Intensive Care Med (2000) 26: 275–285

REVIEW

B. De Jonghe
D. Cook
C. Appere-De-Vecchi
G. Guyatt
M. Meade
H. Outin

Using and understanding sedation scoring systems: a systematic review

Accepted: 20 December 1999

B. De Jonghe (✉) · C. Appere-De-Vecchi · H. Outin
Service de Reanimation Medicale, Centre Hospitalier de Poissy-Saint-Germain, 10 rue du Champ-Gaillard, 78300 Poissy, France
e-mail: bdj@club-internet.fr
Tel.: + 33-1 39-27 52 02
Fax: + 33-1 39-27 44 46

D. Cook
St. Joseph's Hospital, Hamilton, Ontario, Canada

G. Guyatt
McMaster University, Hamilton, Ontario, Canada

Maureen Meade
Hamilton General Hospital, Hamilton, Ontario, Canada

Abstract *Objective*: To systematically review instruments for measuring the level and effectiveness of sedation in adult and pediatric ICU patients.
Study identification: We searched MEDLINE, EMBASE, the Cochrane Library and reference lists of the relevant articles. We selected studies if the sedation instrument reported items related to consciousness and one or more additional items related to the effectiveness or side effects of sedation.
Data abstraction: We extracted data on the description of the instrument and on their measurement properties (internal consistency, reliability, validity and responsiveness).
Results: We identified 25 studies describing relevant sedation instruments. In addition to the level of consciousness, agitation and synchrony with the ventilator were the most frequently assessed aspects of sedation. Among the 25 instruments, one developed in pediatric ICU patients (the Comfort Scale), and 3 developed in adult ICU patients (the Ramsay scale, the Sedation-Agitation-Scale and the Motor Activity Assessment Scale), were tested for both reliability and validity. None of these instruments were tested for their ability to detect change in sedation status over time (responsiveness).
Conclusion: Many instruments have been used to measure sedation effectiveness in ICU patients. However, few of them exhibit satisfactory clinimetric properties. To help clinicians assess sedation at the bedside, to aid readers critically appraise the growing number of sedation studies in the ICU literature, and to inform the design of future investigations, additional information about the measurement properties of sedation effectiveness instruments is needed.

Key words Critical illness · Sedatives · Measurement · Reliability · Validity · Responsiveness

Introduction

Most critically ill patients require sedation, analgesia or both for at least part of their stay in the intensive care unit (ICU). Sedation can minimize agitation, promote synchrony with the ventilator, and help to relieve the anxiety and discomfort associated with the high technology environment of the ICU. The presence of an endotracheal tube, the performing of various diagnostic tests, and interventions such as tracheal suctioning, mobilization, and transportation may necessitate either intermittent or continuous administration of sedative drugs. Sedation is integral to the management of certain ICU patients, including those with severe intracranial hypertension or severe respiratory failure.

In 1981, a survey of 34 ICUs in Great Britain and Ireland found considerable variability in sedative agents used [1]; however, the common target level of sedation

was reflected in the fact that 67% of respondents believed that patients should ideally be "detached from the ICU environment". A follow-up study in 1987 found that 69% of respondents would prefer to have their patients "asleep but easily rousable" [2]. In a survey conducted in the United States in 1991, 84% of ICUs reported frequent use (20–70% of patients) or routine use (> 70% of patients) of sedatives for mechanically ventilated patients [3]. Written standard protocols for sedation were used in 33% of the 49 respondent ICUs in a very recent survey of sedation practice in Denmark [4].

Inappropriate administration of sedation has potentially serious consequences. Insufficient sedation may lead to life-threatening agitation precipitating myocardial ischemia or ventilator dysynchrony. Excessive sedation may create prolonged alteration of consciousness, which could lead to an increased duration of mechanical ventilation [5]. This in turn may predispose to an increased risk of ventilator associated pneumonia [6], ventilator associated lung injury [7], and critical illness neuromuscular abnormalities[8].

Methods used to achieve and evaluate tolerance to the ICU environment are often determined by tradition or by convenience. Just as administration of vasoactive agents is titrated to patient-specific pathophysiology, administration of sedative drugs should be titrated to patient-specific objectives. Accordingly, intensivists require tools that measure the effectiveness of sedation in individual patients in relation to the objectives of sedation. Such instruments should ideally be simple and user-friendly at the bedside, yet should have also undergone rigorous development and appropriate testing to demonstrate validity, reliability and responsiveness.

Since the first description of the widely used Ramsay scale in 1974 [9], several different sedation scoring instruments have been used in clinical investigations in the ICU as well as in daily practice. These instruments usually include descriptions of level of consciousness, and often descriptions of agitation, pain, or synchrony with the ventilator. The objectives of this systematic review are to summarize the available sedation scales, to highlight the domains that they explore, to present their clinimetric properties, and to consider the implications of these findings for clinical practice and research.

Methods

Study identification

To identify studies, we searched MEDLINE and EMBASE from January 1966 to July 1999, using the following text words and key words: *sedation* or *stress* or *anxiety* or *agitation*, *ICU* or *intensive care* or *critically ill* or *artificial ventilation*, *score* or *scale* or *assessment*. In the Cochrane Library, we searched both the Clinical Trials Registry for randomized trials and the Cochrane database of Systematic Reviews. We had no language restrictions. The titles (and the abstracts, when available), in the MEDLINE, EMBASE and Cochrane databases and the reference lists of all primary articles and review articles were reviewed independently in duplicate. We also searched our personal files for relevant studies. Any additional relevant articles were identified and retrieved.

Study selection

Two reviewers independently applied the following selection criteria:

1. Population: Adult or pediatric ICU patients
2. Sedation scores: Studies had to report the original description of a bedside clinical instrument measuring the effect of sedation as used by ICU health care workers. Studies were selected if the instrument reported: (a) items related to consciousness; and (b) one or more items related to domains targeted by sedation (e.g., relief of agitation, ventilator dysynchrony) or to sedation side effects (e.g., hypotension, tachycardia).
3. Design: Eligible studies were published as full-text articles or in abstract form. Eligible studies provided either the original description of the instrument with the corresponding context of its utilization, or data about internal consistency, validity, reliability, and/or responsiveness of a previously published instrument.

We excluded studies reporting instruments describing only consciousness, and instruments that did not include a categorical scale. We also excluded studies that evaluated two different domains at different times in the ICU, and studies that used a sedation instrument for which a comprehensive description of sedation was not provided (e.g., excellent, good, adequate, poor sedation). We omitted studies reporting an original instrument with no information on its clinical use in the study. We did not consider instruments for patient self-evaluation and paraclinical tests to measure the level of sedation (e.g., EEG bispectral analysis [10], auditory evoked potentials [11], RR variability [12]), used either as an original tool or as an instrument to test validity of a clinical score. Finally, we excluded studies conducted primarily in the recovery room or day surgery units.

Data abstraction

Two reviewers abstracted data concerning the population, the study design, the *domains* of sedation effectiveness and/or side effects, and the *internal consistency*, *reliability, validity* and *responsiveness* of the instruments. Disagreements between reviewers were resolved by discussion and consensus.

These instruments measure the degree of sedation (which authors sometime refer to as "sedation scores" or "sedation scales") and are typically made up of one or more items. Each item has a number of response options, which can be measured as categorical variables, either numerical (e.g., with a 5- or 7-point scale) or non-numerical. An instrument may have *domains* or *dimensions* (which investigators sometimes call "*subscales*" or "*subscores*"). Typically, item scores are added up to form domain scores, which are then summed to create *total scores* for the entire instrument.

For the purpose of this review, we defined *internal consistency* as the extent to which items correlated with one another. Ideally, internal consistency is measured using Cronbach's alpha. We defined *reliability* as the consistency with which a measure discriminates between patients at a single point in time. *Inter-rater reliability*, a frequently assessed clinimetric property, is reflected by the

extent to which the duplicate observations give identical results. We considered either a weighted kappa or an intraclass correlation relating the between-person variance to the total variance as the most appropriate statistics for measurement of reliability. Pearson's correlation measures association rather than agreement, i.e., systematic differences between two observations (e.g., scores systematically and consistently higher or lower than each other) will not attenuate the correlation.

We defined *validity* as the extent to which an instrument is truly measuring the degree of sedation. *Criterion validity* refers to the relationship between an instrument and a gold or reference standard. There is no such standard available for sedation, and therefore criterion validity is not relevant. A construct is a theoretically derived notion of the domains we wish to measure. An understanding of the construct will lead to expectations about how an instrument should behave if it is valid. *Construct validity,* therefore, involves comparisons between measures, and examination of the logical relationship that should exist between a measure and characteristics of patients and patient groups. In this case, correlations between the instrument score and other measures of the degree of sedation that were relatively high and similar to a priori predictions provide strong evidence of construct validity.

We defined *responsiveness* as the extent to which an instrument can detect important changes in sedation, even if those changes are small.

Results

Populations, study designs and instrument descriptions of the 25 eligible instruments are presented in Table 1 [9, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36]. The results of one study were presented in two separate publications [22, 37] of which only one was selected [22]. One study was published in Spanish [38], one in German [15] and the remainder in English. Three studies were specifically designed to develop and validate a new sedation scale [18, 21, 36]. Two other studies also contained a pilot study designed to validate the scale [25, 31]. Two studies exclusively reported the development of a new scale [17, 23]. All the other studies were randomized controlled trials (RCT) or case series of ICU patients, using a sedation score to assess the effectiveness and/or side effects of a sedation regimen [9, 13, 14, 15, 16, 19, 22, 24, 26, 27, 28, 29, 30, 32, 33], or studies of a sedation algorithm assessment [35]. Twenty studies concerned adult ICU patients [9, 13, 14, 15, 16, 17, 20, 22, 23, 24, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36] and 5 enrolled pediatric ICU patients [18, 19, 21, 25, 31].

In addition to consciousness which was evaluated in all instruments, scales mostly included items evaluating agitation [9, 13, 14, 17, 19, 20, 21, 22, 23, 24, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36] and/or ventilator synchrony [13, 14, 15, 16, 19, 20, 21, 24, 25, 28, 29, 30]. Instruments described other aspects such as pain [13, 19, 21], anxiety [9, 24, 30], reaction to endotracheal aspiration [17, 19, 20, 25, 28, 30, 31], or muscle tone [21], less frequently. Two studies used precise hemodynamic variables as part of the sedation measurement [21, 35]. Most instruments were composed of single categorical items measured as numerical [9, 13, 14, 17, 22, 23, 24, 26, 27, 30, 31, 33, 34, 35, 36]. Eight scales incorporated several items which result in a wide range of possible scores [15, 16, 18, 19, 20, 21, 25, 28, 29, 34, 35, 36].

Nine studies provided evidence on the methodological properties of nine sedation instruments (Table 2) [18, 21, 25, 31, 36, 38, 39, 40, 41]. Five studies published in full text form described the original instrument along with results of their evaluation: the Vancouver Sedative Recovery Scale [18], the Comfort Scale [21], two other pediatric sedation scales [25, 31], and the Motor Activity Assessment Scale [36]. Clinimetric properties were evaluated in reports published after the original description of the other scores [9, 16, 23, 27].

Internal consistency was documented to be high for the two scales for which it was measured: the Vancouver Sedative Recovery Scale (Cronbach's alpha = 0.87) [18], and the Comfort Scale (Cronbach's alpha = 0. 90) [21].

The reliability of the Ramsay scale [38, 40], the GCS modified by Cook and Palma [38], the Vancouver Sedative Recovery Scale [18], the Comfort Scale [21], the New Sheffield Sedation Scale [41], the Sedation-Agitation Scale [40], the pediatric sedation scale described by Hughes [25] and by Parkinson [31], and the Motor Activity Assessment Scale [36] was high, reflected in overall correlation coefficients of 0.84 to 0.98, and kappa values of 0.73 to 0.94. In the eight reliability studies assessing these nine scoring systems, all evaluations were paired, excepted for the Comfort Scale, which was assessed by three raters simultaneously (one ICU nurse, one research assistant and the study investigator) [21]. Paired evaluations always involved one ICU nurse familiar with the score. The second rater was either another ICU nurse [36, 40, 41], a research nurse [25, 31] or, alternatively, another ICU nurse and a research assistant [18], or another ICU nurse and a doctor [38].

The validity of these sedation instruments was evaluated by comparing the index instrument with another instrument also designed to measure sedation effectiveness. However, the «reference» instruments used in these validity assessments have not themselves been independently tested for validity. We found that the Ramsay scale has been validated against the Glasgow Coma Scale modified by Cook and Palma and vice versa [39]. The Ramsay Scale has also been validated against the Sedation-Agitation Scale and vice versa [40]. The Comfort Scale has been validated against a visual analogue scale [21]. The Motor Activity Assessment Scale has been validated against a visual analogue scale, against blood pressure and heart rate variations, and against the occurrence of "agitation-related events" [36]. In all analyses, the correlations were high, and many reached

Table 1 Instruments used for assessing sedation in critically ill patients

	Author	Name of Instrument	Conditions measured						Instrument structure			Study design	Population
			Consciousness	Agitation	Ventilator synchrony	Pain	Other	Physiologic variables	# Subscales	Total # items	Range of response options		
1	Ramsay 1974 [9]	Ramsay Scale	√	√	–	–	Anxiety	–	–	1 numerical	1 to 6	Case series of alphaxalonealphadolone	30 adult pts (26 ventilated)
2	Yate 1986 [13]	–	√	√	√	√	–	–	–	1 numerical	1 to 4	RCT of alfentanil vs pethidine	30 ventilated post cardiac surgery adult pts
3	Cohen 1987 [14]	–	√	√	√	–	–	–	–	1 numerical	0 to 4	Case series of alfentanil	16 ventilated adult pts
4	Adams 1988 [15]	–	√		√	–	–	–	–	2 numerical	2 to 8	RCT of ketamine vs fentanyl	16 ventilated adults pts
5	Cook 1989 [16]	GCS modified by Cook and Palma (1)	√	–	√	–	–	–	–	4 numerical	4 to 18	Case series of propofol	10 ventilated adult pts
6	O'Sullivan 1990 [17]	Cambridge Sedation Scale	√	√	–	–	Reaction to tracheal suction	–	–	1 numerical	0 to 6	Scale development study	Adult pts
7	Macnab 1991 [18]	Vancouver Sedative Recovery Scale (VSRS)	√	–	–	–	–	–	3 (response, eyes & movement)	12 (either numerical or yes/no)	0 to 22	Scale development and validation study	91 post-operative or trauma paediatric pts
8	Hartwig 1991 [19]	–	√	√	√	√	Reaction to tracheal suction	–	5 (motor, mimic, eyes, respiration & reaction to tracheal suction)	5 numerical	8 to 25	Case series of midazolam	24 ventilated paediatric pts
9	Harris 1991 [20]	Harris scale	√	√	√	–	Reaction to tracheal suction	–	–	3 numerical	3 to 14	Case series of the effect of haemofiltration on sedation level and blood propofol concentration	10 adult pts
10	Ambuel 1992 [21]	Comfort Scale	√	√	√	√ Facial expression	Muscle tone	MABP & HR	–	8 numerical	8 to 40	Scale development and validation study	37 ventilated paediatric pts
11	Chaudhri 1992 [22]	–	√	√	–	–	–	–	–	1 numerical	1 to 6	RCT of propofol vs midazolam	40 ventilated post cardiac bypass surgery adult pts
12	Laing 1992 [23]	New Sheffield Sedation Scale	√	√	–	–	–	–	–	1 numerical	1 to 6	Scale development study	Adult pts (number not reported)
13	Spencer 1994 [24]	–	√	√	√	–	Anxiety	–	–	1 numerical	1 to 6	Correlation study of EEG spectral analysis with a clinical sedation score	23 ventilated adult pts

Table 1 Continued

Author		Name of Instrument	Conditions measured						Instrument structure			Study design	Population
			Consciousness	Agitation	Ventilator synchrony	Pain	Other	Physiologic variables	# Subscales	Total # items	Range of response options		
14	Hughes 1994 [25]	–	√	–	√	–	Reaction to tracheal suction	RR	–	8 (4 numerical + 3 yes/no + the RR value)	For the continuous items: in infants: 4 to 19, in children: 4 to 15	Case series of midazolam (assessment of incidence of hallucinations and prolonged ventilation)	53 ventilated paediatric pts (19 were enrolled in a reliability study of the score)
15	Miller 1994 [26]	–	√	√	–	–	–	–	–	1 numerical	0 to 6	RCT of 3 different dosages of midazolam	30 post abdominal surgery adult pts
16	Riker 1994 [27]	Sedation Agitation Scale (SAS)	√	√	–	–	–	–	–	1 numerical	– 3 to + 3	Case series of haloperidol	8 ventilated adult pts
17	Eddleston 1995 [28]	–	√	√	√	–	Reaction to tracheal suction	–	–	4 numerical	1 to 17	Case series of propofol during continuous haemodiafiltration	10 ventilated adult pts
18	Chamorro 1996 [29]	–	√	√	√	–	–	–	2 (level of sedation & effectiveness)	5 numerical & 4 yes/no	Subscale A: 7 to 19 subscale B: 0 to 4	RCT of propofol vs midazolam	98 ventilated adult pts
19	Mallick 1996 [30]	–	√	√	√	–	anxiety, reaction to tracheal suction	–	–	1 numerical	1 to 6	RCT of endotracheal instillation of lidocaine vs placebo	30 ventilated adults pts
20	Parkinson 1997 [31]	–	√	√	–	–	Reaction to tracheal suction	–	–	1 numerical	1 to 5	RCT of chloral hydrate & promethazine vs midazolam and reliability study	44 (reliability study) & 44 (RCT) paediatric pts
21	Weinbroum 1997 [32]	–	√	√	–	–	–	–	2 (consciousness & agitation)	1 numerical & 1 yes/no	Numerical: 1 to 5	RCT of propofol vs midazolam	67 ventilated adult pts
22	Sanchez-Izquierdo-Riera 1998 [33]	–	√	√	–	–	–	–	–	1 numerical	1 to 4	RCT of midazolam vs propofol vs combination of the 2 drugs	100 ventilated trauma adult pts
23	Checketts 1998 [34]	–	√	√	–	–	–	–	–	1 numerical	0 to 6	RCT of PCA vs computer-controlled infusion of alfentanil	105 ventilated post cardiac bypass surgery adult pts
24	Brown 1998 [35]	–	√	√	–	–	–	HR and RR	–	1 numerical	0 to 5	Prospective assessment of a sedation algorithm for sleep	78 adult pts
25	Devlin 1999 [36]	Motor Activity Assessment Scale (MAA)	√	√	–	–	–	–	–	1 numerical	0 to 6	Scale development and validation study	25 ventilated adult surgical pts

pts: patients
RCT: randomized controlled trial
GCS: Glasgow Coma Scale
MABP: mean arterial blood pressure
HR: heart rate
RR: respiratory rate
PCA: Patient-controlled analgesia
1 Also called: "Newcastle Sedation Scale"

Table 2 Scales used for assessing sedation in critically ill patients which underwent formal methodological evaluation

	Name	Original report	Validation study	Population	Validation process			
					Number of assessments	Internal consistency	Reliability	Validity
1	Ramsay scale	Ramsay , 1974 [9]	Carrasco, 1993 (Abstract) [39]	102 adult pts	1040 measurements (# of raters unclear)	–	–	Validity vs GCS modified by Cook & Palma: correlation coefficient $r = 0.89$ to 0.92 on 4 different occasions, p value NR
			Riker, 1999 [40]	45 adult pts	69 assessments by 2 raters	–	Correlation coefficient $r = 0.87$, $p < 0.001$ Weighted kappa 0.87, $p < 0.001$	Validity vs SAS: correlation coefficient $r = 0.91$, $p < 0.001$
			Gimeno, 1999 [38]	30 adult pts	2 raters (# of assessments unclear)	–	Weighted kappa 0.79, $p < 0.0001$	–
2	GCS modified by Cook & Palma	Cook, 1989 [16]	Carrasco, 1993 (Abstract) [39]	102 adult pts	1040 measurements (# of raters unclear)	–	–	Validity vs Ramsay scale: correlation coefficient $r = 0.89$ to 0.92 on 4 different occasions, p value NR
			Gimeno, 1999 [38]	30 adult pts	2 raters (# of assessments unclear)	–	Weighted kappa 0.94, $p < 0.00001$	
3	Vancouver Sedative Recovery Scale (VSRS)	Macnab, 1991 [18]	Macnab, 1991 [18]	91 postoperative or trauma paediatric pts	91 assessments by 2 raters	Cronbach's alpha 0.87	Intraclass correlation coefficient $r = 0.90$; intraclass correlation coefficient for each of the 12 separate items $r = 0.67$ to 0.89	–
4	Comfort scale	Ambuel, 1992 [21]	Ambuel, 1992 [21]	37 ventilated paediatric pts	37 assessments by 3 raters	Cronbach's alpha 0.90	Correlation coefficient 0.84, $p < 0.01$ (1) Intraclass correlation coefficient for each of the 8 separate items: 0.51 to 0.75	Validity vs expert VAS assessment: Correlation coefficient $r = 0.75$, $p < 0.01$ (1)
5	New Sheffield Sedation Scale	Laing, 1992 [23]	Olleveant, 1998 [41]	50 adult patients	100 assessments by 2 raters	–	Kappa 0.73, p value NR	–
6	–	Hughes, 1994 [25]	Hughes, 1994 [25]	19 paediatric pts	38 assessments by to raters	–	Correlation coefficient $r = 0.94$, p value NR	–
7	Sedation Agitation Scale (SAS)	Riker, 1994 [27]	Riker, 1999 [40]	45 adult pts	69 assessments by 2 raters	–	Correlation coefficient $r = 0.91$, $p < 0.001$ Weighted kappa 0.92, $p < 0.001$	Validity vs Ramsay scale: correlation coefficient $r = 0.91$, $p < 0.001$; validity vs Harris scale: correlation coefficient $r = 0.93$, $p < 0.001$
8	–	Parkinson, 1997 [31]	Parkinson, 1997 [31]	44 paediatric pts	77 assessments by 2 raters	–	Correlation coefficient $r = 0.98$, p value NR	–
9	Motor Activity Assessment Scale (MAAS)	Devlin, 1999 [36]	Devlin, 1999 [36]	25 adult surgical pts	400 assessments by 2 raters	–	Kappa 0.83 (95 % Cl 0.72–0.94)	Validity vs expert VAS assessment: GEE $p < 0.001$ Validity vs change in HR: GEE $p < 0.001$ Validity vs change in BP: GEE $p < 0.001$ Validity vs recent occurrence of agitation-related events: GEE $p < 0.001$

pts: patients
#: number
GCS: Glasgow coma scale
NR: not recorded
SAS: Sedation Agitation Scale
VAS: visual analogue scale
GEE: generalized estimating equation approach to regression
HR: heart rate
BP: blood pressure
1 Pearson's correlation

Table 3 Description of scales tested for both reliability and validity: the Ramsay scale [9], the Comfort Scale [21], the Sedation-Agitation-Scale [27] and the Motor Activity Assessment Scale [36]

Ramsay Scale [9]		
Awake levels:		
patient anxious or agitated or both	1	
patient co-operative, orientated and tranquil	2	
patient responds to commands only	3	
Asleep levels:		
a brisk response to a light glabellar tap	4	
a sluggish response to a light glabellar tap	5	
no response	6	
Comfort Scale [21]		
Alertness:		
Deeply asleep	1	
Lightly asleep	2	
Drowsy	3	
Fully awake and alert	4	
Hyper-alert	5	
Calmness/Agitation:		
Calm	1	
Slightly anxious	2	
Anxious	3	
Very anxious	4	
Panicky	5	
Respiratory response:		
No coughing and no spontaneous respiration	1	
Spontaneous respiration with little or no response to ventilation	2	
Occasional cough or resistance to ventilator	3	
Actively breathes against ventilator or coughs regularly	4	
Fights ventilator; coughing or choking	5	
Physical movement:		
No movement	1	
Occasional, slight movement	2	
Frequent, slight movement	3	
Vigorous movement limited to extremities	4	
Vigorous movement including torso and head	5	
Blood pressure:		
Blood pressure below baseline	1	
Blood pressure consistently at baseline	2	
Infrequent elevations of 15 % or more (1–3 episodes)	3	
Frequent elevations of 15 % or more (more than 3 episodes)	4	
Sustained elevation $\geq$ 15 %	5	
Heart rate:		
Heart rate below baseline	1	
Heart rate consistently at baseline	2	
Infrequent elevations of 15 % or more (1–3 episodes)	3	
Frequent elevations of 15 % or more (more than 3 episodes)	4	
Sustained elevation $\geq$ 15 %	5	
Muscle tone:		
Muscle totally relaxed	1	
Reduced muscle tone	2	
Normal muscle tone	3	
Increased muscle tone and flexion of fingers and toes	4	
Extreme muscle rigidity and flexion of fingers and toes	5	
Facial tension:		
Facial muscles totally relaxed	1	
Facial muscle tone normal; no facial muscle tension evident	2	
Tension evident in some facial muscles	3	
Tension evident throughout facial muscles	4	
Facial muscles contorted and grimacing	5	
Sedation-Agitation Scale [27]		
Dangerous agitation	Pulling at ET tube, trying to remove catheters, climbing over bed rail, striking at staff, trashing side-to-side	7

Table 3 Continued

Very agitated	Does not calm, despite frequent verbal reminding of limits; requires physical restraints, biting ET tube	6
Agitated	Anxious or mildly agitated, attempting to sit up, calms down to verbal instructions	5
Calm and cooperative	Calm, awakens easily, follows command	4
Sedated	Difficult to arouse, awakens to verbal stimuli or gentle shaking but drifts off again, follows simple commands	3
Very sedated	Arouses to physical stimuli but does not communicate or follows command, may move spontaneously	2
Unrousable	Minimal or no response to noxious stimuli, does not communicate or follows command	1
Motor Activity Assessment Scale [36]		
Unresponsive	Does not move with noxious stimulus	0
Responsive only to noxious stimuli	Open eyes OR raises eyebrows OR turns head toward stimulus OR moves limb with noxious stimulus	1
Responsive to touch or name	Open eyes OR raises eyebrows OR turns head toward stimulus OR moves limb when touched or name is loudly spoken	2
Calm and cooperative	No external stimulus is required to elicit movement AND patient is adjusting sheets or clothes purposefully and follows command	3
Restless and cooperative	No external stimulus is required to elicit movement AND patient is picking at sheets or tubes OR uncovering self and follows command	4
Agitated	No external stimulus is required to o elicit movement AND attempting to sit up OR moves limbs out of bed AND does not consistently follow commands (e.g., will lie down when asked but soon reverts back to attempts to sit up or move limbs out of bed)	5
Dangerously agitated, uncooperative	No external stimulus is required to o elicit movement AND patient is pulling at tubes or catheters OR trashing side to side OR striking at staff OR trying to climb out of bed AND does not calm down when asked	6

statistical significance. We identified no studies of sedation effectiveness that measured responsiveness.

Discussion

Numerous instruments to evaluate both consciousness and one or more other domains relevant to sedation have been developed and employed to measure the effect of sedative drugs in clinical ICU investigations. Some of these instruments have been tested to determine their clinimetric properties, but none have been tested concomitantly for reliability, validity and responsiveness.

In addition to measuring consciousness, the most frequently incorporated domains related to the use of sedatives in the instruments we reviewed were agitation and tolerance to mechanical ventilation. Assessment of anxiety, which might result from invasive procedures, uncomfortable therapeutic interventions, or fear and uncertainty about prognosis, was part of only three scores we identified [9, 24, 30]. This omission in most studies may reflect the difficulty that caregivers have in assessing anxiety in ICU patients. However, recognition and treatment of anxiety may improve patient wellbeing and minimize the risk of adverse consequences such as serious hemodynamic disturbances. Verbal and gestural communications, through which patients may express their degree of anxiety, are rarely possible in severely ill mechanically ventilated ICU patients. Agitation and ventilator dysynchrony, which might reflect anxiety, are indirectly related parameters; hemodynamic variables such as tachycardia or increased blood pressure are non-specific in this population. In the three studies in this systematic review in which anxiety was part of an instrument [9, 24, 30], clinicians were asked to specify

whether the patient was or was not anxious, without precise descriptors. This, too, may be a testimony to the difficulty in accurately labeling anxiety in semiconscious or unconscious patients, and the challenges in evaluating treatment for anxiety.

Most of the sedation instruments in this systematic review are constituted by one item with a categorical grading. Such an instrument can be simple to use at the bedside. However, the use of a one-item instrument might not be appropriate when different conditions (e.g., consciousness and agitation) are assessed in the same item. For example, a patient can be concomitantly agitated (level 1 on the Ramsay Scale) and exhibit only a brisk response to a light glabellar tap (level 4 on the Ramsay Scale). Thus, the use of a single item to assess two or more different aspects of sedation can lead to loss of clinically important information and systematic or random measurement error.

The Ramsay Scale (Table 3) was first published in 1974 to describe the effect of alphaxalone-alphadolone in a series of 30 ICU patients using a 6-point scale ranging from anxious or agitated to asleep with no response [9]. Since this publication, the Ramsay scale has been used by many investigators and was employed in 20 of 31 randomized controlled trials comparing sedative agents with respect to quality of sedation, or duration of mechanical ventilation [42]. Although this scale may be widely considered as the best instrument for measuring sedation in critically ill patients, data on its clinimetric properties have been available only recently. The score exhibits a satisfactory inter-rater reliability [38, 40] and a high correlation with the GCS modified by Cook and Palma [39] and with the SAS [40]. The high correlations observed between these distinct instruments provides some measure of validation, but this is tempered by the extent to which the instruments include items that are very similar in content and structure.

Apart from the Ramsay scale, three instruments included in this systematic review were tested for reliability and validity simultaneously in a full-text article with a comprehensive description of the validation process (Table 3). One instrument was developed in the pediatric ICU population [21], the two others in the adult ICU population [36, 40]. The Comfort Scale is constituted from eight items with response options ranging from 1 to 5. This was developed to measure not only the level of consciousness but also other parameters such as face grimacing, muscle tone, physiological values and the level of agitation, all considered as possibly reflecting tolerance to the pediatric ICU environment [21]. The Comfort Scale exhibited a good inter-rater reliability when assessed with a Pearson's correlation (coefficient 0.84), and also when evaluated with a more rigorous analysis (intraclass correlation coefficients for each of the 8 separate items 0.51 to 0.75). The investigators of the Comfort Scale also provided evidence for validity in that the correlation between their scores and an independent expert assessment on a visual analogue scale yielded a Pearson's correlation of 0.75. This validation is stronger than other validation measures because the visual analogue scale is different in structure and content from the index instrument. Nevertheless, validation could have been strengthened further by additional comparisons to a number of alternative measures of sedation.

The Sedation-Agitation Scale (SAS) developed to assess consciousness and agitation in adult ICU patients is comprised of one item, with response options ranging from 1 to 7 [40]. Excellent inter-rater reliability was found (kappa 0.92), and construct validity has been demonstrated against 2 other different sedation scales (correlation coefficient with Ramsay scale $r = 0.91$ and with Harris scale $r = 0.93$).

The recently reported Motor Activity Assessment Scale (MAAS), similar in its structure to the SAS, is comprised of one item with response options ranging from 0 to 6 and was developed in surgical ICU patients [36]. Good inter-rater reliability was noted (kappa 0.83) and high correlation was found with a visual analogue scale, suggesting construct validity. Construct validity is also supported by a significant correlation between the MAAS and heart rate variations, respiratory rate variations, and the number of agitation-related events recorded during the 30 min before the assessment.

Since the level of consciousness and sedation requirements of ICU patients vary over time, the capacity of an instrument to detect change in a patient's clinical condition is a desirable measurement attribute [43]. The responsiveness of such a scale to sedative initiation, modification in the drug dosage, and withdrawal of sedation, is therefore an important property that the instrument should demonstrate. However, we identified no adult or pediatric sedation instrument in this systematic review that was tested for responsiveness.

Rigorous measurement of the level of sedation in critically ill patients is a challenge. The widespread use of sedatives, and the growing number of randomized controlled trials evaluating different classes of drugs and different drugs in the same class have highlighted the need to measure sedation effectiveness more precisely. The concept of "quality of sedation" may be elusive but is nonetheless important. Moreover, depth of sedation is increasingly recognized as contributing to delayed weaning from mechanical ventilation and associated ICU complications. The studies we included in this review acknowledged the central role of bedside nurses in making such assessments in practice, and suggest their central role as potential collaborators in this field of research. As more cost-effectiveness studies on sedation are being conducted, we will need rigorously developed and evaluated instruments. The high reliabil-

ity exhibited by the Ramsay scale, the SAS and the MAAS in the adult population, and by the Comfort scale in the pediatric population, and their high correlation with one another (and/or with the clinicians' global rating) suggest that they provide satisfactory measures of sedation at one point in time. However, none of these instruments have been evaluated with respect to their ability to detect changes in sedation status over time. Further exploration of the measurement properties of these sedation instruments would strengthen the confidence we have in using them to monitor the sedation of individual patients in the intensive care unit.

References

1. Merriman HM (1981) The techniques used to sedate ventilated patients. A survey of methods used in 34 ICUs in Great Britain. Intensive Care Med 7: 217–24
2. Bion JF, Ledingham IM (1987) Sedation in intensive care – a postal survey (letter). Intensive Care Med 13: 215–6
3. Hansen-Flaschen JH, Brazinsky S, Basile C, Lanken PN (1991) Use of sedating drugs and neuromuscular blocking agents in patients requiring mechanical ventilation for respiratory failure. A national survey. JAMA 266: 2870–5
4. Christensen BV, Thunedborg LP (1999) Use of sedatives, analgesics and neuromuscular blocking agents in Danish ICUs 1996/97. A national survey. Intensive Care Med 25: 186–91
5. Kollef MH, Levy NT, Ahrens TS, Schaiff R, Prentice D, Sherman G (1998) The use of continuous i.v. sedation is associated with prolongation of mechanical ventilation. Chest 114: 541–8
6. Cook DJ, Walter SD, Cook RJ, Griffith LE, Guyatt GH, Leasa D, Jaeschke RZ, Brun-Buisson C (1998) Incidence of and risk factors for ventilator-associated pneumonia in critically ill patients. Ann Intern Med 129: 433–40
7. Meade MO, Cook DJ, Kernerman P, Bernard G (1997) How to use articles about harm: the relationship between high tidal volumes, ventilating pressures, and ventilator-induced lung injury. Crit Care Med 25: 1915–22
8. DeJonghe B, Cook D, Sharshar T, Lefaucheur J, Carlet J, Outin H (1998) Acquired neuromuscular disorders in critically ill patients: a systematic review. Intensive Care Med 24: 1242–1250
9. Ramsay M, Savege T, Simpson B, Goodwin R (1974) Controlled sedation with Alphaxalone-Alphadolone. BMJ ii: 656–9
10. Simmons LE, Riker RR, Prato BS, Fraser GL (1999) Assessing sedation during intensive care unit mechanical ventilation with the Bispectral Index and the Sedation-Agitation Scale. Crit Care Med 27: 1499–504
11. Schulte-Tamburen AM, Scheier J, Briegel J, Schwender D, Peter K (1999) Comparison of five sedation scoring systems by means of auditory evoked potentials. Intensive Care Med 25: 377–82
12. Haberthur C, Lehmann F, Ritz R (1996) Assessment of depth of midazolam sedation using objective parameters. Intensive Care Med 22: 1385–90
13. Yate PM, Thomas D, Short SM, Sebel PS, Morton J (1986) Comparison of infusions of alfentanil or pethidine for sedation of ventilated patients on the ITU. Br J Anaesth 58: 1091–9
14. Cohen AT, Kelly DR (1987) Assessment of alfentanil by intravenous infusion as long-term sedation in intensive care. Anaesthesia 42: 545–8
15. Adams HA, Biscoping J, Russ W, Bachmann B, Ratthey K, Hempelmann G (1988) Untersuchungen zur sedativ-analgetischen medikation beatmungspflichtiger intensivpatienten. Anaesthesist 37: 268–76
16. Cook S, Palma O (1989) Propofol as a sole agent for prolonged infusion in intensive care. J Drug Dev 4 (suppl 2): 65–67
17. O'Sullivan G, Park G (1990) The assessment of sedation in critically ill patients. Clin Intensive Care 1: 116–122
18. Macnab AJ, Levine M, Glick N, Susak L, Baker-Brown G (1991) A research tool for measurement of recovery from sedation: the Vancouver Sedative Recovery Scale. J Pediatr Surg 26: 1263–7
19. Hartwig S, Roth B, Theisohn M (1991) Clinical experience with continuous intravenous sedation using midazolam and fentanyl in the pediatric intensive care unit. Eur J Pediatr 150: 784–8
20. Harris CE, O'Donnell C, Macmillan RR, Mostafa SB (1991) Use of propofol by infusion for sedation of patients undergoing haemofiltration: assessment of the effect of haemofiltration on the level of sedation and on blood propofol concentration. J Drug Dev 4 (suppl 3): 37–39
21. Ambuel B, Hamlett KW, Marx CM, Blumer JL (1992) Assessing distress in pediatric intensive care environments: the COMFORT scale. J Pediatr Psychol 17: 95–109
22. Chaudhri S, Kenny GN (1992) Sedation after cardiac bypass surgery: comparison of propofol and midazolam in the presence of a computerized closed loop arterial pressure controller. Br J Anaesth 68: 98–99
23. Laing AS (1992) The applicability of a new sedation scale for intensive care. Intensive Crit Care Nurs 8: 149–52
24. Spencer EM, Green JL, Willatts SM (1994) Continuous monitoring of depth of sedation by EEG spectral analysis in patients requiring mechanical ventilation. Br J Anaesth 73: 649–54
25. Hughes J, Gill A, Leach HJ, Nunn AJ, Billingham I, Ratcliffe J, Thornington R, Choonara I (1994) A prospective study of the adverse effects of midazolam on withdrawal in critically ill children. Acta Paediatr 83: 1194–9
26. Miller DR, Martineau RJ, Hull KA, Vallee F, Le Bel M (1994) Optimizing sedation following major vascular surgery: a double-blind study of midazolam administered by continuous infusion. Can J Anaesth 41: 782–93
27. Riker RR, Fraser GL, Cox PM (1994) Continuous infusion of haloperidol controls agitation in critically ill patients. Crit Care Med 22: 433–40
28. Eddleston JM, Pollard BJ, Blades JF, Doran B (1995) The use of propofol for sedation of critically ill patients undergoing haemodiafiltration. Intensive Care Med 21: 342–7
29. Chamorro C, de Latorre FJ, Montero A, Sanchez-Izquierdo JA, Jareno A, Moreno JA, Gonzalez E, Barrios M, Carpintero JL, Martin-Santos F, Otero B, Ginestal R (1996) Comparative study of propofol versus midazolam in the sedation of critically ill patients: results of a prospective, randomized, multicenter trial. Crit Care Med 24: 932–9
30. Mallick A, Smith SN, Bodenham AR (1996) Local anaesthesia to the airway reduces sedation requirements in patients undergoing artificial ventilation. Br J Anaesth 77: 731–4

31. Parkinson L, Hughes J, Gill A, Billingham I, Ratcliffe J, Choonara I (1997) A randomized controlled trial of sedation in the critically ill. Paediatr Anaesth 7: 405–10
32. Weinbroum AA, Halpern P, Rudick V, Sorkine P, Freedman M, Geller E (1997) Midazolam versus propofol for long-term sedation in the ICU: a randomized prospective comparison. Intensive Care Med 23: 1258–63
33. Sanchez-Izquierdo-Riera JA, Caballero-Cubedo RE, Perez-Vela JL, Ambros-Checa A, Cantalapiedra-Santiago JA, Alted-Lopez E (1998) Propofol versus midazolam: safety and efficacy for sedating the severe trauma patient. Anesth Analg 86: 1219–24
34. Checketts MR, Gilhooly CJ, Kenny GN (1998) Patient-maintained analgesia with target-controlled alfentanil infusion after cardiac surgery: a comparison with morphine PCA. Br J Anaesth 80: 748–51
35. Brown G, Scott W (1998) An assessment of a sedative algorithm for sleep in an intensive care unit. Off J Can Assoc Crit Care Nurs 9: 20–4
36. Devlin JW, Boleski G, Mlynarek M, Nerenz DR, Peterson E, Jankowski M, Horst HM, Zarowitz BJ (1999) Motor Activity Assessment Scale: a valid and reliable sedation scale for use with mechanically ventilated patients in an adult surgical intensive care unit. Crit Care Med 27: 1271–5
37. Kenny G (1989) Investigation of propofol for sedation following cardiac bypass surgery. J Drug Dev 2 (suppl 2): 125–6
38. Gimeno G, Alcolea M, Arana M, Bartolome N, Cruspinera A, Figueras MJ, Mas T, Peregrin A, Ventura R (1999) Son todas las escalas de sedacion igual de utiles para la valoracion de enfermeria? Enferm Intensiva 10: 3–12
39. Carrasco G, Molina R, Costa J, Soler J, Paniagua J, Cabre L (1993) Usefulness of sedation scales in ICU. A comparative randomized study in patients sedated with propofol, midazolam or opiates plus benzodiazepines. Intensive Care Med 18(suppl 2): 158
40. Riker RR, Picard JT, Fraser GL (1999) Prospective evaluation of the Sedation-Agitation Scale for adult critically ill patients. Crit Care Med 27: 1325–9
41. Olleveant N, Humphris G, Roe B (1998) A reliability study of the modified new Sheffield Sedation Scale. Nurs Crit Care 3: 83–8
42. Ostermann ME, Seiferling RA, Keenan SP, Sibbald WJ (1999) A systematic review of the evidence supporting the use of sedative agents in the intensive care unit: which agent should be used? JAMA (in press)
43. Guyatt G, Walter S, Norman G (1987) Measuring change over time: assessing the usefulness of evaluative instruments. J Chronic Dis 40: 171–8