Journal of Neurology (2022) 269:3204–3215
https://doi.org/10.1007/s00415-021-10905-y

ORIGINAL COMMUNICATION

Towards consensus on visual pursuit and visual fixation in patients with disorders of consciousness. A Delphi study

Berno U. H. Overbeek[1,2,3] · Jan C. M. Lavrijsen[1] · Simon van Gaal[4] · Daniel Kondziella[5,6] · Henk J. Eilander[1] · Raymond T. C. M. Koopmans[1,7]

Received: 20 July 2021 / Revised: 10 November 2021 / Accepted: 11 November 2021 / Published online: 10 January 2022

Abstract

Background The aim of this Delphi study was to reach consensus about definition, operationalization and assessment of visual pursuit (VP) and visual fixation (VF).

Methods In a three-round international Delphi study, clinical and research experts on disorders of consciousness indicated their level of agreement on 87 statements using a 5-point Likert scale. Consensus for agreement was defined by a median of 5, an interquartile range (IQR) ≤ 1, and ≥ 80% indicating moderate or strong agreement.

Results Forty-three experts from three continents participated, 32 completed all three rounds. For VP, the consensus statements with the highest levels of agreement were on the term 'pursuit of a visual stimulus', the description 'ability to follow visually in horizontal and/or vertical plane', a duration > 2 s, tracking in horizontal and vertical planes, and a frequency of more than 2 times per assessment. For VF, consensus statements with the highest levels of agreement were on the term 'sustained VF', the description 'sustained fixation in response to a salient stimulus', a duration of > 2 s and a frequency of 2 or more times per assessment. The assessment factors with the highest levels of agreement were personalized stimuli, the use of eye tracking technology, a patient dependent time of assessment, sufficient environmental light, upright posture, and the necessity to exclude ocular/oculomotor problems.

Conclusion This first international Delphi study on VP and VF in patients with disorders of consciousness provides provisional operational definitions and an overview of the most relevant assessment factors.

Keywords Visual pursuit · Visual fixation · Disorders of consciousness · Delphi study

Introduction

Unresponsive wakefulness syndrome (UWS), originally called vegetative state (VS) [1], and minimally conscious state (MCS) [2] are among the worst outcomes in survivors of acquired brain injury. Patients in UWS/VS show no signs of consciousness [3], whereas MCS patients demonstrate reproducible, minimal signs of consciousness [2]. The diagnosis of MCS is based on one or more of the following behaviors: following simple commands, gestural and/or verbal yes/no responses, intelligible verbalization and purposeful behavior [2]. However, the inconsistent presence of these minimal signs of consciousness can lead to misdiagnosis, especially when the systematic use of clinical assessment tools (e.g. the Coma Recovery Scale-Revised; CRS-R [4]) is omitted, as demonstrated by several studies showing that around 40% of patients presumed to be in UWS/VS were found to be at least in MCS [5, 6].

✉ Berno U. H. Overbeek
berno.overbeek@radboudumc.nl

1 Department of Primary and Community Care, Radboud Institute of Health Sciences, Radboud University Medical Center, Nijmegen, The Netherlands

2 Kalorama, Beek Ubbergen, The Netherlands

3 Azora, Terborg, The Netherlands

4 Faculty of Social and Behavioural Sciences, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

5 Department of Neurology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

6 Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

7 Center for Specialized Geriatric Care, Joachim en Anna, Nijmegen, The Netherlands

A correct diagnosis of MCS is important regarding prognosis [7], ability to perceive pain [8], and effect of treatments such as intensive neurorehabilitation [9, 10], pharmacological treatment with amantadine or zolpidem [11, 12], transcranial direct current stimulation (tDCS) [13], or deep-brain stimulation [14]. Moreover, compared to UWS/VS, a diagnosis of MCS may lead to other medical-ethical considerations because of experience of (some) awareness [15–17].

Visual pursuit (VP) and visual fixation (VF) have been recognized as being among the first signs of emerging consciousness and therefore considered indicative of MCS [2]. Incorporation of VP into the diagnostic criteria was motivated [18] by a higher incidence of VP in MCS patients [19] and better prognostic outcomes [20, 21]. For VF, no supporting data were reported for the incorporation into the diagnostic criteria [18]. The importance of these responses for a diagnosis of MCS has been demonstrated in several studies [22–26] investigating the occurrence of behavioral responses on the CRS-R [4]. This scale consists of six subscales and is the most recommended scale in clinical practice and endorsed by recently published guidelines [27–29]. Visual responses were the most frequently observed signs of consciousness [22–26]. A MCS diagnosis based exclusively on the CRS-R visual subscale score was found in 27–55% [6, 22–25].

The influence of VP and VF on the prevalence rate has been demonstrated in a Dutch prevalence study on UWS/VS [6]. In this study, the diagnosis of reported UWS/VS patients was verified with the CRS-R and 15 of 41 patients (37%), who were diagnosed as UWS/VS, were actually in MCS. Among these 15 patients (11 MCS −, 4 MCS +), 8 showed VP and 1 VF. VP was the only observed sign of consciousness in 6 patients, VF was not observed as a single sign of consciousness.

However, despite the importance for the diagnosis and prevalence of MCS, debate still exists about the clinical significance of VP and VF. In 1994, 8 years before the introduction of MCS as a specific level of consciousness, the Multi Society Task Force on Persistent Vegetative State (MSTF) advised to be cautious in diagnosing UWS/VS. VP and VF could be either considered as visual orienting reflexes or indicative of a transition to 'a state of awareness' [3]. We reviewed the literature about this topic over more than 20 years and found no agreed-upon definitions and many different assessment scales with different operational criteria [30]. In the European guideline on the diagnosis of coma and other disorders of consciousness (DoC), probing for voluntary eye movements has been recommended [28]. However, the clinical significance of VP and VF was not discussed extensively in this guideline, nor in the US guideline [28, 29].

Therefore, a Delphi study was conducted with the aim to reach more consensus about definition, operationalization and assessment of VP and VF.

Methods

The Delphi technique was used to investigate the level of consensus on definition, operationalization and assessment items of VP and VF. A Delphi study is a multi-staged survey which aims to achieve consensus on an important issue [31]. The result is an expert opinion about a subject where previously no such opinion existed. In the development of clincial guidelines, the Delphi technique is used as a consensus builiding process to capture expert opinions and experiences [29, 32, 33].

This Delphi study consisted of three rounds, since this approach provides the best balance between obtaining a high level of consensus and maintaining a high response rate [31]. A two or three-round Delphi is in line with most Delphi studies [34]. Prior to the start of the study, it was agreed that a third round would take place if consensus was reached on less than two thirds of the statements.

Based on an extensive literature review [30], four researchers from the Netherlands (BO, JL, HE, RK) initiated the study and developed the protocol and survey. Since international consensus was aimed for, four other researchers from different countries were approached to form an international core group for further review of the protocol and survey. This core group consisted of 3 elderly care physicians (a medical specialty in the Netherlands, formerly known as nursing home medicine [35]), 2 clinical neuropsychologists, a neuroscientist who had developed an assessment scale for patients with DoC, an experienced neurologist in acute and critical care, and a neuroscientist with expertise in human eye movements.

After the core group approved the final version of the survey, an international panel of experts was identified and selected. Participants were recruited by searching the reference list of our integrative review [30], by approaching members of the International Brain Injury Association (IBIA) special interest groups on diagnosis/prognosis and treatment in DoC patients, and by the collaborative network of the core group members. Potential participants were considered as experts if they met one ore more of the following criteria: (1) had published about DoC over the last 10 years, (2) developed assessment scale(s) for DoC patients which include visual functions over the last 10 years, (3) had clinical experience in the assessment and treatment of DoC patients in the acute, post-acute or chronic phases over the last 5 years, (4) were experienced in the neurobiology of eye movements. Expert panel members received an online information letter, gave digital informed consent before starting the first-round survey and received online access to the integrative review [30].

In the survey, definitions and descriptions of both VP and VF that were found in the literature review [30] were presented, as well as the operational criteria direction of tracking, frequency and duration of response. Regarding the assessment factors, statements about assessment conditions, type of stimulus and additional techniques were presented. The statements were categorized in seven domains, according to the study aims: terminology of VP, descriptions of VP, items of operational criteria of VP, terminology of VF, descriptions of VF, items of operational criteria of VF and factors facilitating or confounding the assessment of VP and VF. The survey was sent to the participants electronically using LimeSurvey (www.limesurvey.org) (Supplementary information).

For each statement, the level of agreement was indicated on a five-point Likert scale which is generally used in Delphi studies [31]. The five levels were: strongly disagree (1), moderately disagree (2), neutral (3), moderately agree (4), and strongly agree (5). In addition to these levels the option 'undecided' could be chosen by experts, when they thought they did not have sufficient expertise to indicate their level of (dis)agreement. Apart from indicating the level of agreement, suggestions for additional statements and comments could be written under each domain. The suggestions and comments were used for possible additions and/or revisions of statements in the second round. The first round was open for 3 weeks in June 2019. Before starting this round, the following demographic data were collected: country of residence, profession, experience with treatment of DoC patients, work setting, use of assessment scale(s) and use of additional techniques. The second round was open for 8 weeks over the summer period (July–September 2019). In this round, participants received feedback to their answers so they could see their own results compared to the median scores of the entire expert panel. The third round was open for 3 weeks (October–November 2019) and had the same methodology as the second round.

Data analysis

Measures of central tendency and dispersion are the most frequently used statistics in Delphi Studies [31]. Consensus was calculated by combining median values, interquartile ranges (IQR), and percentage of agreement. The median is a measure of central tendency and the IQR is a measure of variability, based on dividing the scores into quartiles and calculating the difference between the first and third quartile. A low IQR is preferred as this indicates little spread of the scores. For each statement, the median, IQR and percentage of agreement were calculated after each round, using SPSS Statistics version 25.0 (Armonk, NY: IBM Corp).

Consensus for agreement was defined as a median score of 5, an IQR ≤ 1, and $\geq 80\%$ scoring 4 or 5. Consensus for disagreement was defined as a median of 1, an IQR ≤ 1, and $\geq 80\%$ scoring 1 or 2. The consensus percentage of 80% is in line with a systematic review of 98 Delphi studies which demonstrated that the level of consensus was most frequently expressed in percentages with a median threshold for consensus of 75% [34].

Ethical review

The protocol was reviewed by the medical-ethical committee of the Radboud University Medical Center and was not considered to be subject to the Dutch Medical Research Act Involving Human Subjects (1998). According to the medical-ethical committee, further medical-ethical evaluation was not indicated (file number 2019–5225).

Results

Expert panel

Of the 99 invited experts from three continents, 43 (44.8%) agreed to participate. Thirty-two participants (74%) completed all 3 rounds (Fig. 1). Experts worked as medical doctors (specialized in neurology, neurocritical care, neurosurgery, rehabilitation and elderly care medicine) (33%), neuropsychologists (33%), neuroscientists (12%), speech therapists (9%), occupational therapists (9%) and researchers (7%). Forty percent combined a clinical profession with a research function. The experts worked in hospitals, rehabilitation clinics and nursing homes or in a combination of these settings. The mean experience of working with DoC patients was 12.8 years. Assessment scales were used widely and as many as 26 different scales were mentioned, of which the CRS-R was most frequently used. Slightly more than half of the experts used additional techniques, which were mainly used in hospitals (77%), sparsely in rehabilitation clinics (14%) and not at all in nursing homes. The most frequently used technique was electroencephalogram (EEG) (91%) (Table 1).

Consensus development

The first-round survey of the Delphi consisted of 73 statements (Supplementary information). Based on the suggestions and comments of the expert panel after the first round, 14 new statements were added to the second-round survey, and 1 statement was slightly rephrased (Table 2). The number of statements with consensus in the three rounds was 14 (19%), 3 (4.1%) and 23 (32.9%) respectively. In total, consensus was reached on 40/87 (46%) statements. The process

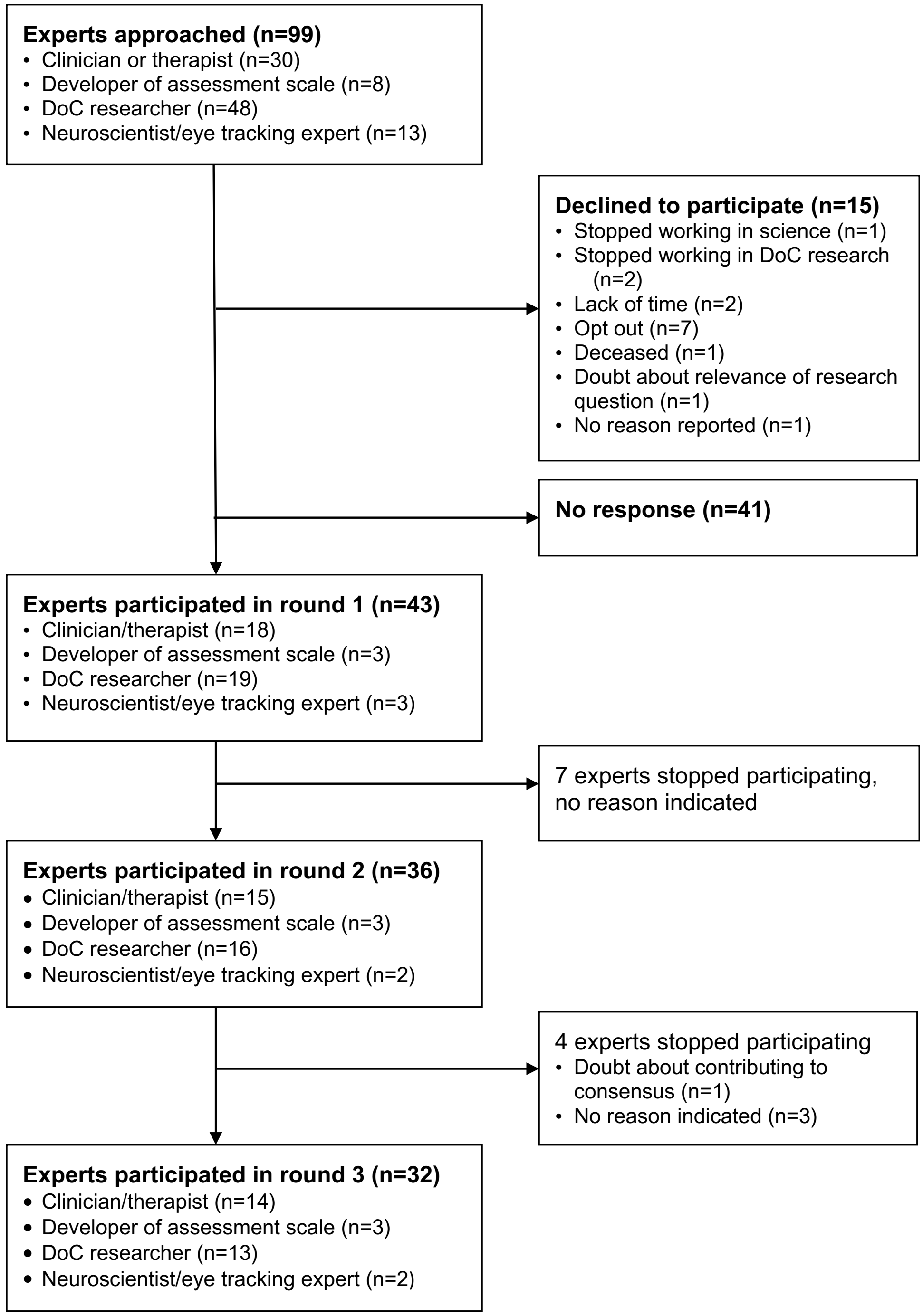

Fig. 1 Flow chart of participation of experts

of consensus development is described in Fig. 2 and statements with consensus on agreement and disagreement are presented in Tables 3 and 4 respectively.

Round 1

For VP, no consensus was reached on terminology and description. Regarding operational criteria, consensus was reached on testing VP in the horizontal plane (92.3%). For VF, no consensus was reached on terminology either, but consensus was reached on the description 'sustained fixation that occurs in relation to a salient stimulus' (83.7%). Also, consensus was reached on the 'frequency of response' (83.3%) as an operational criterion and more specifically on the presence of > 2 responses per assessment (82.4%). Regarding the assessment factors, consensus was reached on the type of stimulus (100%), and in more detail on the use of a personalized object (100%), the presence of a person (97.5%) in particular a relative (97.4%), a photo of a relative (94.7%) and the use of a

Table 1 Characteristics of expert panel and diagnostics used ($n = 43$)

Residence, number (%)	
Europe	33 (77)
North America	7 (16)
Asia	3 (7)
Experience with treatment of DoC patients (years), mean (range)	12.8 (0–35)
Profession, number (%)	
Medical doctor	14 (33)
Neurology[a]	8 (57)
Neurosurgery	2 (14)
Rehabilitation	2 (14)
Elderly care physician	2 (14)
Neuropsychologist	14 (33)
Neuroscientist	5 (12)
Speech therapist	4 (9)
Occupational therapist	3 (7)
Researcher	3 (7)
Clinical profession combined with research function	17 (40)
Professional setting, number (%)	
Hospital	14 (33)
Rehabilitation clinic	12 (28)
Nursing home	4 (9)
Combination of settings	9 (23)
Not reported	4 (9)
Use of assessment scale, number (%)	40 (93)
Most frequently used[b]	
Coma Recovery Scale revised (CRS-R)	36 (90)
Glasgow Outcome Scale (GOS)/Glasgow Outcome Scale Extended (GOSE)	6 (15)
Full Outline of Responsiveness Score (FOUR)	5 (13)
Wessex Head Injury Matrix (WHIM)	5 (13)
Use of additional diagnostic techniques, number (%)	22 (51)
Settings	
Hospital	17 (77)
Rehabilitation clinic	3 (14)
Nursing home	0
Research setting (lab)	2 (9)
Most frequently used techniques	
Electroencephalogram (EEG)	20 (91)
Functional Magnetic Resonance Imaging (fMRI)	12 (55)
Positron Emission Tomography (PET)	11 (50)
Magnetic Resonance Imaging (MRI)	7 (32)
Evoked and event related potentials	9 (18)

[a]2 neurologists specifically reported working as neuro-intensivists

[b]The majority of the experts used a combination of assessment scales; therefore, numbers and percentages exceed 43 (100%)

mirror (90.0%). Consensus was reached on physical settings and circumstances as facilitating assessment factor (93.0%), in particular sufficient environmental light (87.8%). No consensus was reached on a specific assessment time. Consensus was reached on excluding oculomotor (97.7%) and ocular (97.6%) problems as confounding factors.

Round 2

For VP, no consensus was reached on terminology and description. In the comments after round 1, it was suggested that VP should be tested in both horizontal and vertical planes and therefore this was added as a statement, which reached consensus (96.0%). For VF, no consensus

Table 2 Revised and newly added statements in the second round

Terminology of visual pursuit
Ocular following
Purposeful eye movements
Pursuit of a visual stimulus
Visual following
Descriptions of visual pursuit
Ability to follow visually in horizontal *and/or* vertical plane (through right, left, upper and lower visual fields)[a]
Ability to follow a moving mirror or any moving stimulus (object, person) with the eyes only
Slow tracking movements of the eyes to keep a small moving stimulus on the fovea
Smoothly follow a moving object with the entire head and/or eyes
Smooth pursuit eye movements that occur in direct response to moving stimuli in horizontal and/or vertical plane through right, left, upper and/or lower visual fields
Tracking eye movements following objects and/or people
Items for operational criteria of visual pursuit
Visual pursuit in horizontal and/or vertical plane[b]
Terminology of visual fixation
Sustained visual fixation
Factors facilitating or confounding the assessment of VP and VF
Assess patient in different positions like lying, sitting (if possible), standing (if possible e.g. in tilt table, standing frame)
The best moments for assessment of VP and VF are patient dependent
Visual evoked potentials

[a]Minor rephrasing of statement: horizontal and vertical plane replaced by horizontal and/or vertical plane

[b]Experts suggested to both include horizontal and vertical plane as operational criterion, instead of horizontal and vertical plane separately

was reached on terminology. Consensus was reached on duration of the response as operational criterion (97.1%). Concerning the time of assessment, a patient dependent time of assessment was suggested and therefore added as a statement, which reached consensus (96.4%).

Round 3

For VP, consensus was reached on 2 terms, with the highest level of consensus reached on the term 'pursuit of a visual stimulus' (90.3%). Consensus was reached on 5 descriptions, with the highest level of consensus on: 'ability to follow visually in horizontal and/or vertical plane (through right, left, upper and lower visual fields)' (90.6%). Regarding the operational criteria of VP, consensus was reached on the duration (90.3%) and frequency (90.6%) of the response. More specifically, VP should have a duration of > 2 s (80.8%), and a frequency of > 2 VP responses per assessment (83.3%). Rejected statements were the description 'smoothly follow a moving object, with the entire head' (93.3%), and 2 operational criteria: a duration of 1 s (96.3%) and a frequency of 1 response per assessment (93.3%). For VF, consensus was reached on the term 'sustained visual fixation' (87.5%), on a duration of > 2 s (93.3%), and on a frequency of 2 responses per assessment (93.3%). A duration of 1 s (87.1%) and a frequency of 1 response per assessment (90.0%) were rejected. Regarding the assessment, consensus was reached on the following facilitating assessment factors: the use of a photo (100%) in particular of the patient him/herself (96.6%), the use of eye tracking technology (100%), and an upright sitting position (86.7%). At the end of this round, consensus was reached on all domains.

Discussion

This is the first study to reach consensus on definitions, descriptions, operational criteria and relevant assessment factors of VP and VF in DoC patients. The provisional operational definitions of VP and VF are based on a combination of the highest levels of agreement on terminology, description and operational criteria (Boxes 1 and 2). In addition to the provisional operational definition, this study provides an overview of the assessment factors that need to be considered.

We formulated the following provisional operational definition for VP: pursuit of a visual stimulus in which the patient follows visually in horizontal and/or vertical plane (through right, left, upper and lower visual fields) with a duration of more than 2 s, tested in the horizonal and/or vertical planes, and observed more than 2 times in a single assessment session. The newly proposed term 'pursuit of a visual stimulus' is more in line with the functional definition of VP, which is described as allowing clear vision of

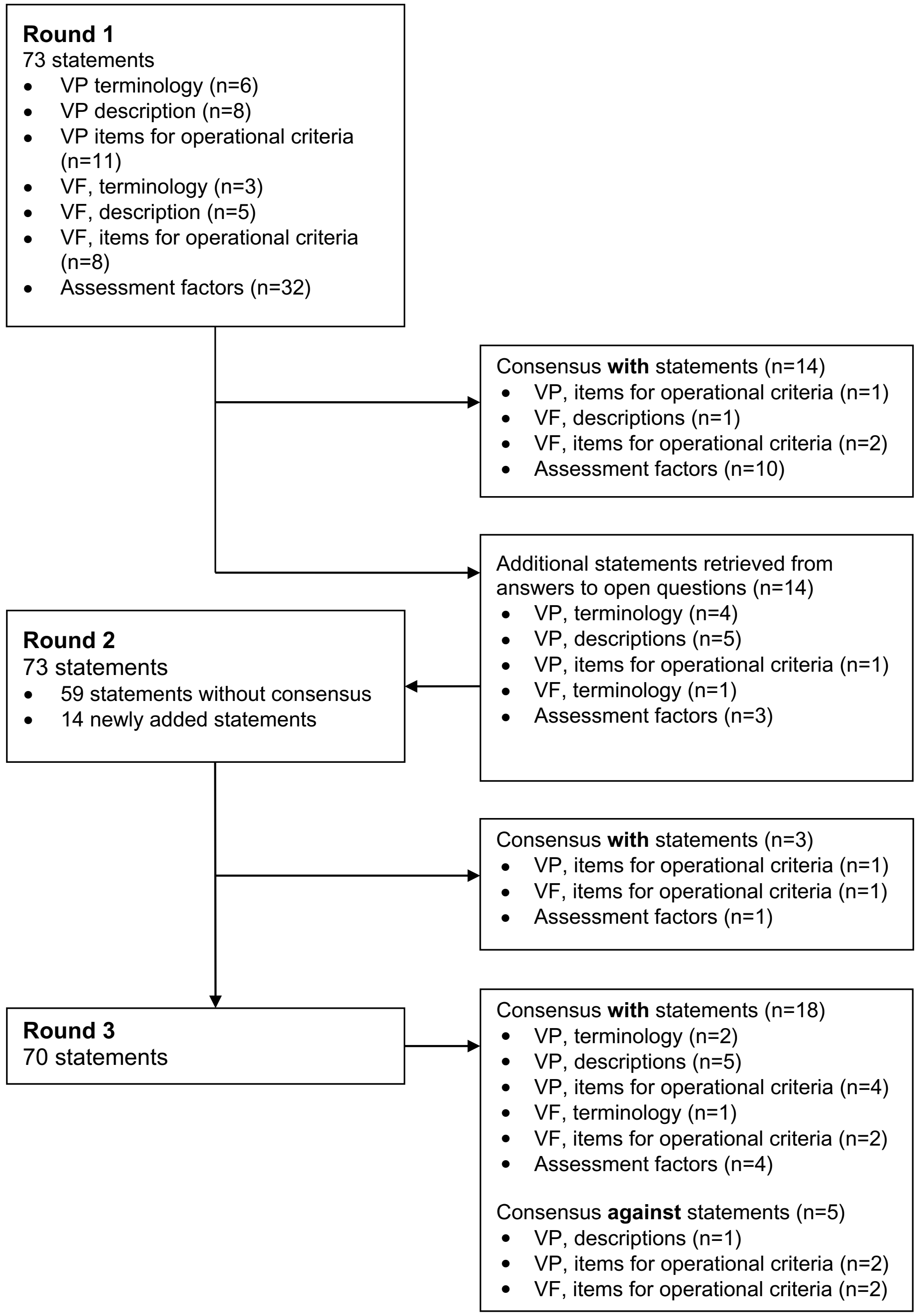

Fig. 2 Consensus development process

the object as it moves within the environment [36]. There was agreement to add the duration of VP as an additional operational criterion. This is useful, since as early as in 1994, 'sustained VP' was reported by the MSTF as a differentiating sign between UWS/VS and a state of emergence out of UWS/VS (i.e. a higher level of consciousness) [3]. Some years later, it was discussed whether UWS/VS patients were capable of visually tracking and it was reported that lack of an accepted operational definition contributed to the ambiguity of the status of VP [37]. The addition of a duration of 2 s increases clarification of the concept of 'sustained VP' and emphasizes that the VP response should be of sufficient long duration, since a short duration (i.e. of 1 s) was rejected.

We formulated the following provisional operational definition of VF: sustained fixation on a salient visual stimulus, in which the patient fixates more than 2 s and observed 2 times or more in a single assessment session. The adjective

Table 3 Statements with consensus on agreement ($n = 35$)

Domain	Agreement (%)
Terminology of visual pursuit	
Pursuit of a visual stimulus	90.3
Pursuit eye movements	87.1
Description of visual pursuit	
Ability to follow visually in horizontal and/or vertical plane (through right, left, upper and lower visual fields)	90.6
Ability to follow a moving mirror or any moving stimulus (object, person) with the eyes only	87.5
Smooth pursuit eye movements that occur in direct response to moving stimuli in horizontal and/or vertical plane through right, left, upper and/or lower visual fields	83.9
Tracking eye movements following objects and/or people	83.3
Pursuit eye movements that occur in direct response to moving stimuli	81.3
Items for operational criteria of visual pursuit	
Duration of response	90.3
More than 2 s	80.8
Horizontal plane	92.3
Horizontal and/or vertical plane	96.0
Frequency of VP within one assessment session	90.6
More than 2 VP responses per assessment	83.3
Terminology of visual fixation	
Sustained visual fixation	87.5
Description of visual fixation	
Sustained fixation that occurs in direct response to a salient stimulus	83.7
Items for operational criteria of visual fixation	
Duration of response	97.1
More than 2 s	93.3
Frequency of response	83.3
2 VF responses per assessment	93.3
More than 2 VF responses per assessment	82.4
Factors facilitating or confounding the assessment of VP and VF	
Facilitating factors	
Type of stimulus	100[a]
Personalized object	100
Photo	100
Photo of the patient him/herself	96.6
Photo of a relative	94.7
Persons	97.5
Relatives	97.4
Mirror	90.0
Use of eye tracking	100
The best moments for assessment of VP and VF are patient dependent	96.4
Physical settings and circumstances	93.0
Sufficient environmental light	87.8
Patient in an upright sitting position	86.7
Confounding factors	
Oculomotor problems (i.e. problems of the eye muscles and/or their innervation)	97.7
Ocular problems (i.e. problems of the eyes themselves)	97.6

[a]Facilitating factors grouped by highest percentages of main items

Table 4 Statements with consensus on disagreement ($n=5$)

Domain	Disagreement (%)
Description of visual pursuit	
Smoothly follow a moving object, with the entire head	93.3
Items for operational criteria of visual pursuit	
1 s	96.3
1 VP response per assessment	93.3
Items for operational criteria of visual fixation	
1 VF response per assessment	90.0
1 s	87.1

'sustained' was added to both terminology and description. This supports the previously mentioned distinction between UWS/VS and MCS by a brief and sustained VF response respectively [2]. Like VP, a VF response should be of sufficient long duration, a short duration (i.e. of 1 s) was rejected. Moreover, VF in the context of DoC reflects refixation of gaze on a moving stimulus, which has to be differentiated from a more general functional definition of VF which is the maintenance of a stationary stimulus on the fovea [36].

To detect VP and VF optimally, the following facilitating factors are important in the assessment: the use of personalized stimuli such as meaningful objects, photos, persons, or a mirror which enables reflection of one's own face, a personally optimized assessment time, sufficient light and upright posture as physical circumstances and eye tracking as potential accessory diagnostic technique. The relevance of using personalized stimuli is in line with a recent review that concluded that the use of personalized stimuli results in detection of more behavioral and electrophysiological responses [38]. Personalization not only accounted for the applied stimuli, but also for the time of assessment. Several standard time intervals were proposed without reaching consensus on any interval. A patient dependent time of assessment was suggested by several experts with the motivation that personal factors such as fatigue, daily individual schedule and individual circadian rhythm can influence the optimal time of assessment. Sufficient light and upright body posture are important physical settings and circumstances. The use of bright light can lead to higher levels of consciousness on the CRS-R [39] and an upright position is associated with better arousal levels and a greater behavioral repertoire [40]. Ocular and/or oculomotor problems were considered as confounding assessment factors and can influence clinical assessment [41], as shown by an absence of visual subscale scores on the CRS-R in about 20% of MCS patients [23]. In this regard, screening on ocular/oculomotor problems is necessary. Eye tracking was considered a useful accessory diagnostic technique, which has been shown to be helpful in differentiating between UWS/VS and MCS [42]. However, eye tracking has not been fully developed yet in the study of disorders of consciousness, and the measurement of eye movements in DoC patients needs further research [43].

Box 1 Highest percentages of agreement for visual pursuit and visual fixation

Domains	Visual pursuit Highest level of agreement	Visual fixation Highest level of agreement
Terminology	Pursuit of a visual stimulus	Sustained visual fixation
Description	Ability to follow visually in horizontal and/or vertical plane (through right, left, upper and lower visual fields)	Sustained fixation that occurs in direct response to a salient stimulus
Operational criteria	Duration > 2 s	Duration > 2 s
	Testing in horizontal and/or vertical plane	
	> 2 responses per assessment	≥ 2 responses per assessment

Box 2 Highest percentages of agreement on assessment factors

Main assessment factors with highest levels of agreement	Detailed assessment factors with highest level of agreement
Type of stimulus	Objects with a personal meaning
	Photos of patient him/herself or relatives
	Presence of familiar persons at the assessment session, especially relatives
	Mirror
Eye tracking technology	Use eye tracking technology
Time of assessment	Best time of assessment is patient dependent
Setting and circumstances	Provide sufficient environmental light
	Assess patient in an upright position
Confounding factors	Take oculomotor problems into account
	Take ocular problems into account

Strengths and limitations

The strength of this study is that a highly experienced panel of experts, from 3 continents, succeeded on reaching consensus on definitions, descriptions and operational criteria of VP and VF for the first time. This consensus development made the proposition of operational definitions of VP and VF possible. However, there are some limitations to address. First, despite the international character of the expert panel, there was an underrepresentation of neuroscientists with expertise on eye movements because only 23% of the recruited experts agreed to participate, compared to 40–50% of the other expert categories. Therefore, valuable input and new perspectives from this field may have been missed. However, since these experts were mainly from outside the DoC research field, the experts who participated are a good representation of current clinical practice and research field on DoC. Second, according to the Delphi methodology, experts indicated their level of (dis)agreement for each statement but did not have the possibility to indicate an order of preference. This could have been helpful in reaching more consensus, since choosing the most appropriate statement was sometimes difficult because of minor differences in phrasing and/or content. On the other hand, indicating agreement on separate statements led to consensus on all domains. Moreover, the consensus statements are strongly supported since consensus requirements were set high.

Future directions

The proposed operational definitions from this Delphi study provide a basis for reaching international consensus on the diagnostic assessment of VP and VF. To achieve this, two further steps are necessary. Firstly, consensus is needed on the proposed definitions of VP and VF and secondly, further research is needed into reliability and validity of the established operational definitions.

Based on the unanimous consensus that eye tracking is a facilitating factor for detecting VP and VF, we recommend further research into the use of eye tracking, a promising technique that might be useful to quantify VP and VF assessments in practice.

From a broader perspective on this subject, more fundamental research is required into the neural basis of VP and VF and the relation to consciousness. The association between visual awareness and eye movements is subject of debate and it has been shown that the eyes react to moving objects even when the object is presented below the threshold of awareness [44]. The need for further research into the neural basis of VP and VF and the relation to consciousness also stresses the need for a gold standard measure for determining the level of consciousness. Until now, clinical evaluation is considered as the gold standard for assessing patients with DoC [45]. However, this approach leads to high misdiagnosis rates because of limitations on patient, examiner and environmental level [45]. The use of imaging and electrophysiological techniques has been extensively investigated and their use is recommended, especially when command following is not observed clinically [28]. However, there are challenges regarding availability (i.e. these techniques are not routinely available) and uncertainty about the clinical and prognostic significance of the use of these techniques which can raise ethical dilemmas [46].

The provisional operational definitions for VP and VF resulting from this Delphi study largely correspond to the representation of these responses in the CRS-R administration and scoring guideline [47]. In each CRS-R subscale, the name of the response, assessment methods and scoring criteria are mentioned. These items are also incorporated in our operational definitions and like the CRS-R, our operational definition contain clear criteria for establishing the presence of VP and/or VF. However, the operational definitions proposed in this study provide a more comprehensive representation of: (1) what the response entails, and (2) what exactly a patient is capable of when demonstrating pursuit of a visual stimulus and/or sustained visual fixation. Concerning the assessment, a more extensive approach is recommended, in which accurate assessment of individually observed behavior is added to standardized methods such as the CRS-R. In a recently published review on behavioral assesment in DoC patients, standardized and individual appraoces have been discussed and a combination of these two approaches was recommended [48].

In conclusion, this Delphi study provides provisional operational definitions of VP and VF and an overview of the most relevant assessment factors. The use of unambiguous operational definitions in the assessment procedure and an adaptation of the assessment procedure to individual circumstances is crucial for improvement of diagnostic accuracy and establishing reliable prevalence rates of MCS and other DoC, which is essential for allocating appropriate treatment to DoC patients and determining treatment services.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00415-021-10905-y.

Acknowledgements We thank all the experts for participating in this study. We thank the IBIA special interest group on diagnosis/prognosis of DoC for their help in recuitment of experts and discussing the study. We thank 'Stichting St. Joannes de Deo' and 'Het Coma-Hospitium' for funding this study. We thank Anita Oude Bos, research assistant of the Radboud University Medical Center for developing and processing the online survey and calculating the results. We thank Camille Chatelle PhD and Theresa Bender-Pape PhD for reviewing the survey.

Author contributions BO contributed to the design of the study protocol, acquisition and analysis of the data and drafted the manuscript.

JL contributed to the design of the study protocol, analysis of the data and drafted a significant proportion of the manuscript. SvG contributed to the review of the study protocol, interpretatation of the data and revision of the manuscript. DK contributed to the review of the study protocol, interpretatation of the data and revision of the manuscript. HE contributed to the design of the study protocol, analysis of the data and drafted a significant proportion of the manuscript. RK contributed to the design of the study protocol, analysis of the data and drafted a significant proportion of the manuscript.

Funding This study was funded by 'Het Coma-Hospitium' and 'Stichting St. Joannes de Deo'.

Data availability Anonymized data will be shared upon reasonable request.

Declarations

Conflicts of interest The authors declare that they have no conflict of interest.

Ethical approval The study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendsments or comparable ethical standards.

Informed consent Informed consent was obtained from the participating experts prior to the start of the study.

References

1. Laureys S, Celesia G, Cohadon F, Lavrijsen J, León-Carrión J, Sannita W, Sazbon L, Schmutzhard E, von Wild K, Zeman A, Dolce G (2010) Unresponsive wakefulness syndrome: a new name for the vegetative state or apallic syndrome. BMC Med 8:68. https://doi.org/10.1186/1741-7015-8-68
2. Giacino JT, Ashwal S, Childs N, Cranford R, Jennett B, Katz DI, Kelly JP, Rosenberg JH, Whyte J, Zafonte RD, Zasler ND (2002) The minimally conscious state: definition and diagnostic criteria. Neurology 58:349–353. https://doi.org/10.1212/wnl.58.3.349
3. Multi-Society Task Force on PVS (1994) Medical aspects of the persistent vegetative state (1). N Engl J Med 330:1499–1508. https://doi.org/10.1056/NEJM199405263302107
4. Giacino JT, Kalmar K, Whyte J (2004) The JFK Coma Recovery Scale-Revised: measurement characteristics and diagnostic utility. Arch Phys Med Rehabil 85:2020–2029. https://doi.org/10.1016/j.apmr.2004.02.033
5. Schnakers C, Vanhaudenhuyse A, Giacino J, Ventura M, Boly M, Majerus S, Moonen G, Laureys S (2009) Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment. BMC Neurol 9:35. https://doi.org/10.1186/1471-2377-9-35
6. van Erp WS, Lavrijsen JC, Vos PE, Bor H, Laureys S, Koopmans RT (2015) The vegetative state: prevalence, misdiagnosis, and treatment limitations. J Am Med Dir Assoc 16:85 e89-85 e14. https://doi.org/10.1016/j.jamda.2014.10.014
7. Luaute J, Maucort-Boulch D, Tell L, Quelard F, Sarraf T, Iwaz J, Boisson D, Fischer C (2010) Long-term outcomes of chronic minimally conscious and vegetative states. Neurology 75:246–252. https://doi.org/10.1212/WNL.0b013e3181e8e8df
8. Boly M, Faymonville ME, Schnakers C, Peigneux P, Lambermont B, Phillips C, Lancellotti P, Luxen A, Lamy M, Moonen G, Maquet P, Laureys S (2008) Perception of pain in the minimally conscious state with PET activation: an observational study. Lancet Neurol 7:1013–1020. https://doi.org/10.1016/S1474-4422(08)70219-9
9. Eilander HJ, Wijnen VJ, Scheirs JG, de Kort PL, Prevo AJ (2005) Children and young adults in a prolonged unconscious state due to severe brain injury: outcome after an early intensive neurorehabilitation programme. Brain Inj 19:425–436. https://doi.org/10.1080/02699050400025299
10. Seel RT, Douglas J, Dennison AC, Heaner S, Farris K, Rogers C (2013) Specialized early treatment for persons with disorders of consciousness: program components and outcomes. Arch Phys Med Rehabil 94:1908–1923. https://doi.org/10.1016/j.apmr.2012.11.052
11. Giacino JT, Whyte J, Bagiella E, Kalmar K, Childs N, Khademi A, Eifert B, Long D, Katz DI, Cho S, Yablon SA, Luther M, Hammond FM, Nordenbo A, Novak P, Mercer W, Maurer-Karattup P, Sherer M (2012) Placebo-controlled trial of amantadine for severe traumatic brain injury. N Engl J Med 366:819–826. https://doi.org/10.1056/NEJMoa1102609
12. Singh R, McDonald C, Dawson K, Lewis S, Pringle AM, Smith S, Pentland B (2008) Zolpidem in a minimally conscious state. Brain Inj 22:103–106. https://doi.org/10.1080/02699050701829704
13. Thibaut A, Bruno MA, Ledoux D, Demertzi A, Laureys S (2014) tDCS in patients with disorders of consciousness: sham-controlled randomized double-blind study. Neurology 82:1112–1118. https://doi.org/10.1212/WNL.0000000000000260
14. Schiff ND, Giacino JT, Kalmar K, Victor JD, Baker K, Gerber M, Fritz B, Eisenberg B, Biondi T, O'Connor J, Kobylarz EJ, Farris S, Machado A, McCagg C, Plum F, Fins JJ, Rezai AR (2007) Behavioural improvements with thalamic stimulation after severe traumatic brain injury. Nature 448:600–603. https://doi.org/10.1038/nature06041
15. Wilkinson D, Savulescu J (2013) Is it better to be minimally conscious than vegetative? J Med Ethics 39:557–558. https://doi.org/10.1136/medethics-2012-100954
16. Cranford RE (1998) The vegetative and minimally conscious states: ethical implications. Geriatrics 53(Suppl 1):S70-73
17. Fins JJ, Bernat JL (2018) Ethical, palliative, and policy considerations in disorders of consciousness. Neurology 91:471–475. https://doi.org/10.1212/WNL.0000000000005927
18. Giacino JT, Kalmar K (2005) Diagnostic and prognostic guidelines for the vegetative and minimally conscious states. Neuropsychol Rehabil 15:166–174. https://doi.org/10.1080/09602010443000498
19. Giacino JT, Kalmar K (1997) The vegetative and minimally conscious states: a comparison of clinical features and functional outcome. J Head Trauma Rehabil 12:36–51. https://doi.org/10.1097/00001199-199708000-00005
20. Shiel A, Horn SA, Wilson BA, Watson MJ, Campbell MJ, McLellan DL (2000) The Wessex Head Injury Matrix (WHIM) main scale: a preliminary report on a scale to assess and monitor patient recovery after severe head injury. Clin Rehabil 14:408–416. https://doi.org/10.1191/0269215500cr326oa
21. Ansell BJ, Keenan JE (1989) The Western Neuro Sensory Stimulation Profile: a tool for assessing slow-to-recover head-injured patients. Arch Phys Med Rehabil 70:104–108
22. Noe E, Olaya J, Navarro MD, Noguera P, Colomer C, Garcia-Panach J, Rivero S, Moliner B, Ferri J (2012) Behavioral recovery in disorders of consciousness: a prospective study with the Spanish version of the Coma Recovery Scale-Revised. Arch Phys Med Rehabil 93(428–433):e412. https://doi.org/10.1016/j.apmr.2011.08.048
23. Estraneo A, Moretta P, Cardinale V, De Tanti A, Gatta G, Giacino JT, Trojano L (2015) A multicentre study of intentional behavioural responses measured using the Coma Recovery

Scale-Revised in patients with minimally conscious state. Clin Rehabil 29:803–808. https://doi.org/10.1177/0269215514556002

24. Wannez S, Gosseries O, Azzolini D, Martial C, Cassol H, Aubinet C, Annen J, Martens G, Bodart O, Heine L, Charland-Verville V, Thibaut A, Chatelle C, Vanhaudenhuyse A, Demertzi A, Schnakers C, Donneau AF, Laureys S (2018) Prevalence of coma-recovery scale-revised signs of consciousness in patients in minimally conscious state. Neuropsychol Rehabil 28:1350–1359. https://doi.org/10.1080/09602011.2017.1310656
25. Bagnato S, Boccagni C, Sant'Angelo A, Fingelkurts AA, Fingelkurts AA, Galardi G (2017) Longitudinal assessment of clinical signs of recovery in patients with unresponsive wakefulness syndrome after traumatic or nontraumatic brain injury. J Neurotrauma 34:535–539. https://doi.org/10.1089/neu.2016.4418
26. Martens G, Bodien Y, Sheau K, Christoforou A, Giacino JT (2020) Which behaviours are first to emerge during recovery of consciousness after severe brain injury? Ann Phys Rehabil Med 63:263–269. https://doi.org/10.1016/j.rehab.2019.10.004
27. American Congress of Rehabilitation Medicine BI-ISIGDoCTF, Seel RT, Sherer M, Whyte J, Katz DI, Giacino JT, Rosenbaum AM, Hammond FM, Kalmar K, Pape TL, Zafonte R, Biester RC, Kaelin D, Kean J, Zasler N (2010) Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research. Arch Phys Med Rehabil 91:1795–1813. https://doi.org/10.1016/j.apmr.2010.07.218
28. Kondziella D, Bender A, Diserens K, van Erp W, Estraneo A, Formisano R, Laureys S, Naccache L, Ozturk S, Rohaut B, Sitt JD, Stender J, Tiainen M, Rossetti AO, Gosseries O, Chatelle C, Ean Panel on Coma DoC (2020) European Academy of Neurology guideline on the diagnosis of coma and other disorders of consciousness. Eur J Neurol 27:741–756. https://doi.org/10.1111/ene.14151
29. Giacino JT, Katz DI, Schiff ND, Whyte J, Ashman EJ, Ashwal S, Barbano R, Hammond FM, Laureys S, Ling GSF, Nakase-Richardson R, Seel RT, Yablon S, Getchius TSD, Gronseth GS, Armstrong MJ (2018) Practice guideline update recommendations summary: Disorders of consciousness: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology; the American Congress of Rehabilitation Medicine; and the National Institute on Disability, Independent Living, and Rehabilitation Research. Neurology 91:450–460. https://doi.org/10.1212/WNL.0000000000005926
30. Overbeek BUH, Eilander HJ, Lavrijsen JCM, Koopmans R (2018) Are visual functions diagnostic signs of the minimally conscious state? an integrative review. J Neurol 265:1957–1975. https://doi.org/10.1007/s00415-018-8788-9
31. Keeney S, Hasson F, McKenna HP (2011) The Delphi technique in nursing and health research. Wiley-Blackwell, Oxford
32. Junger S, Payne SA, Brine J, Radbruch L, Brearley SG (2017) Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med 31:684–706. https://doi.org/10.1177/0269216317690685
33. Giacino JT, Whyte J, Nakase-Richardson R, Katz DI, Arciniegas DB, Blum S, Day K, Greenwald BD, Hammond FM, Pape TB, Rosenbaum A, Seel RT, Weintraub A, Yablon S, Zafonte RD, Zasler N (2020) Minimum competency recommendations for programs that provide rehabilitation services for persons with disorders of consciousness: a position statement of the american congress of rehabilitation medicine and the national institute on disability, independent living and rehabilitation research traumatic brain injury model systems. Arch Phys Med Rehabil 101:1072–1089. https://doi.org/10.1016/j.apmr.2020.01.013
34. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, Wales PW (2014) Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies. J Clin Epidemiol 67:401–409. https://doi.org/10.1016/j.jclinepi.2013.12.002
35. Koopmans RT, Lavrijsen JC, Hoek F (2013) Concrete steps toward academic medicine in long term care. J Am Med Dir Assoc 14:781–783. https://doi.org/10.1016/j.jamda.2013.08.004
36. Leigh RJ, Zee DS (2015) The neurology of eye movements. Oxford University Press, New York
37. Giacino JT (1997) Disorders of consciousness: differential diagnosis and neuropathologic features. Semin Neurol 17:105–111. https://doi.org/10.1055/s-2008-1040919
38. Magliacano A, De Bellis F, Galvao-Carmona A, Estraneo A, Trojano L (2019) Can salient stimuli enhance responses in disorders of consciousness? A systematic review. Curr Neurol Neurosci Rep 19:98. https://doi.org/10.1007/s11910-019-1018-8
39. Blume C, Lechinger J, Santhi N, del Giudice R, Gnjezda MT, Pichler G, Scarpatetti M, Donis J, Michitsch G, Schabus M (2017) Significance of circadian rhythms in severely brain-injured patients: a clue to consciousness? Neurology 88:1933–1941. https://doi.org/10.1212/WNL.0000000000003942
40. Wilson BA, Dhamapurkar S, Tunnard C, Watson P, Florschutz G (2013) The effect of positioning on the level of arousal and awareness in patients in the vegetative state or the minimally conscious state: a replication and extension of a previous finding. Brain Impairment 14:475–479. https://doi.org/10.1017/BrImp.2013.34
41. Chatelle C, Bodien YG, Carlowicz C, Wannez S, Charland-Verville V, Gosseries O, Laureys S, Seel RT, Giacino JT (2016) Detection and interpretation of impossible and improbable coma recovery scale-revised scores. Arch Phys Med Rehabil 97(1295–1300):e1294. https://doi.org/10.1016/j.apmr.2016.02.009
42. Trojano L, Moretta P, Masotta O, Loreto V, Estraneo A (2018) Visual pursuit of one's own face in disorders of consciousness: a quantitative analysis. Brain Inj 32:1549–1555. https://doi.org/10.1080/02699052.2018.1504117
43. Ting WK, Perez Velazquez JL, Cusimano MD (2014) Eye movement measurement in diagnostic assessment of disorders of consciousness. Front Neurol 5:137. https://doi.org/10.3389/fneur.2014.00137
44. Spering M, Carrasco M (2015) Acting without seeing: eye movements reveal visual processing without awareness. Trends Neurosci 38:247–258. https://doi.org/10.1016/j.tins.2015.02.002
45. Schnakers C (2020) Update on diagnosis in disorders of consciousness. Expert Rev Neurother 20:997–1004. https://doi.org/10.1080/14737175.2020.1796641
46. Royal College of Physicians (2020) Prolonged disorders of consciousness following sudden onset brain injury: National clinical guidelines. RCP, London
47. Giacino JT, Kalmar K (2004) Coma Recovery Scale-Revised. Administration and Scoring Guidelines. In: Injuries CfH (ed), Edison, New Jersey. http://www.tbims.org/combi/crs/CRS%20Syllabus.pdf
48. Fitzpatrick-DeSalme E, Long A, Patel F, Whyte J (2021) Behavioral assessment of patients with disorders of consciousness. J Clin Neurophysiol. https://doi.org/10.1097/WNP.0000000000000666