J Rehabil Med 2021; 53: jrm0000X

ORIGINAL REPORT

A DELPHI PROCEDURE ON REHABILITATION OUTCOME FOR PATIENTS WITH MODERATE TO SEVERE TRAUMATIC BRAIN INJURY; FIRST PHASE OF THE NEUROTRAUMATOLOGY QUALITY REGISTRY (NET-QURE)

Lianne D. PEPPEL, MSc[1,2], Majanka H. HEIJENBROK-KAL, PhD[1,2], Thomas A. VAN ESSEN, MD[3,4], Godard C. W. DE RUITER, MD, PhD[3], Wilco C. PEUL, MD, PhD[3] and Gerard M. RIBBERS, MD, PhD[1,2]

From the [1]Rijndam Rehabilitation, Rotterdam; [2]Department of Rehabilitation Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam; [3]University Neurosurgical Center Holland, Leiden University Medical Center, Leiden; Haaglanden Medical Center and Haga Teaching Hospital, The Hague; and [4]Center for Medical Decision Sciences, Department of Public Health, Erasmus MC – University Medical Center Rotterdam, Rotterdam, The Netherlands

Objective: **To select a set of rehabilitation outcome instruments for a national Neurotrauma Quality Registry (Net-QuRe) among professionals involved in the care of patients with traumatic brain injury.**

Design: **A 3-round online Delphi procedure.**

Subjects: **Eighty professionals from multiple disciplines working in 1 of the 8 participating rehabilitation centres were invited to participate. The response rate varied from 70% to 76% per round.**

Methods: **For the Delphi procedure, multiple outcome categories were defined based on the International Classification of Functioning, Disability and Health (ICF) with concomitant measurement instruments. For each category we strived for consensus on one instrument of at least 75%.**

Results: **After the first round, consensus was reached for the category subjective cognitive functioning. After the second round for quality of life, pain, general functioning, anxiety and depression, general psychological functioning, communication (impairment), and personal factors. Finally, after the third round, consensus was reached for activities of daily living, participation, self-awareness, and aphasia. No consensus was reached for the categories motor function, cognitive function, comorbidity, fatigue, and employment status.**

Conclusion: **Consensus was reached in 12 out of 17 outcome categories. A Delphi procedure seems to be a feasible method to collectively select measurement instruments for a multicentre study.**

Key words: Delphi procedure; outcome measures; traumatic brain injury.

Accept 28 Oct, 2021; Epub ahead of print Nov 10, 2021

J Rehabil Med 2021; 53: jrm002XX

Correspondence address: Lianne D. Peppel, Rijndam Rehabilitation, Rotterdam, The Netherlands. E-mail: lpeppel@rijndam.nl

LAY ABSTRACT

The aim of this study was to select a set of rehabilitation outcome instruments for a national Neurotrauma Quality Registry (Net-QuRe) among professionals involved in the care of patients with traumatic brain injury. Eighty professionals from multiple disciplines working in 1 of the 8 participating rehabilitation centres were invited to participate in a 3-round online Delphi procedure. Consensus was reached for the categories subjective cognitive functioning, quality of life, pain, general functioning, anxiety and depression, general psychological functioning, communication (impairment), personal factors, activities of daily living, participation, self-awareness, and aphasia. No consensus was reached for the categories motor function, cognitive function, comorbidity, fatigue, and employment status. A Delphi procedure seems to be a feasible method to collectively select measurement instruments for a multicentre study.

Worldwide, 13 million people sustain moderate to severe traumatic brain injury (TBI) annually. Most patients will have life-long disabilities after surviving moderate to severe traumatic brain injury (1). These disabilities vary in nature and severity. Domains that are often affected are cognition, self-awareness, motor function, language and communication. Moreover, patients may encounter problems in activities of daily life, may have fatigue, depression, anxiety or pain, which in turn may influence quality of life, social participation, or general and psychological functioning (2–4). Due to the variety of disabilities, care pathways for patients with TBI differ from one patient to another. To gain insight into the chain of care and outcomes on several domains after moderate to severe TBI, the Neurotraumatology Quality Registry (Net-QuRe) was initiated, in which patients with moderate to severe TBI are followed from hospital admission through the healthcare chain up to 2 years post-injury as an add-on study of Center-TBI (5). The aim of Net-QuRe is to evaluate treatment strategies and to optimize the quality of the total healthcare chain for patients with moderate to severe TBI. In contrast to the USA, in the Netherlands a national initiative like The Traumatic Brain Injury Model Systems (TBIMS) (6), whereby patients with TBI are followed in a standardized way to examine outcome and the course of recovery does not yet exist. Centre-specific protocols targeting outcome measurement are scarce and differ among Dutch rehabilitation centres.

jrm.v53.760

doi: 10.2340/jrm.v53.760

To measure outcome in Net-QuRe, suitable measurement instruments needed to be identified from the numerous instruments available from the literature (7–9). For example, in systematic reviews, Tate et al. (2013) identified 728 behavioural assessment instruments for patients with TBI (7). Van Heugten et al. (2019) identified 381 neuropsychological instruments (8), and Laxe et al. (2012) identified 283 unique instruments in 193 papers on TBI (9). These instruments can be categorized according to the International Classification of Functioning, Disability and Health (ICF) (9, 10). The ICF provides a standard framework for the description of functioning and disability for health conditions. Its components are classified in body structures and functions, activities, participation, personal, and external factors.

Several initiatives have focused on defining core sets for TBI research. In 2010, the TBI Outcomes Working group of the National Institute of Neurological Disorders and Stroke (NINDS) selected a common set of outcome measures, consisting of a core, a supplemental, and an emerging part for TBI (11, 12). However, these core sets include instruments, such as the Functional Independence Measure, that are not common practice outside the USA. More recently, Honan et al. (2019) recommended 56 instruments out of 115 instruments in psychosocial research targeting moderate to severe TBI (13). Together these articles provided a complete overview of existing outcome measures. Selecting the best set from this abundance of outcome measures is a challenge.

To identify a set of outcome measures on ICF categories for patients with moderate to severe TBI to be used in Net-QuRe a Delphi procedure was performed, an iterative process based on expert opinion suitable to reach consensus (14). It is a flexible procedure concerning the number of rounds, consensus level and the number of participants (15, 16). By using expert opinion, experience of professionals, and existing knowledge for selecting a customized set of outcome measures we intended to create support for using the set of measurement instruments in both clinical care and research, and to minimize the chance of missing data in Net-QuRe. Choices made concerning the Delphi procedure are explained and questioned in this article.

METHODS

The ethics procedure for this study was waived by the medical ethics committee of Erasmus MC, University Medical Center Rotterdam (MEC-2015-395).

Design

Between 10 July and 30 September 2015 an online Delphi procedure was performed among experts from 8 rehabilitation centres in the Netherlands to select outcome measures for Net-QuRe (Netherlands Trial Register NL5761). In Net-QuRe detailed information concerning treatment and clinical data from the acute phase are collected according to Center-TBI. This Delphi procedure will focus solely on follow-up and rehabilitation outcome. Guidelines from the literature were used to make choices concerning the number of rounds, consensus level, and the number of participants (14, 16, 17). In this Delphi procedure questionnaires are completed anonymously, answers from each round are presented in the next round, and participants review their initial response based on the group response of the previous round (14). The procedure consisted of 3 Delphi rounds among professionals with expertise in treatment of patients with TBI. The procedure ended after 3 rounds or when a consensus level of 75% was reached. First-choice instrument references are given in Appendix I. The results from the Delphi procedure were used by the research committee to decide on the final outcome set for Net-QuRe.

Participants

Eight members of the Dutch Special Interest Group on TBI, which represents all Dutch rehabilitation centres, were asked to provide a list of 2 physiatrists, 2 psychologists, 2 physiotherapists, 2 occupational therapists, 1 or 2 speech and language therapists, 1 or 2 social workers, 1 or 2 nurse/nurse practitioners and, if available, an expert with specific knowledge of 1 of the domains or knowledge of TBI. Two elderly care physicians working in skilled nursing homes and not being a member of the Special Interest Group were invited to participate because of their knowledge of acquired brain injury and outcome measures. All experts were instructed about the Delphi procedure via an information letter.

Development of the online Delphi questionnaires

Based on the ICF model, interviews with physiatrists, and expert opinion, the following outcome categories were selected for the Delphi procedure: objective cognitive functioning, subjective cognitive functioning, aphasia, (cognitive) communication, self-awareness, general psychological functioning, anxiety and depression, pain, fatigue, physical functioning, activities and participation, activities of daily living, employment status, personal factors, general functioning, quality of life, and comorbidity.

Subsequently, a literature search was performed for each category to identify potential measurement instruments for patients with moderate to severe TBI. For this study, afore-mentioned review studies and core sets were used as a starting point for our overview of available instruments (7, 9, 11, 12). From there, extra instruments were added by screening references lists of relevant articles, and by checking at a website containing over 400 measurement instruments used in healthcare (18). All measurement instruments for which a translation into the Dutch language was available were included in the first round of the Delphi procedure. In the first round, instruments that were not listed could be added to the list by the experts.

Questionnaires were developed using an online survey tool named Limesurvey (15). With this tool it is possible to keep track

of responses without showing the answers of a participant. Thus anonymity was guaranteed. Participants were invited by email and could enter the online questionnaire by using a personal token received by email. This token was sent out automatically by the system. For an optimal response, reminders were sent by the system during the procedure. In Appendix II examples of questions asked and feedback given to participants is provided.

Delphi rounds

Each round, participants answered questions about their profession, institution, and the years of experience they had with patients with TBI. Experts could participate in each round irrespective of their contribution to the previous round.

Round 1. For each Delphi category a list with potential measurement instruments was presented. Participants were asked if they were familiar with the instruments. In addition, extra instruments could be added by participants if they were missing in their opinion. By doing so, the list of outcome measures became complete. The main task for each participant for each Delphi category was to select a first- and second-choice instrument for each Delphi category, with the instruction to take into account psychometric quality, sensitivity to measure change, time efficiency, and suitability for patients with moderate to severe TBI. Participants could select instruments from the list or from the new instruments they added to the list. One of the standard answering options was "I am not the right person to answer questions regarding this category". In this case, further questions regarding the concerning Delphi category were skipped automatically. Finally, there was the option to post a comment.

The procedure for cognitive and physical functioning was different from the other Delphi categories. In addition to a screening test, it is almost impossible to select 1 instrument that measures cognitive functioning extensively. The same holds for motor functioning. Therefore, instead of selecting a first- and second-choice cognitive- and physical functioning instrument, participants were asked to select a combination of multiple instruments. The physical functioning combination had to contain 1–4 physical functioning instruments. For the cognitive functioning category, participants were asked which test battery for people with TBI, with a maximum duration of 30 min, they would select in addition to the Montreal Cognitive Assessment (MoCA) (19), which is a general screening instrument.

Round 2. A detailed table with the results of the first round was given at the start of the second round for each Delphi category. If consensus was reached after the first round, this particular category was excluded from the second round. Similar to the first round, participants were asked to select a first- and second-choice instrument. This time the list with instruments was updated based on the results from the first round. In the second round it was not possible to add instruments anymore (14). For the Delphi categories of cognitive- and physical functioning, all combinations given in the first round, formed the list of combinations from which 2 preferred combinations could be chosen in the second round. Again, it was not possible to add new combinations at this stage.

Round 3. In this round participants were asked to select only their first-choice instrument. Furthermore, the third round was comparable to the second round. Categories upon which consensus was reached after the second round were excluded. Instruments that were not chosen by at least 1 expert as first- or second-choice instrument in the first or second round were excluded from the list for the third round.

Statistical analysis

To summarize responses and to characterize the group of experts, descriptive statistics were performed using the Statistical Package for the Social Sciences (SPSS) statistics, version 23.

RESULTS

Participants

One representative of each of the 8 rehabilitation centres, provided a list of 12–15 professionals from different disciplines to be invited for the Delphi procedure. Eighty experts were invited, of whom 72 responded in at least 1 round (overall response rate 90%) and 40 experts (50%) participated in all 3 rounds. In the first round, 76% ($n=61$) of the experts participated, 70% ($n=56$) in the second round, and 74% ($n=9$) in the third round. The mean number of years of experience in the field of acquired brain injury was 14.4 (SD 7.4) years. Details about the professional background of participants are shown in Table I.

First round

In total, 99 instruments were presented to the experts in the first round of the Delphi procedure, divided into 17 categories. Consensus was reached by 80% of participants on using the Cognitive Failure Questionnaire (CFQ) in the category subjective cognitive functioning in the first round. Three out of the 99 (3%) instruments were not familiar to any participant. Another 6 instruments were not selected as first- or second-choice and were also removed. Ninety of 99 (91%) original instruments were selected for the second round. In addition, 30 new instruments were added to the list and 13 unique combinations of tests were created for the

Table I. Number of participants per discipline and per round

	Invited	Round 1	Round 2	Round 3
Discipline	*n* (%)	*n* (%)	*n* (%)	*n* (%)
Physiatrist	15	11 (18.0)	10 (17.9)	7 (11.9)
Psychologist	13	7 (11.5)	8 (14.3)	8 (13.6)
Physiotherapist	12	10 (16.4)	9 (16.1)	11 (18.6)
Occupational therapist	12	11 (18.0)	10 (17.9)	11 (18.6)
Speech and language therapist	9	7 (11.5)	7 (12.5)	8 (13.6)
Nurse/nurse practitioner	8	5 (8.2)	3 (5.4)	5 (8.5)
Social worker	6	5 (8.2)	4 (7.1)	5 (8.5)
Cognitive trainer	1	1 (1.6)	1 (1.8)	1 (1.7)
Researcher	1	1 (1.6)	1 (1.8)	1 (1.7)
Reintegration coordinator	1	1 (1.6)	1 (1.8)	0 (0.0)
Elderly care physician	2	2 (3.3)	2 (3.6)	2 (3.4)
Total	80 (100)	61 (100)	56 (100)	59 (100)

Table II. Instrument selection procedure in first, second and third round

	Instruments		
Category	R1, *n*	R2, *n* old/new/total	R3, *n*
General functioning	9	9/1/10	n.a.
General psychological functioning	7	5/2/7	n.a.
Fatigue	5	5/3/8	7
Anxiety & depression	6	5/0/5	n.a.
Physical functioning	16	16/21[c]/21[c]	12[c]
Aphasia	6	6/4/10	7
Communication (impairment)	1	1/4/5	n.a.
Cognitive functioning	0	0/13[c]/13[c]	7
Subjective cognitive functioning	2	n.a.	n.a.
Disease awareness	2	2/3/5	5
Pain	7	7/2[b]/9	n.a.
Comorbidity	2	2/1/3	2
Activities of daily living	7	7/4/11	11
Participation	11	11/2/13	9
Employment status	2	2/3/5	4
Personal factors	7	5/0/5	n.a.
Quality of life	9	7/1[a]/8	n.a.
Total	99	90/64/138	64

[a]Targeting quality of life in persons with dementia in nursing home. [b]Targeting patients with mental disorders. [c]Combinations; n.a.: not applicable, consensus was reached in previous round; R: round.

category cognitive functioning and 21 combinations for physical functioning. Seventeen physical functioning combinations contained the Berg Balance Scale. Further details about the instrument selection procedure are shown in Table II.

Second round

In the second round, consensus was reached for 7 categories: for quality of life the Quality of Life after Brain Injury scale (QOLIBRI) was selected as first-choice instrument, for pain the visual analogue scale (VAS), for general functioning the Utrecht Scale for the Evaluation of clinical Rehabilitation (USER), for anxiety and depression the Hospital Anxiety and Depression Scale (HADS), for general psychological functioning the Symptom Checklist-90 (SCL-90), for communication (impairment) the Screening Test for Cognition and Communication (STCC), and for personal factors the Utrecht Coping List (UCL). Consensus levels reached are shown in Table III. From the 89 instruments/combinations that were presented in the remaining categories, 25 were not selected for the third round and were removed from the list. From the 21 test combinations for physical functioning 12 test combinations were selected for the third round. In addition, 7 out of 13 cognitive functioning test batteries were chosen for the next round.

Third round

In the third and final round, consensus was reached for 4 categories: for activities of daily living the Barthel Index was selected as first-choice instrument, for activities and participation the USER-Participation, for self-awareness the Awareness Questionnaire (AQ), and for aphasia the ScreeLing (Table III). No consensus was reached for physical functioning, cognitive functioning, comorbidity, fatigue and employment status. In total, consensus was reached for 12 out of 17 categories. All selected measurement instruments, are validated and available in the English literature, except the STCC, which is a Dutch validated instrument (see Appendix I for references). Although no consensus was reached for cognitive functioning, 2 options were preferred: 1) the Allen Cognitive Level Screening (ACLS) and 2) the combination of Stroop, Trail Making Test (TMT), 15 Word Test (15WT), Visuoconstruction

Table III. Number of respondents per Delphi round and consensus level reached

Category	1st Round, (*n*)	2nd Round, (*n*)	3rd Round, (*n*)	Round consensus reached	Consensus level % (*n*)	1st choice instrument
Subjective cognitive functioning	15	n.a.	n.a.	1	80.0 (12)	CFQ
Quality of life	21	22	n.a.	2	86.4 (19)	QOLIBRI
Pain	29	30	n.a.	2	80.0 (24)	VAS
General functioning	29	29	n.a.	2	79.3 (23)	USER
Anxiety & depression	21	18	n.a.	2	94.4 (17)	HADS
General psychological functioning	11	13	n.a.	2	76.9 (10)	SCL-90
Communication (impairment)	8	13	n.a.	2	76.9 (10)	STCC
Personal factors	12	17	n.a.	2	76.5 (13)	UCL
ADL	28	33	36	3	77.8 (28)	BI
Participation	28	31	34	3	76.5 (26)	USER-P
Disease-awareness	13	18	19	3	84.2 (16)	AQ
Aphasia	15	12	10	3	90.0 (9)	ScreeLing
Physical functioning	22	30	27	n.a.	44.4 (12)	BBS, FAC, TCT,BFM
Cognitive functioning	17	17	27	n.a.	40.7 (11)/40.7 (11)	ACLS/Combi A
Comorbidity	4	9	11	n.a.	72.7 (8)	CIRS
Fatigue	37	33	35	n.a.	57.1 (20)	VAS
Employment status	5	9	11	n.a.	45.5 (5)/45.5 (5)	ERS/FML

CFQ: Cognitive Failures Questionnaire; QOLIBRI: Quality of Life after Brain Injury; VAS: visual analogue scale; USER: Utrecht Scale for the Evaluation of Clinical Rehabilitation; HADS: Hospital Anxiety and Depression Scale; SCL-90: Symptom Checklist-90; STCC: Screening Test for Cognition and Communication; UCL: Utrecht Coping List; BI: Barthel Index; USER-P: Utrecht Scale for the Evaluation of Clinical Rehabilitation – Participation; AQ: Awareness Questionnaire; BBS: Berg Balance Scale; FAC: Functional Ambulation Categories; TCT: Trunk Control Test; BFM: Brunnstrom Fugl-Meyer assessment; Combi A: Stroop, TMT, 15WT, visuoconstruction, zoo map BADS; ACLS: Allen Cognitive Level Screening; TMT: Trail Making Test; 15WT: 15 Word Test; CIRS: Cumulative Illness Rating Scale; ERS: Employability Rating Scale; FML: Functionele mogelijkheden Lijst. First-choice instrument references are available in Appendix II.

and ZOO map (BADS). Both options were preferred by 40.7% of the participants. Equivalent outcomes were found for the category employment status. Both the Functionele Mogelijkheden Lijst (FML) and Employability Rating Scale (ERS) were preferred by 45.5% of the participants. For physical functioning, the combination of the Berg Balance Scale (BBS), the Functional Ambulation Categories (FAC), the Trunk Control Test (TCT), and the Brunnstrom Fugl-Meyer Assessment (FMA) was preferred by 44% of the participants. Although 72.7% is below the threshold of 75% consensus, the Cumulative Illness Rating Scale (CIRS) was the most preferred instrument to measure comorbidity. With a consensus level of 57.1% for the VAS scale, agreement was low for fatigue. In 91% of the categories, the first-choice instrument was also the most familiar instrument.

DISCUSSION

An online Delphi procedure was performed aimed at defining outcome measures for the national Neurotraumatology Quality Registry (Net-QuRe), which was initiated to evaluate treatment strategies and to optimize the quality of the total healthcare chain for patients with moderate to severe TBI, as an add-on study of Center-TBI (5). This study identified multiple preferred outcome measures for TBI assessment within several pre-defined outcome categories, selected by experts in the field. In this study consensus on outcome measures was reached in 12 of 17 categories.

Multiple outcome measures were listed in the Delphi questionnaire and further divided into outcome categories based on the ICF model (10). Choosing ICF levels was challenging. For example, the decision to select both employment status and participation as separate categories. In the ICF model employment status is part of the broader category participation, but for the current study it was of interest to focus in more detail on employment status. The fact that participants suggested "new" instruments that were already listed in another category could be a consequence of this decision. Outcome categories form the backbone of a Delphi procedure targeting outcome measures, and should therefore be selected with care. Not surprisingly, selecting outcome categories/domains instead of instruments is the main goal of Delphi studies performed by other researchers, for example targeting outcome domains in patients with cancer (20), eczema (21), psoriatic arthritis (22), and ankylosing spondylitis (23). The categories selected in other studies, partly cover the 17 categories selected for Net-QuRe (20–23). For 10 out of 12 categories, the first-choice instrument was also the most familiar instrument. It could be debated whether this is desirable. In case the most familiar measure is the most used because of its psychometric quality, there is no need to change. In all other cases, being aware of new instruments is of high importance. The use of a measurement instrument is predicted by its psychometric quality, practicality, and having positive attitudes regarding the added value of an instrument over clinical judgement. Of these, attitudes and practicality seem to be the strongest predictors (24). This Delphi procedure focused mainly on the preferences of clinical experts, who were instructed to take into account psychometric quality, sensitivity to measure change, time efficiency, and suitability for patients with moderate to severe TBI. Because using validated measurement instruments is the norm among clinicians and researchers working in the Netherlands, detailed psychometric information about the listed outcome measures was not provided to the participants, which is a limitation of this study.

The strength of a Delphi procedure is the fact that it is based on expert opinion, with the ultimate goal of reaching consensus as a group. By making the procedure anonymous the method tackles problems that are related to decision-making in committees of groups, whereby 1 person or a group of persons dominate the process (25). This online Delphi procedure is also a cost- and time-efficient method, because there is no need to travel for face-to-face meetings, and respondents may choose their own moment for completing the online questionnaires.

A Delphi procedure is a flexible approach, for which methodological choices must be made by the researchers. In a systematic review, 15 other studies were examined that used a Delphi procedure to determine outcome measures. Methodological variability was found across these studies (16). Professionals from different disciplines from 8 rehabilitation centres were invited to create a multidisciplinary respondent group of 80 experts. Instead of inviting participants to respond to a specific category based on their profession, professionals were allowed to decide for each category whether their knowledge was sufficient to respond or to go on to the next section. This procedure minimized the risk of excluding participants who acquired their knowledge by experience. Secondly, experts could participate in each round regardless of involvement in the previous round. Whether this influenced the consensus reached remains unclear, but it certainly resulted in a high response rate per round compared with the response rate in other Delphi studies targeting outcome measures, focusing on pulmonary arterial hypertension (44%) (26), induction of labour (54%) (27), scleroderma (49%) (28), and primary healthcare (33%) (29). Another explanation for the high response

rate in the current study could be the 2-step invitation procedure. We first selected contact persons in the participating centres, who subsequently asked their colleagues if they were willing to participate. This approach was also found in 2 other Delphi studies targeting outcome measures for patients with head and neck cancer (30), and balance outcome measures in cerebellar ataxia (31), in which the response rates were comparable with this study (respectively, 71% and 87%). A limitation of this approach is that it could have resulted in selection bias, because respondents were not randomly selected. In this study the focus was on the first-choice instrument of each category; therefore we did not combine the first- and second-choice instruments. The second-choice instrument was used to provide researchers and participants with additional information about their preferences.

A strength of an online Delphi procedure is that participants could save the results and continue later. However, this procedure could lead to some duplication during data-export. This was identified during the second round, having only a marginal influence on the response feedback of the first round. Nevertheless, this is a limitation of this procedure.

In conclusion, consensus was reached by experts on 12 out of 17 categories of measurement instruments. The Delphi procedure is a feasible method to select measurement instruments for a research project. It appears to be a valuable method on the borderline of qualitative and quantitative research. Although this study focuses on outcome measures that are preferred in the Netherlands, the methods and challenges described will inform other researchers who are planning to perform a Delphi procedure targeting outcome measures.

ACKNOWLEDGEMENTS

The authors thank Basalt, UMCG Centrum voor Revalidatie, Heliomare, Libra Revalidatie, RMC Groot Klimmendaal, Reade, Rijndam Revalidatie, and independent experts. Part of this work was presented at the 11th World Congress on Brain Injury, The Hague, The Netherlands, 2 March 2016. This study was funded by the Dutch Brain Foundation (PS2014.06).

The authors have no conflicts of interest to declare.

REFERENCES

1. Ponsford JL, Downing MG, Olver J, Ponsford M, Acher R, Carty M, et al. Longitudinal follow-up of patients with traumatic brain injury: outcome at two, five, and ten years post-injury. J Neurotrauma 2014; 31: 64–77.
2. Andelic N, Howe EI, Hellstrom T, Sanchez MF, Lu J, Lovstad M, et al. Disability and quality of life 20 years after traumatic brain injury. Brain Behav 2018; 8: e01018.
3. Dahm J, Ponsford J. Long-term employment outcomes following traumatic brain injury and orthopaedic trauma: a ten-year prospective study. J Rehabil Med 2015; 47: 932–940.
4. Khan F, Amatya B, Judson R, Chung P, Truesdale M, Elmalik A, et al. Factors associated with long-term functional and psychological outcomes in persons with moderate to severe traumatic brain injury. J Rehabil Med 2016; 48: 442–448.
5. Van Essen TA, Volovici V, Cnossen MC, Kolias A, Ceyisakar I, Nieboer D, et al. Comparative effectiveness of surgery in traumatic acute subdural and intracerebral haematoma: study protocol for a prospective observational study within CENTER-TBI and Net-QuRe. BMJ Open 2019; 9: e033513.
6. Tso S, Saha A, Cusimano MD. The Traumatic Brain Injury Model Systems National Database: a review of published research. Neurotrauma Rep 2021; 2: 149–164.
7. Tate RL, Godbee K, Sigmundsdottir L. A systematic review of assessment tools for adults used in traumatic brain injury research and their relationship to the ICF. NeuroRehabilitation 2013; 32: 729–750.
8. van Heugten C, Caldenhove S, Crutsen J, Winkens I. An overview of outcome measures used in neuropsychological rehabilitation research on adults with acquired brain injury. Neuropsychol Rehabil 2019: 1–26.
9. Laxe S, Tschiesner U, Zasler N, Lopez-Blazquez R, Tormos JM, Bernabeu M. What domains of the International Classification of Functioning, Disability and Health are covered by the most commonly used measurement instruments in traumatic brain injury research? Clin Neurol Neurosurg 2012; 114: 645–650.
10. World Health Organization (WHO). International Classification of Functioning, Disability and Health (ICF). Geneva: WHO; 2001.
11. Wilde EA, Whiteneck GG, Bogner J, Bushnik T, Cifu DX, Dikmen S, et al. Recommendations for the use of common outcome measures in traumatic brain injury research. Arch Phys Med Rehabil 2010; 91: 1650–1660 e17.
12. Hicks R, Giacino J, Harrison-Felix C, Manley G, Valadka A, Wilde EA. Progress in developing common data elements for traumatic brain injury research: version two--the end of the beginning. J Neurotrauma 2013; 30: 1852–1861.
13. Honan CA, McDonald S, Tate R, Ownsworth T, Togher L, Fleming J, et al. Outcome instruments in moderate-to-severe adult traumatic brain injury: recommendations for use in psychosocial research. Neuropsychol Rehabil 2019; 29: 896–916.
14. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000; 32: 1008–1015.
15. Limesurvey. 2015 [accessed 2015 April]. Available from: https://www.limesurvey.org/.
16. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. PLoS Med 2011; 8: e1000393.
17. Powell C. The Delphi technique: myths and realities. J Adv Nurs 2003; 41: 376–382.
18. meetinstrumenten in de zorg. 2015 [accessed 2015 May]. Available from: https://meetinstrumentenzorg.nl.
19. Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005; 53: 695–699.
20. Howell D, Fitch M, Bakker D, Green E, Sussman J, Mayo S, et al. Core domains for a person-focused outcome measurement system in cancer (PROMS-Cancer Core) for routine care: a scoping review and Canadian Delphi Consensus. Value Health 2013; 16: 76–87.
21. Schmitt J, Langan S, Stamm T, Williams HC. Core outcome domains for controlled trials and clinical recordkeeping in eczema: international multiperspective Delphi consensus process. J Invest Dermatol 2011; 131: 623–630.
22. Taylor WJ. Preliminary identification of core domains for

outcome studies in psoriatic arthritis using Delphi methods. Ann Rheum Dis 2005; 64 Suppl 2: ii110–2.
23. Zochling J, Sieper J, van der Heijde D, Braun J. Development of a core set of domains for data collection in cohorts of patients with ankylosing spondylitis receiving anti-tumor necrosis factor-alpha therapy. J Rheumatol 2008; 35: 1079–1082.
24. Jensen-Doss A, Hawley KM. Understanding barriers to evidence-based assessment: clinician attitudes toward standardized assessment tools. J Clin Child Adolesc Psychol 2010; 39: 885–896.
25. Jones J, Hunter D. Consensus methods for medical and health services research. BMJ 1995; 311: 376–380.
26. Distler O, Behrens F, Pittrow D, Huscher D, Denton CP, Foeldvari I, et al. Defining appropriate outcome measures in pulmonary arterial hypertension related to systemic sclerosis: a Delphi consensus study with cluster analysis. Arthritis Rheum 2008; 59: 867–875.
27. Dos Santos F, Drymiotou S, Antequera Martin A, Mol BW, Gale C, Devane D, et al. Development of a core outcome set for trials on induction of labour: an international multistakeholder Delphi study. BJOG 2018; 125: 1673–1680.
28. Gazi H, Pope JE, Clements P, Medsger TA, Martin RW, Merkel PA, et al. Outcome measurements in scleroderma: results from a delphi exercise. J Rheumatol 2007; 34: 501–509.
29. Hutchinson A, Fowler P. Outcome measures for primary health care: what are the research priorities? Br J Gen Pract 1992; 42: 227–231.
30. Kirschneck M, Sabariego C, Singer S, Tschiesner U. Assessment of functional outcomes in patients with head and neck cancer according to the International Classification of Functioning, Disability and Health Core Sets from the perspective of the multi-professional team: results of 4 Delphi surveys. Head Neck 2014; 36: 954–968.
31. Winser SJ, Smith C, Hale LA, Claydon LS, Whitney SL. Balance outcome measures in cerebellar ataxia: a Delphi survey. Disabil Rehabil 2015; 37: 165–170.

Appendix I. First-choice instrument references.

Cognitive Failures Questionnaire (CFQ) (S1)
Quality of Life after Brain Injury (QOLIBRI) (S2)
Visual analogue scale (VAS) (S3)
Utrecht Scale for the Evaluation of clinical Rehabilitation (USER) (S4)
Hospital Anxiety and Depression Scale (HADS) (S5)
Symptom Checklist-90 (SCL-90) (S6)
ScreeningTest for Cognition and Communication (STCC) (S7)
Utrecht Coping List (UCL) (S8, S9)
Barthel Index (BI) (S10)
Utrecht Scale for the Evaluation of clinical Rehabilitation – Participation (USER-P) (S11)
Awareness Questionnaire (AQ) (S12)
ScreeLing (S13)
Berg Balance Scale (BBS) (S14)
Functional Ambulation Categories (FAC) (S15)
Trunk Control Test (TCT) (S16)
Brunnstrom Fugl-Meyer assessment (BFM) (S17)
Allen Cognitive Level Screening (ACLS) (S18)
Stroop test (S19)
Visuoconstruction, zoo map BADS (S20)
Trail Making Test (TMT) (S21)
15 Word Test (15WT) (S22)
Cumulative Illness Rating Scale (CIRS) (S23)
Employability Rating Scale (ERS) (S24)

REFERENCES

S1. Broadbent DE, Cooper PF, FitzGerald P, Parkes KR. The Cognitive Failures Questionnaire (CFQ) and its correlates. Br J Clin Psychol 1982; 21: 1–16.
S2. von Steinbuchel N, Wilson L, Gibbons H, Hawthorne G, Hofer S, Schmidt S, et al. Quality of Life after Brain Injury (QOLIBRI): scale validity and correlates of quality of life. J Neurotrauma 2010; 27: 1157–1165.
S3. Freyd M. The graphic rating scale. J Educ Psychol 1923; 14: 83–102.
S4. Post MW, van de Port IG, Kap B, Berdenis van Berlekom SH. Development and validation of the Utrecht Scale for Evaluation of Clinical Rehabilitation (USER). Clin Rehabil 2009; 23: 909–917.
S5. Zigmond AS, Snaith RP. The Hospital Anxiety and Depression Scale. Acta Psychiatr Scand 1983; 67: 361–370.
S6. Derogatis LR, Lipman RS, Covi L. SCL-90: an outpatient psychiatric rating scale – preliminary report. Psychopharmacol Bull 1973; 9: 13–28.
S7. Paemeleire F. [The ScreeningTest for Cognition and Communication (STCC): a new instrument for adults with acquired brain injury (ABI).] Logopedie 2014; 27: 51–65.
S8. Schreurs PJG, Tellegen B, van de Willige G. [Health, stress and coping: development of the Utrecht Coping List.]1984; 12: 101–117.
S9. Turner H, Bryant-Waugh R, Peveler R, Bucks RS. A psychometric evaluation of an English version of the Utrecht Coping List. Eur Eat Disord Rev 2012; 20: 339–342.
S10. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. Md State Med J 1965; 14: 61–65.
S11. Post MW, van der Zee CH, Hennink J, Schafrat CG, Visser-Meily JM, van Berlekom SB. Validity of the Utrecht Scale for evaluation of rehabilitation-participation. Disabil Rehabil 2012; 34: 478–485.
S12. Sherer M, Bergloff P, Boake C, High W Jr, Levin E. The Awareness Questionnaire: factor structure and internal consistency. Brain Inj 1998; 12: 63–68.
S13. Doesborgh SJ, van de Sandt-Koenderman WM, Dippel DW, van Harskamp F, Koudstaal PJ, Visch-Brink EG. Linguistic deficits in the acute phase of stroke. J Neurol 2003; 250: 977–982.
S14. Berg K, Wood-Dauphinee S, Williams JI, Gayton D. Measuring balance in the elderly: preliminary development of an instrument. Physiotherapy Canada 1989; 41: 304–311.
S15. Holden MK, Gill KM, Magliozzi MR, Nathan J, Piehl-Baker L. Clinical gait assessment in the neurologically impaired. Reliability and meaningfulness. Phys Ther 1984; 64: 35–40.
S16. Collin C, Wade D. Assessing motor impairment after stroke: a pilot reliability study. J Neurol Neurosurg Psychiatry 1990; 53: 576–579.
S17. Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. a method for evaluation of physical performance. Scand J Rehabil Med 1975; 7: 13–31.
S18. Earhart CA, Allen CK, Blue T. Allen Diagnostic Module. Colchester: S&S Worldwide; 1993 (revised 2003).
S19. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935; 18: 643–662.
S20. Wilson BA, Alderman N, Burgess PW, Emslie H, Evans JJ. Behavioural assessment of the dysexecutive syndrome (BADS). London: Thames Valley Test Co.; 1996.
S21. Reitan RM, Wolfson D. The Halstead–Reitan Neuropsychological Test Battery: theory and clinical interpretation. Tucson: Neuropsychology Press; 1985.
S22. Rey A. l'examen clinique en psychologie. Paris: Presse Universitaire de France; 1964.
S23. Linn BS, Linn MW, Gurel L. Cumulative Illess Rating Scale. J Amer Geriatr Soc 1968; 16: 622–626.
S24. Ben-Yishay Y, Silver SM, Piasetsky E, Rattok J. Relationship between employability and vocational outcome after intensive holistic cognitive rehabilitation. J Head Trauma Rehab 1987; 2: 35–48.

Appendix II. Examples of questions asked, and feedback given in this Delphi procedure.

Total number of questions

Round 1: 86

Round 2: 49

Round 3: 23

General questions

In which centre do you work? (Answer options are listed + option 'other, namely')

What is your profession? (Answer options are listed + option 'other, namely')

How many years experience do you have (approximately) with the target group of patients

with acquired brain injury? (open answer field)

Instructions

Various measurement instruments are available to measure outcome and change over time and selecting the best tool is a challenge. A suitable measurement instrument must be valid and reliable, must have high sensitivity to change, must be time efficient, and suitable for patients with brain injury. We perform a Delphi procedure to obtain consensus on the best set of measurement instruments, taking all these elements into account. The procedure consists of several rounds, with a maximum of 3. For each round, we ask you to complete the online survey in which your preference for a suitable measurement instrument should be indicated for each outcome category.

Category: Activities and Participation

Which of the following validated instruments are familiar to you? (multiple answers apply)

- Life-Habits (Life-H)
- Impact on Participation and Autonomy Questionnaire (IPA)
- Participation Scale (P-Scale)
- Model of Human Occupation Screening Tool (MOHOST)
- World Health Organization Disability Assessment Schedule (WHODAS)
- London Handicap Scale
- Community Integration Questionnaire (CIQ)
- Utrecht Activities List (UAL) an adapted version of the Craig Handicap Assessment and Reporting Technique (CHART)
- Frenchay Activities Index (FAI)
- Utrecht Scale for the Evaluation of clinical rehabilitation – participation (USER-P)
- Sickness Impact Profile-68 (SIP-68)
- I do not know any of the instruments mentioned above (this answer excluded all other options)

Do you know of any other instruments missing from this list, please specify? (4 open answer fields)

We are looking for a suitable "Activities and Participation" instrument for patients with moderate to severe traumatic brain injury, what is your first choice? (see instructions; single answer applies).

- I am not the right person to answer questions about "Activities and Participation". *(In this case Limesurvey automatically skipped all other questions regarding this category)*
- Life-Habits (Life-H)
- Impact on Participation and Autonomy Questionnaire (IPA)
- Participation Scale (P-Scale)
- Model of Human Occupation Screening Tool (MOHOST)
- World Health Organization Disability Assessment Schedule (WHODAS)
- London Handicap Scale
- Community Integration Questionnaire (CIQ)
- Utrecht Activities List (UAL) an adapted version of the Craig Handicap Assessment and Reporting Technique (CHART)
- Frenchay Activities Index (FAI)
- Utrecht Scale for the Evaluation of clinical rehabilitation – participation (USER-P)
- Sickness Impact Profile-68 (SIP-68)
- Other, please specify

What is your second choice? (see instructions; single answer applies)

- Life-Habits (Life-H)
- Impact on Participation and Autonomy Questionnaire (IPA)
- Participation Scale (P-Scale)
- Model of Human Occupation Screening Tool (MOHOST)
- World Health Organization Disability Assessment Schedule (WHODAS)
- London Handicap Scale
- Community Integration Questionnaire (CIQ)
- Utrecht Activities List (UAL) an adapted version of the Craig Handicap Assessment and Reporting Technique (CHART)
- Frenchay Activities Index (FAI)
- Utrecht Scale for the Evaluation of clinical rehabilitation – participation (USER-P)
- Sickness Impact Profile-68 (SIP-68)
- Other, please specify

Do you have any comments concerning instruments targeting "Activities and Participation"?

Feedback given to respondents in round 2 concerning the category "Activities and Participation"

Instrument	Number of times 1st choice	Number of times 2nd choice
Utrecht Scale for the Evaluation of clinical rehabilitation – participation (USER-P)	15 (51, 72%)	3 (10, 34%)
Impact on Participation and Autonomy Questionnaire (IPA)	2 (6, 90%)	4 (13, 79%)
Sickness Impact Profile-68 (SIP-68)	2 (6, 90%)	2 (6, 90%)
Model of Human Occupation Screening Tool (MOHOST)	1 (3, 45%)	2 (6, 90%)
World Health Organization Disability Assessment Schedule (WHODAS)	1 (3, 45%)	0
Community Integration Questionnaire (CIQ)	1 (3, 45%)	3 (10, 34%)
Utrecht Activities List (UAL)	1 (3, 45%)	2 (6, 90%)
Frenchay Activities Index (FAI)	1 (3, 45%)	4 (13, 79%)
Life Habits (Life-H)	0	1 (3, 45%)
Participation Scale (P-Scale)	0	1 (3, 45%)
London Handicap Scale	0	1 (3, 45%)
Other, please specify:		
Canadian Occupational Performance Measure (COPM)	3 (10, 34%)	2 (6, 90%)
Own questionnaire/inventory list	1 (3, 45%)	0
Occupational Performance History Interview (OPHI)	0	1 (3, 45%)
No answer given	1 (3, 45%)	3 (10, 34%)
Total	29 (100%)	29 (100%)