HHS Public Access

Author manuscript

Arch Phys Med Rehabil. Author manuscript; available in PMC 2018 August 16.

Published in final edited form as:

Arch Phys Med Rehabil. 2016 August ; 97(8): 1295–1300.e4. doi:10.1016/j.apmr.2016.02.009.

Detection and interpretation of impossible and improbable Coma Recovery Scale-Revised scores

C Chatelle, PhD[1,2,3,*], **YG Bodien, PhD**[1,*], **C Carlowicz, BA**[1], **S. Wannez, MSc**[3], **V. Charland-Verville, MSc**[3], **O. Gosseries, MSc**[3,4], **S Laureys, MD, PhD**[3], **RT Seel, PhD**[5], and **JT Giacino, PhD**[1]

[1]Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital and Harvard Medical School, Boston, MA, USA

[2]Laboratory for NeuroImaging of Coma and Consciousness, Massachusetts General Hospital, Boston, MA, USA

[3]Coma Science Group, GIGA and Cyclotron Centre, University and University Hospital of Liège, Belgium

[4]Departments of Psychology and Psychiatry, University of Wisconsin, Madison, WI, USA

[5]Crawford Research Institute, Shepherd Center, Atlanta, GA, USA

Abstract

Objectives—To determine the frequency with which specific Coma Recovery Scale-Revised (CRS-R) subscale scores co-occur as a means of providing clinicians and researchers with an empirical method of assessing CRS-R data quality.

Design—We retrospectively analyzed CRS-R subscale scores in hospital inpatients diagnosed with DoC to identify impossible and improbable subscore combinations as a means of detecting inaccurate and unusual scores. Impossible subscore combinations were based on violations of CRS-R scoring guidelines. To determine improbable subscore combinations, we relied on the Mahalanobis distance which detects outliers within a distribution of scores. Subscore pairs that were not observed at all in the database (i.e., frequency of occurrence = 0%) were also considered improbable.

Setting—Specialized DOC program and University hospital.

Participants—1190 patients diagnosed with DoC (coma= 76, VS= 464, MCS= 586, EMCS= 64; 794 males; mean age= 43±20 years; traumatic etiology= 747; time post injury= 162±568 days).

Interventions—Not applicable.

Corresponding author: Camille Chatelle, Laboratory for NeuroImaging of Coma and Consciousness, Massachusetts General Hospital, 175 Cambridge street, Suite 300, Boston, MA, 02114, Tel: +32 484 079 361, cchatelle@partners.org.

*Contributed equally to this work.

Main Outcome Measure(s)—Impossible and improbable CRS-R subscore combinations.

Results—Of the 1190 CRS-R profiles analyzed, 4.7% were excluded because they met scoring criteria for impossible co-occurrence. Among the 1137 remaining profiles, 12.2% (41/336) of possible subscore combinations were classified as improbable.

Conclusions—Clinicians and researchers should take steps to ensure the accuracy of CRS-R scores. To minimize the risk of diagnostic error and erroneous research findings, we have identified 9 impossible and 36 improbable CRS-R subscore combinations. The presence of any one of these subscore combinations should trigger additional data quality review.

Keywords

Assessment; disorders of consciousness; vegetative state; minimally conscious state; data quality

Introduction

Patients surviving severe injury to the brain may remain in a coma for up to several weeks before transitioning into either a vegetative state (VS, also coined “unresponsive wakefulness syndrome”[1]) or a minimally conscious state. Individuals in VS show periods of wakefulness of varying duration but do not demonstrate any behavioral signs of consciousness [2]. MCS is a severely altered state of consciousness in which the person demonstrates minimal but definite behavioral evidence of comprehension of simple commands, intelligible verbalizations, gestural or verbal yes-no responses, object manipulation or non-reflexive behaviors that occur in contingent relation to specific environmental stimuli (e.g., visual pursuit) [3]. Emergence from MCS (EMCS) is marked by the reemergence of a reliable yes-no communication system and/or functional object use [3]. Detecting behavioral signs of awareness and differentiating between these disorders of consciousness (DoC) can be challenging and has led to the development of standardized approaches to diagnostic assessment [4, 5]. The Coma Recovery Scale-Revised (CRS-R) [6] has strong evidence of reliability and validity for assessment of patients with DoC, based on a recent systematic review completed by the Clinical Practice Committee of the American Congress of Rehabilitation Medicine [7].

The CRS-R consists of 23 hierarchically-organized items parcellated into 6 subscales designed to interrogate functional brain networks responsible for mediating auditory, visual, motor, language and arousal functions. Weighted scores are assigned to reflect the presence or absence of specific behaviors, ranging from brain stem reflexes to those that are cognitively-mediated (see table 1). All assessment procedures and scoring criteria are operationally-defined and the diagnostic criteria for coma, VS, MCS and EMCS are embedded within the scale. The total score can be used to gauge the general trajectory of recovery over time as higher scores reflect progressively increasing levels of neurobehavioral function [8].

The hierarchical framework of the items included in the CRS-R is supported by psychometric studies demonstrating the properties of unidimensionality (i.e., all items on the scale are related to a single underlying construct), monotonicity (i.e., as the total score

increases, the score on any single item increases or remains stable), mutual independence (i.e., the only source of correlation between two or more subscales is the underlying construct measured by the scale as a whole) and invariant item ordering (i.e., for any given ability level, the order of difficulty of items remains the same) [9, 10].

In view of the broad use of the CRS-R in research and clinical practice, we were interested in developing an empirical approach to data quality analysis. More specifically, our objective was to develop a methodology that could be used to alert the examiner to erroneous or unusual scores. Based on the previously-described psychometric characteristics of the CRS-R, the probability of receiving a specific score on a given subscale should be largely related to the scores received on the other subscales. Thus, establishing the incidence of specific subscale score combinations may serve to identify rare subscore combinations that could indicate an invalid assessment due to use of improper administration or scoring procedures. Alternatively, improbable subscore combinations may signal the presence of an underlying functional impairment, which may have diagnostic relevance. For example, a very low score on the auditory subscale coupled with a high score on the motor subscale raises the possibility of an underlying aphasia or impairment in auditory processing. Detection of highly-improbable subscore combinations can serve as a “red flag,” triggering the need for further investigation.

The primary aim of this retrospective study was to determine the probability with which specific CRS-R subscale score combinations occur as a means to establish an empirical method of data quality analysis. We hypothesized that subscore combinations that fail to respect the hierarchical structure of the scale (e.g., scores that concurrently fall at the floor and ceiling of two different subscales) will have a low probability of occurrence. We also identified a list of impossible subscore combinations. That is, scores that, in combination, violate the CRS-R’s standardized scoring procedures. For example, object recognition on the Visual subscale cannot co-occur with auditory localization on the Auditory subscale. The presence of object recognition requires command-following, however, scoring auditory localization as the best response on the Auditory subscale implies the absence of command-following.

Methods

Demographic data and CRS-R scores were retrospectively obtained from the databases of two specialized inpatient rehabilitation programs serving patients with DoC in the United States (n=767) and an acute care hospital located in Belgium (n=423). The study was approved by the Institutional Review Board (USA) or Ethics Committee (Belgium) of each site. Inclusion criteria were history of severe acquired brain injury with DoC at the time of assessment (i.e., coma, VS, MCS or EMCS), age 16 or older, and fluent in English or French (or translator available during the assessment). CRS-R scores were obtained on admission for 1190 patients diagnosed with DoC (coma= 76, VS= 464, MCS= 586, EMCS= 64; 794 males; mean age= 43±20 years; traumatic etiology= 747; time post injury= 162±568 days; mean admission CRS-R total score= 8.5±5.1).

To identify impossible subscore combinations, we reviewed the CRS-R Administration and Scoring Manual (http://www.tbims.org/combi/crs/) and identified subscores which cannot co-occur without violating the scoring rules of one or both items (see table 2) leaving 336 possible subscore combinations. All assessments that contained at least one impossible subscore combination were excluded.

We used a 4-step process to determine improbable subscore combinations (see table 3). First, we calculated the Mahalanobis distance for each of the 336 pairs of subscores. The Mahalanobis distance is a statistical measure based on a chi-squared distribution that is commonly used to detect outliers. It measures the distance of a point P from the centroid (i.e., multidimensional mean) of a distribution D, given the covariance (i.e., multidimensional variance) of D (Mahalanobis 1936). If P is at the centroid of D, the distance will be zero, and the more it differs from the centroid, the higher the distance (i.e., the greater the standard deviation) will be. The advantage of this method is that is takes into account the existing correlation between variables (i.e., CRS-R subscales) included in the distribution. We then calculated whether any of the Mahalanobis distances (i.e. test statistics) exceeded the alpha critical chi-square values at $p<.05$, $p<.02$, $p<.01$, and $p<.001$ ($df=2$) and determined, for each of the thresholds, the number of subjects identified as having at least one improbable score combination. We selected $p<.001$ as the criterion threshold as this cutoff identified a maximum of 5% of the subjects with one or more improbable subscore combinations. Subscore pairs that were not observed at all in the database (i.e., frequency of occurrence= 0%) were also considered improbable. Finally, we removed subscore combinations that were statistically improbable but not clinically atypical. The latter category was comprised exclusively of subscore combinations in which both scores fell at the ceiling (e.g., consistent command following coincident with functional communication). These pairs occurred infrequently because CRS-R scores were obtained on admission when patients were most compromised and unlikely to attain scores at the ceiling of more than one subscale.

Results

A total of 9 subscore combinations were identified as impossible based on the CRS-R administration and scoring guidelines. Impossible subscore combinations are shown in table 2. Of the 1190 independent CRS-R profiles analyzed, 4.7% (53) were excluded from further analyses because they met the criteria for impossible scoring.

Among the 1137 remaining profiles, 12.2% of the observed combinations (41/336) were classified as improbable. These 41 combinations were visually-inspected and 1.5% (5/336) were removed from the dataset because they were not considered clinically atypical. Of the 36 remaining combinations, 5.0% (17/336) were not observed at all in the dataset and 5.6% (19/336) fell below the $p<.001$ threshold. We pooled the subscore combinations that were not observed at all with those that fell at the $p<.001$ level to arrive at the final list of 36 “improbable” subscore combinations. These 36 pairs accounted for 10.7% of all subscore combinations in the dataset (see table 3).

CRS-R users who are interested in viewing the frequency of occurrence of all 336 subscore combinations in association with the corresponding Mahalanobis distance and probability level should refer to supplementary table S1. This table enables users to adjust the stringency of the cut-off for flagging subscore combinations that may require further analysis.

Discussion

The CRS-R is a well-established standardized behavioral assessment measure designed specifically for use in patients with DoC. Although psychometric studies have consistently demonstrated strong interrater and test-retest reliability [6, 11], CRS-R scores are subject to inaccuracy attributable to examiner error and other confounding factors that can lead to misinterpretation of results. The aim of this study was to develop an empirical method of flagging subscore combinations that require further scrutiny, either because they violate CRS-R administration and scoring guidelines (i.e., “impossible”), or because they rarely or never co-occur (i.e., “improbable”). We have tabulated the impossible and improbable subscore combinations to serve as a reference for CRS-R users. Knowledge of impossible and improbable subscore combinations may aid investigators responsible for conducting CRS-R data quality analysis. Clinicians engaged in diagnostic assessment may rely on improbable score combinations to signal the presence of an underlying neurological or physical impairment that may require further assessment.

Of some concern, we found that impossible subscore combinations occurred in 4.7% of our cases (n=53), suggesting that examiner error or encoding is not infrequent. When impossible subscore combinations are observed, the investigator should attempt to confirm that the CRS-R administration and scoring guidelines were adhered to during data acquisition and verify that values were properly recorded on data forms. We also retrospectively interrogated a large CRS-R dataset (n=1190) and employed quantitative methods to determine the probability with which CRS-R subscores co-occurred. Approximately 11% of all possible subscore combinations were identified as improbable, either because they were not observed at all (5%) or because the frequency of occurrence fell at the $p < .001$ threshold (6 %). Improbable subscore combinations may alert the examiner to scoring errors or unusual neurobehavioral findings that should be further investigated. Low probability subscore combinations may reflect the impact of peripheral injuries or focal disruption of specific cortical pathways on behavior and, thus, may be diagnostically-relevant. For example, the improbable combination of a score of 4 on the Auditory subscale (i.e., consistent command-following) with a score of 0 on the Motor subscale (i.e., no motor response) may be observed in patients with quadriparesis or generalized spasticity who retain language comprehension but are unable to engage efferent motor pathways. A second example is the combination of a score of 2 on the Communication subscale (i.e., reliable yes-no communication) with a score of 1 on the Oromotor/Verbal subscale (i.e., no intelligible speech). These scores may be observed in patients who retain sufficient cognitive capacity to answer basic questions reliably but cannot verbally communicate responses as the result of oromotor weakness or apraxia of speech. Prospective studies of the putative causes of improbable score combinations conducted in patients with known functional impairments would further inform the diagnostic utility of improbable scores.

When improbable subscore combinations are noted in the context of research, we suggest the investigator attempt to verify that the CRS-R administration and scoring guidelines were adhered to during data acquisition and check for data transcription errors. In the clinical domain, attempts should be made to replicate the unusual finding using a second blinded examiner.

Study Limitations

The findings from this study should be viewed in the context of several limitations. First, our sample was under-represented at the lower and upper limit of the CRS-R range. That is, in comparison to the number of subjects who were in VS and MCS, significantly fewer met criteria for EMCS. This is not unexpected given that the CRS-R was intended to monitor recovery from coma through reemergence of communication. This natural skew in the distribution likely accounts for some co-occurring high subscale scores exceeding the threshold for improbability. As a result, because we did not meet the assumption of multivariate normal distribution, our findings should be interpreted with caution. Second, we did not investigate whether improbable subscore combinations differ between patients with different demographic or injury characteristics. Replication in a larger sample will be necessary to discern whether improbable subscore combinations vary as a function of age, injury severity, chronicity or other factors. Finally, as with all retrospective analyses, we could not control for factors that may have influenced the results, including subject selection bias and the training background and level of experience of the examiners. We encourage prospective studies to investigate whether improbable scores differ relative to patient (e.g., blindness, deafness, aphasia, apraxia, etc.) and examiner (e.g., novice v. expert) characteristics (Lovstad, et al. 2010).

Conclusion

Clinicians and researchers should take steps to ensure the accuracy of CRS-R scores. To minimize the risk of diagnostic error and erroneous research findings, we developed an empirical approach to identify impossible and improbable CRS-R subscore combinations. This procedure can be used to alert the examiner to the need for additional data quality review.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was conducted with support from a grant from the National Institute on Disability, Independent Living and Rehabilitation Research, Department of Health and Human Services, Administration for Community Living (90DP0039-01-00), The James S. McDonnell Foundation, Harvard Catalyst | The Harvard Clinical and Translational Science Center (National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health Award UL1 TR001102), financial contributions from Harvard University and its affiliated academic healthcare centers, The Belgian American Educational Foundation, Wallonie Bruxelles International Concerted Research Action, the Belgian Funds for Scientific Research (FRS), European Commission and Fondazione Europea di Ricerca Biomedica. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funders. The authors have no conflicts of interest.

Abbreviations

CRS-R	Coma Recovery Scale-Revised
DoC	Disorders of Consciousness
VS	Vegetative State
MCS	Minimally Conscious State
EMCS	Emerged from Minimally Conscious State

References

1. Laureys S, et al. Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. BMC Med. 2010; 8:68. [PubMed: 21040571]
2. The Multi-Society Task Force on PVS. Medical aspects of the persistent vegetative state (1). The New England Journal of Medicine. 1994; 330(21):1499–508. [PubMed: 7818633]
3. Giacino JT, et al. The minimally conscious state: definition and diagnostic criteria. Neurology. 2002; 58(3):349–53. [PubMed: 11839831]
4. Schnakers C, et al. Diagnostic accuracy of the vegetative and minimally conscious state: Clinical consensus versus standardized neurobehavioral assessment. BMC Neurology. 2009; 9(1):35. [PubMed: 19622138]
5. Stender J, et al. Quantitative rates of brain glucose metabolism distinguish minimally conscious from vegetative state patients. J Cereb Blood Flow Metab. 2015; 35(1):58–65. [PubMed: 25294128]
6. Giacino J, Kalmar K, Whyte J. The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. Arch Phys Med Rehabil. 2004; 85(12):2020–2029. [PubMed: 15605342]
7. Seel RT, et al. Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research. Arch Phys Med Rehabil. 2010; 91(12):1795–813. [PubMed: 21112421]
8. Bodien Y, et al. Sensitivity and Specificity of the Coma Recovery Scale-Revised Total Score in Detection of Conscious Awareness. Archives of Physical Medicine and Rehabilitation. In Press.
9. La Porta F, et al. Can we scientifically and reliably measure the level of consciousness in vegetative and minimally conscious States? Rasch analysis of the coma recovery scale-revised. Arch Phys Med Rehabil. 2013; 94(3):527–535e1. [PubMed: 23127303]
10. Gerrard P, Zafonte R, Giacino JT. Coma Recovery Scale-Revised: evidentiary support for hierarchical grading of level of consciousness. Arch Phys Med Rehabil. 2014; 95(12):2335–41. [PubMed: 25010536]
11. Schnakers C, et al. A French validation study of the Coma Recovery Scale-Revised (CRS-R). Brain Injury. 2008; 22(10):786–792. [PubMed: 18787989]

Table 1

Coma Recovery Scale-Revised

AUDITORY FUNCTION SCALE	
4	Consistent Movement to Command *
3	Reproducible Movement to Command*
2	Localization to Sound
1	Auditory Startle
0	None
VISUAL FUNCTION SCALE	
5	Object Recognition *
4	Object Localization: Reaching *
3	Visual Pursuit *
2	Fixation *
1	Visual Startle
0	None
MOTOR FUNCTION SCALE	
6	Functional Object Use†
5	Automatic Motor Response *
4	Object Manipulation *
3	Localization to Noxious Stimulation *
2	Flexion Withdrawal
1	Abnormal Posturing
0	None/Flaccid
OROMOTOR/VERBAL FUNCTION SCALE	
3	Intelligible Verbalization *
2	Vocalization/Oral Movement
1	Oral Reflexive Movement
0	None
COMMUNICATION SCALE	
2	Functional: Accurate†
1	Non-Functional: Intentional *
0	None
AROUSAL SCALE	
3	Attention
2	Eye Opening w/o Stimulation
1	Eye Opening with Stimulation
0	Unarousable

Table 2

List of impossible subscore combinations based on CRS-R scoring guidelines.

Impossible subscore combinations	
Subscore 1	**Subscore 2**
No auditory response (A0)	Object recognition (V5)
Auditory startle (A1)	Object recognition (V5)
Localization to sound (A2)	Object recognition (V5)
No auditory response (A0)	Intentional communication (C1)
Auditory startle (A1)	Intentional communication (C1)
Localization to sound (A2)	Intentional communication (C1)
No auditory response (A0)	Functional communication (C2)
Auditory startle (A1)	Functional communication (C2)
Localization to sound (A2)	Functional communication (C2)

Table 3

List of improbable CRS-R subscore combinations.

Improbable subscore combinations		
Subscore 1	**Subscore 2**	**Possible contributing factors (when scoring errors are ruled out)**
No auditory response (A0) [+]	Functional object use (M6) [+]	Deafness
No auditory response (A0) [+]	Attention (Ar3) [+]	
Auditory startle (A1) [+]	Functional object use (M6) [+]	Aphasia; Central deafness
Localization to sound (A2) [+]	Attention (Ar3) [+]	
Consistent command following (A4) [+]	Visual startle (V1) [+]	3rd and 4th cranial nerve palsy; Ocular apraxia; Visual agnosia
Consistent command following (A4) [+]	Abnormal posturing (M1) [+]	Severe spasticity
Consistent command following (A4) [+]	Unarousable/No eye-opening (Ar0) [+]	Bilateral ptosis [‡]; Facial oedema [‡]; Eyelid apraxia [‡]
Blink to threat (V1) [+]	Functional object use (M6) [+]	Bilateral optic nerve damage; Terson's syndrome; Cortical blindness
Blink to threat (V1) [+]	Functional communication (C2) [+]	
Visual fixation (V2) [+]	Unarousable (Ar0) [+]	Ptosis [‡]; Eyelid apraxia [‡]
Object localization (V4) [+]	Unarousable (Ar0) [+]	
Object localization (V4) [+]	No motor response (M0) [+]	
Object recognition (V5) [+]	Abnormal posturing (M1) [+]	Severe spasticity
Object recognition (V5) [+]	Unarousable (Ar0) [+]	Ptosis [‡]; Eyelid apraxia [‡]
Abnormal posturing (M1) [+]	Intelligible verbalization (Ve3) [+]	Severe spasticity
Functional object use (M6) [+]	Unarousable (Ar0) [+]	Ptosis [‡]; Eyelid apraxia [‡]
Functional communication (C2) [+]	Unarousable (Ar0) [+]	
Reproducible command following (A3) [†]	Functional communication (C2) [†]	N/A [*]
Consistent command following (A4) [†]	No visual response (V0) [†]	Bilateral optic nerve damage; Terson's syndrome; Cortical blindness
Consistent command following (A4) [†]	No motor response (M0) [†]	Quadriplegia
No visual response (V0) [†]	Functional communication (C2) [†]	Bilateral optic nerve damage; Terson's syndrome; Cortical blindness
Visual fixation (V2) [†]	Functional communication (C2) [†]	3rd and 4th cranial nerve palsy; Ocular apraxia; Visual agnosia
Visual pursuit (V3) [†]	Functional communication (C2) [†]	Ocular apraxia; Visual agnosia
Object localization (V4) [†]	Functional communication (C2) [†]	Visual agnosia, hemineglect
Object recognition (V5) [†]	No motor response (M0) [†]	Quadriplegia
No motor response (M0) [†]	Functional communication (C2) [†]	Quadriplegia
Abnormal posturing (M1) [†]	Functional communication (C2) [†]	Severe spasticity
Flexion withdrawal (M2) [†]	Functional communication (C2) [†]	Severe spasticity, hypertonus or hypotonus
Localization to pain (M3) [†]	Functional communication (C2) [†]	Apraxia

Improbable subscore combinations		
Subscore 1	**Subscore 2**	**Possible contributing factors (when scoring errors are ruled out)**
Object manipulation (M4)[†]	Functional communication (C2)[†]	Severe spasticity, hypertonus or hypotonus; Apraxia
Automatic motor response (M5)[†]	Functional communication (C2)[†]	Object agnosia; Apraxia
No verbal response (Ve0)[†]	Functional communication (C2)[†]	Facial nerve palsy/Oromotor weakness
Oral reflexive movement (Ve1)[†]	Functional communication (C2)[†]	
Vocalization (Ve2)[†]	Functional communication (C2)[†]	
Functional communication (C2)[†]	Eyes open with stimulation (Ar1)[†]	N/A[*]
Functional communication (C2)[†]	Eyes open without stimulation (Ar2)[†]	N/A[*]

[+]Non-observed subscore combinations

[†]p <.001

[*]These combinations are clinically expected and likely to be improbable due to the population included in the sample (mainly VS/MCS patients).

[≠]Awake with preserved vision but unable to open the eyelids