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