 2009.



a)



b)



c)

Supplementary Figure 6: Forest plots – secondary outcomes – adverse events domain: a) self-extubation; b) re-intubation; c) Ventilator associated pneumonia