Factorial validity and comparability of the six translations of the Rivermead Post-Concussion Symptoms Questionnaire translations: results from the CENTER-TBI study

Marina Zeldovich[1]* , Fabian Bockhop[1], Amra Covic[1], Isabelle Mueller[1,2], Suzanne Polinder[3], Ana Mikolic[4,5], Marjolein van der Vlegel[3], Nicole von Steinbuechel[1] and CENTER-TBI participants and investigators

Abstract

Background Comparison of patient-reported outcomes in multilingual studies requires evidence of the equivalence of translated versions of the questionnaires. The present study examines the factorial validity and comparability of six language versions of the Rivermead Post-Concussion Symptoms Questionnaire (RPQ) administered to individuals following traumatic brain injury (TBI) in the Collaborative European NeuroTrauma Effectiveness Research (CENTER-TBI) study.

Methods Six competing RPQ models were estimated using data from Dutch (n = 597), English (n = 223), Finnish (n = 213), Italian (n = 268), Norwegian (n = 263), and Spanish (n = 254) language samples recruited six months after injury. To determine whether the same latent construct was measured by the best-fitting model across languages and TBI severity groups (mild/moderate vs. severe), measurement invariance (MI) was tested using a confirmatory factor analysis framework.

Results The results did not indicate a violation of the MI assumption. The six RPQ translations were largely comparable across languages and were able to capture the same construct across TBI severity groups. The three-factor solution comprising emotional, cognitive, and somatic factors provided the best fit with the following indices for the total sample: χ^2 (101) = 647.04, χ^2/df = 6.41, $p < 0.001$, CFI = 0.995, TLI = 0.994, RMSEA = 0.055, CI$_{90\%}$[0.051, 0.059], SRMR = 0.051.

Conclusion The RPQ can be used in international research and clinical settings, allowing direct comparisons of scores across languages analyzed within the full spectrum of TBI severity. To strengthen the aggregated applicability across languages, further analyses of the utility of the response scale and comparisons between different translations of the RPQ at the item level are recommended.

Keywords Traumatic brain injury, Post-concussion symptoms, Measurement invariance, Rivermead Post-Concussion Symptoms Questionnaire

*Correspondence:
Marina Zeldovich
marina.zeldovich@med.uni-goettingen.de
Full list of author information is available at the end of the article

Zeldovich *et al. Journal of Patient-Reported Outcomes*   (2023) 7:90

Page 2 of 18

Background

Traumatic brain injury (TBI) is a condition characterized by changes in brain functioning caused by external head trauma [1]. It imposes life-long limitations [2] and leads to a range of physical, emotional, and cognitive disabilities, impacting functioning of affected individuals [1]. The burden of TBI extends to family caregivers [3], as well as health and economic systems [2].

Individuals after TBI can especially suffer from a range of post-concussion symptoms (PCS), which may persist much longer than initially expected [4]. These symptoms encompass physical (e.g., headaches or nausea), cognitive (e.g., a diminished ability to concentrate), and emotional/behavioral (e.g., depressiveness or fatigue) impairments [5]. PCS are commonly reported after mild to moderate TBI [6], but individuals following severe TBI also experience similar deficits [7, 8], referred to as PC-like symptoms.

To assess PCS, researchers and clinicians often rely on the subjective experiences of those affected using patient-reported outcome measures (PROMs), such as the Rivermead Post-Concussion Symptoms Questionnaire (RPQ) [6]. The RPQ is widely used to assess self-reported PCS. For the Collaborative European NeuroTrauma Effectiveness Research study (CENTER-TBI; clinicaltrials.gov NCT02210221), which was designed to examine treatment and outcomes of individuals after TBI in 18 countries [9], translations and linguistic validations were performed for the RPQ, resulting in eleven additional versions [10].

In multilingual studies, the equivalence of translated PROMs, in terms of their conceptual alignment with the original version and other translations, cultural relevance, acceptability to the target populations, and psychometrical comparability, is essential for language and country comparisons, as well as data aggregation in multilingual studies [11]. Measurement invariance (MI) analysis is a valuable tool to determine whether the translations of an instrument measure the same construct [12]. In particular, MI analysis investigates if the differences in observed variables across language versions are solely attributed to differences in latent means.

Therefore, the main aim of the present study is to provide empirical evidence of MI for the RPQ in six European languages. The RPQ has been declared a unidimensional PROM, but this property could not be replicated across translations, including English-speaking samples e.g., [13]. Thus, the factorial structure of the RPQ is examined to find the best-fitting model prior to MI analyses. In addition, the study seeks to explore whether the same construct is measured across the spectrum of TBI severity.

Once the assumption of MI is met, differences in RPQ scores between language samples will be due to true differences in self-reported PCS and not to differences in translation allowing for data aggregation and direct comparisons.

Methods

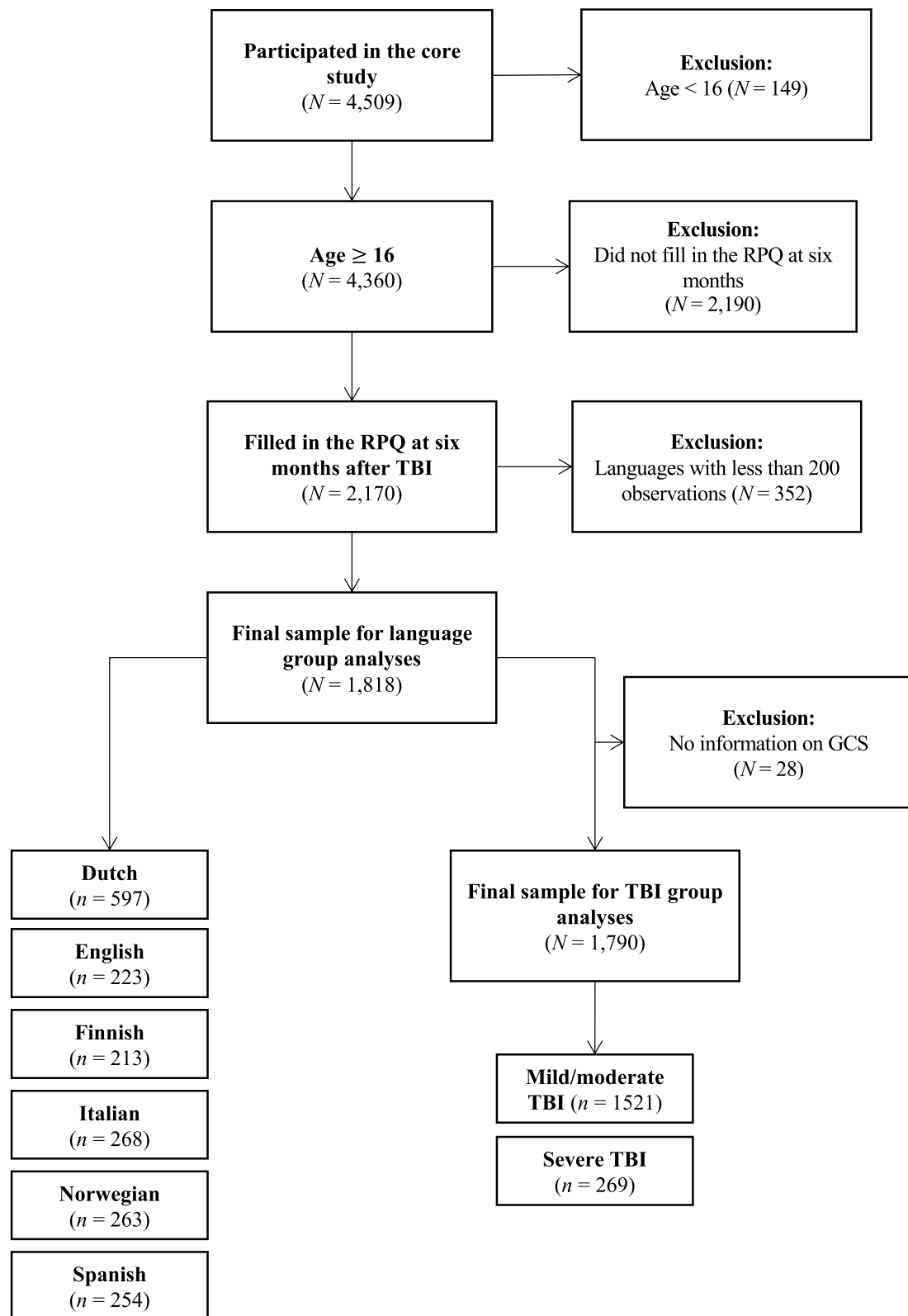### Study design and participants

Data were collected from December 2014 to December 2019 within the CENTER-TBI study involving 63 centers in 18 countries in Europe and Israel. A total of 4,509 individuals after TBI were enrolled in the core study. Inclusion criteria for study participation were clinical diagnosis of TBI and indication for computed tomography (CT) scan, enrollment within 24 h after injury, and informed consent for study participation. Written informed consent was obtained according to the local and national legislations for all patients (either by the patients or the legal representatives) and documented in the electronic case report form. To avoid bias in outcome assessment, patients with severe pre-existing neurological disorders were excluded from the study. Individuals were either seen in the emergency room (ER) and then discharged or either admitted to the hospital ward or intensive care unit (ICU). More detailed description of the study design is provided by Steyerberg et al. [14]. Data were retrieved from the CENTER-TBI database via Neurobot tool (core data set 2.1, November 2019).

The following analyses were limited to individuals aged 16 years of age or older who completed the RPQ six months after TBI. Following the recommendation for MI analyses, we included language samples with at least 200 participants (N = 1,818). Additional analyses on TBI severity involved individuals with available information on the Glasgow Coma Scale (GCS) [15] score at baseline (N = 1,790). For the composition of the study sample, see Fig. 1.

Sample characteristics

Sociodemographic characteristics were collected at study enrollment and included sex, age, education (in groups and years), employment status, marital status, and living situation. Language samples were identified according to the languages spoken in the participating sites. For more details on language sample compositions in the CENTER-TBI study, see von Steinbuechel et al. [13].

The following variables were used to characterize premorbid and injury-related factors: mental health status before the injury (presence vs. absence of prior psychiatric diseases), cause of injury, clinical care pathways (ER, ward, ICU), and TBI severity measured using the Glasgow Coma Scale at baseline (GCS) [15] combined with information on abnormalities on the CT scan

```
┌─────────────────────────┐          ┌─────────────────────────┐
│  Participated in the core│────────▶ │       Exclusion:        │
│          study           │          │  Age < 16 (N = 149)     │
│      (N = 4,509)         │          │                         │
└─────────────────────────┘          └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐          ┌─────────────────────────┐
│      Age ≥ 16            │────────▶ │       Exclusion:        │
│     (N = 4,360)          │          │ Did not fill in the RPQ │
│                          │          │ at six months           │
│                          │          │     (N = 2,190)         │
└─────────────────────────┘          └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐          ┌─────────────────────────┐
│  Filled in the RPQ at six│────────▶ │       Exclusion:        │
│   months after TBI       │          │ Languages with less than│
│     (N = 2,170)          │          │ 200 observations (N=352)│
└─────────────────────────┘          └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ Final sample for language│
│    group analyses        │
│     (N = 1,818)          │
└─────────────────────────┘
```

Fig. 1 Sample attrition

Zeldovich *et al. Journal of Patient-Reported Outcomes*   (2023) 7:90

Page 4 of 18

(uncomplicated mild, complicated mild, moderate, and severe TBI) [15, 16]. The functional recovery status at six months was rated using the Glasgow Outcome Scale – Extended (GOSE) [17]. The total injury severity score (ISS) and the brain injury severity score from the Abbreviated Injury Scale (AIS) [18] assessed total injury severity and brain injury severity, respectively.

The Rivermead Post-Concussion Symptoms Questionnaire
The Rivermead Post-Concussion Symptoms Questionnaire (RPQ) [6] assesses 16 symptoms including headaches, dizziness, nausea and/or vomiting, noise sensitivity, sleep disturbance, fatigue, irritability, depression, frustration, forgetfulness and poor memory, poor concentration, slow thinking, blurred vision, light sensitivity, double vision, and restlessness. Individuals are asked to rate the symptoms over the last 24 h compared with their condition before the TBI using a five-point Likert-type scale (from 0 "not experienced at all" to 4 "a severe problem"). Based on the originally proposed unidimensional factor structure, the total score ranges from 0 to 64 with higher values indicating greater impairment, whereby values of "1" indicating no more of a problem than before TBI are treated as zero. For clinical screening, a cut-off score of 12 can be applied [19]. The factorial structure of the RPQ has so far been the subject of several studies [19–24] and no agreement on an ultimate solution has yet been reached. Initial analyses of the unidimensional structure using data collected in the CENTER-TBI study also revealed rather poor model fit across language samples [13].

Statistical analyses
Descriptive statistics and language-samples comparisons
Prior to statistical analyses, language samples were compared by sample characteristics (esp. injury-related factors) using Kruskal–Wallis tests and pairwise U-tests to ensure their general comparability. For the pairwise U-tests, Vargha and Delaney's effect size [25] was calculated based on the following cut-offs: groups equality (0.50), small (0.35–0.44 or 0.56–0.63), medium (0.30–0.34 or 0.64–0.70), and large effect (beyond 0.29 or 0.71). Distribution of the TBI severity groups in the language samples has been investigated applying a two-dimensional chi-squared test and computing Cramer's V to determine the effect size. For this purpose, we first used the initial distribution containing four TBI severity groups (i.e., uncomplicated mild, complicated mild, moderate, and severe TBI), and then a collapsed classification (i.e., mild/moderate and severe TBI). The effect size was determined using Cohen's taxonomy [26] with values of 0.10, 0.30, and 0.50 representing small, medium, and

large effects, respectively. Furthermore, the distribution of PCS in language samples has been investigated.

Analyses of dimensionality
First, we analyzed the response pattern of the RPQ items within the language samples. The factorial structure of the RPQ translations was then examined separately for each language version using confirmatory factor analysis (CFA) with robust weighted least squares estimator (WLSMV) [27] for ordered categorical data. In the absence of agreement on the factorial structure of the RPQ, six competing models derived from previous research were estimated: the initial one-factor model [6], two two-factor models [19, 20], two three-factor models [21, 22], the first of which [21] was based on research findings presented by Gerber and Schraa [28], and one four-factor [23] model, corresponding to the study findings by Lannsjö and colleagues [24]. For more details on RPQ models estimated in the present study, see Table 1.

The fit of all estimated models was assessed by several goodness of fit indices: χ^2 and degrees of freedom (*df*), as well as the ratio χ^2/df, the comparative fit index (CFI) [29], the Tucker-Lewis index (TLI) [30] the root mean square error of approximation (RMSEA) [31] including 90% confidence interval ($CI_{90\%}$), and the standardized root mean square residual (SRMR). A ratio $\chi^2/df \leq 2$ indicate good model fit [32], CFI and TLI values larger than 0.95 indicate a good fit [33], the RMSEA values less than 0.05 signal a close fit, values from 0.05 to 0.08 a fair fit, between 0.08 and 0.10 a mediocre and above 0.10 a poor fit [34]; the same criteria apply to the CIs. SRMR values less than 0.08 demonstrate a good model fit [33]. Since the cut-off values for the CFI, TLI, and RMSEA have not yet been validated for ordinal data, interpretation should be carried out with caution [35]. Therefore, all fit measures were considered simultaneously to identify the best fitting model.

For the CFI analyses, we first used the raw data to obtain the model fit and then the data with modified items due to missing responses in some categories. Thus, responses 3 "moderate problem" or 4 "severe problem" of the items *Nausea* and *Double vision* were collapsed with the category 2 "mild problem" implying a trichotomized response format when considering 1 "no more of a problem than before TBI" as "1" and a dichotomized scale when treating "1" as "0". For all other items, the original responses were kept to maximally retain information.

Measurement invariance (MI)
The best-fitting factor solution served as the basis for MI analyses across languages and for TBI severity MI analyses. Because the absence of responses in some categories would not allow for invariance testing across groups,

Table 1 Factorial structures of the RPQ investigated in the present study

Item label (shortened version)	One-factor model [6]	Two-factor model [19]	Two-factor model [20]	Three-factor model [21]	Three-factor model [22]	Four-factor model [23]
Headaches	RPQ	Emotional-somatic	RPQ-3	Somatic	General somatic	Vertigo
Dizziness	RPQ	Emotional-somatic	RPQ-3	Somatic	General somatic	Vertigo
Nausea	RPQ	Emotional-somatic	RPQ-3	Somatic	General somatic	Vertigo
Noise sensitivity	RPQ	Emotional-somatic	RPQ-13	Somatic	General somatic	Vertigo
Sleep disturbance	RPQ	Emotional-somatic	RPQ-13	Somatic	General somatic	Mood/somatic
Fatigue	RPQ	Emotional-somatic	RPQ-13	Somatic	Mood/cognition	Mood/somatic
Irritable	RPQ	Emotional-somatic	RPQ-13	Emotional	Mood/cognition	Mood/somatic
Depressed	RPQ	Emotional-somatic	RPQ-13	Emotional	Mood/cognition	Mood/somatic
Frustrated	RPQ	Emotional-somatic	RPQ-13	Emotional	Mood/cognition	Mood/somatic
Forgetful	RPQ	Cognitive	RPQ-13	Cognitive	Mood/cognition	Cognitive
Poor concentration	RPQ	Cognitive	RPQ-13	Cognitive	Mood/cognition	Cognitive
Longer to think	RPQ	Cognitive	RPQ-13	Cognitive	Mood/cognition	Cognitive
Blurred vision	RPQ	Emotional-somatic	RPQ-13	Somatic	Visual somatic	Vision
Light sensitivity	RPQ	Emotional-somatic	RPQ-13	Somatic	Visual somatic	Vision
Double vision	RPQ	–[a]	RPQ-13	Somatic	Visual somatic	Vision
Restless	RPQ	Emotional-somatic	RPQ-13	Emotional	General somatic	Mood/Somatic

[a] Item Double vision was excluded from the model

only modified items were considered. We conducted a multi-group CFA with stepwise increasing constraints following the framework proposed by Wu and Estabrook [36] and updated by Svetina et al. [37] to make it more suitable for Likert-type scales. This approach slightly differs from the classical MI procedure. First, the baseline model testing for configural invariance was fitted. Then, this model was restrained by requiring invariance of thresholds, and thresholds and loadings across the groups. The models were stepwise compared by calculating the chi-square difference test and changes in CFI (ΔCFI) and RMSEA (ΔRMSEA). Models showing non-significant differences ($p \geq 0.05$), ΔCFI < 0.01 [38], and ΔRMSEA ≤ 0.01, which is recommended for groups with unequal sample sizes [39], were considered equivalent. If the difference tests between the models were not significant, the assumption of MI was considered justified. Once the MI assumption was fulfilled, a multi-group CFA approach was again used to examine the differences between mild and moderate/severe TBI across all RPQ translations. Finally, the best-fitting model was estimated and visualized for the total study sample.

According to the original scoring [6], "1" responses indicating that a symptom posed no more of a problem than before TBI should be treated as "0". Because participants explicitly used a five-point Likert-type scale when completing the questionnaire, "1" responses were considered in both CFA and MI analyses. However, we additionally replicated the analyses using the simplified response scale (i.e., treating 1 as 0) to achieve greater congruence with the original scoring procedure. These results are reported in the Additional file 1.

All analyses were carried out with R version 4.0.0. [40] and packages "Table 1" [41] for descriptive analyses and "lavaan" [42] for the CFA and MI testing as well as package "lavaanPlot" for model visualization [43]. The significance level was set at 5% except for pairwise comparisons, for which the Bonferroni correction was performed to avoid alpha inflation $\left(\alpha_{adj} = \frac{.05}{6} = 0.008\right)$.

Results

Sample characteristics

The total sample comprised 1,818 participants (65.4% male) with a mean age of 49.5 ± 19.5 years (*Mdn* = 51, range 16–95) who completed the RPQ at six months after injury. Most individuals sustained a mild TBI (73.1%). The cause of injury was most commonly either a road traffic accident (41.7%) or incidental fall (43.7%). At six months after TBI, more than half of the individuals showed good recovery (GOSE: 7–8) across all language samples. For more details, see Table 2 and Additional file 1: Table S1 – Additional characteristics of the language samples.

Some significant differences were observed between the language samples regarding age, years of education, ISS, GCS at baseline, and GOSE at six months. Dutch participants were significantly older (52.9 ± 19.1) compared to all but the Italian sample. Finnish (13.3 ± 3.16)

Table 2 Characteristics of the language samples

Variable	Group/values	Dutch (N = 597)	English (N = 223)	Finnish (N = 213)	Italian (N = 268)	Norwegian (N = 263)	Spanish (N = 254)	Total (N = 1818)
Sex[a]	Female	229 (38.4%)	74 (33.2%)	84 (39.4%)	84 (31.3%)	83 (31.6%)	75 (29.5%)	629 (34.6%)
	Male	368 (61.6%)	149 (66.8%)	129 (60.6%)	184 (68.7%)	180 (68.4%)	179 (70.5%)	1189 (65.4%)
Age	*M* (*SD*)	52.9 (19.1)	48.3 (17.1)	47.7 (19.6)	50.2 (20.6)	45.8 (19.9)	47.1 (19.4)	49.5 (19.5)
	Mdn [*Min*, *Max*]	57.0 [16.0, 95.0]	51.0 [16.0, 85.0]	50.0 [16.0, 89.0]	53.0 [16.0, 93.0]	48.0 [16.0, 89.0]	44.0 [16.0, 93.0]	51.0 [16.0, 95.0]
Years of education	*M* (*SD*)	14.4 (3.87)	14.9 (3.73)	13.3 (3.16)	12.4 (4.37)	14.1 (3.21)	14.7 (5.60)	14.0 (4.13)
	Mdn [*Min*, *Max*]	15.0 [4.00, 30.0]	14.0 [7.00, 25.0]	12.0 [9.00, 30.0]	13.0 [4.00, 25.0]	14.0 [4.00, 25.0]	15.0 [2.00, 30.0]	14.0 [2.00, 30.0]
	Missing	128 (21.4%)	30 (13.5%)	66 (31.0%)	50 (18.7%)	15 (5.7%)	45 (17.7%)	334 (18.4%)
Injury cause[a]	Incidental fall	287 (48.1%)	88 (39.5%)	98 (46.0%)	95 (35.4%)	111 (42.2%)	116 (45.7%)	795 (43.7%)
	Road traffic accident	238 (39.9%)	104 (46.6%)	70 (32.9%)	135 (50.4%)	105 (39.9%)	107 (42.1%)	759 (41.7%)
	Other	67 (11.2%)	28 (12.6%)	37 (17.4%)	31 (11.6%)	43 (16.3%)	25 (9.8%)	231 (12.7%)
	Missing	5 (0.8%)	3 (1.3%)	8 (3.8%)	7 (2.6%)	4 (1.5%)	6 (2.4%)	33 (1.8%)
Clinical pathways[a]	ER	102 (17.1%)	62 (27.8%)	51 (23.9%)	67 (25.0%)	66 (25.1%)	71 (28.0%)	419 (23.0%)
	Ward	308 (51.6%)	74 (33.2%)	99 (46.5%)	57 (21.3%)	114 (43.3%)	54 (21.3%)	706 (38.8%)
	ICU	187 (31.3%)	87 (39.0%)	63 (29.6%)	144 (53.7%)	83 (31.6%)	129 (50.8%)	693 (38.1%)
GOSE score	3	14 (2.3%)	13 (5.8%)	9 (4.2%)	24 (9.0%)	9 (3.4%)	12 (4.7%)	81 (4.5%)
	4	30 (5.0%)	18 (8.1%)	7 (3.3%)	19 (7.1%)	4 (1.5%)	13 (5.1%)	91 (5.0%)
	5	63 (10.6%)	36 (16.1%)	17 (8.0%)	27 (10.1%)	32 (12.2%)	32 (12.6%)	207 (11.4%)
	6	89 (14.9%)	34 (15.2%)	23 (10.8%)	43 (16.0%)	55 (20.9%)	22 (8.7%)	266 (14.6%)
	7	180 (30.2%)	37 (16.6%)	45 (21.1%)	45 (16.8%)	62 (23.6%)	78 (30.7%)	447 (24.6%)
	8	221 (37.0%)	84 (37.7%)	112 (52.6%)	110 (41.0%)	101 (38.4%)	97 (38.2%)	725 (39.9%)
	Missing	0 (0%)	1 (0.4%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (0.1%)
TBI severity[a]	Uncomplicated mild	275 (46.1%)	83 (37.2%)	96 (45.1%)	75 (28.0%)	110 (41.8%)	80 (31.5%)	719 (39.5%)
	Complicated mild	190 (31.8%)	65 (29.1%)	74 (34.7%)	90 (33.6%)	85 (32.3%)	106 (41.7%)	610 (33.6%)
	Moderate	45 (7.5%)	11 (4.9%)	16 (7.5%)	33 (12.3%)	17 (6.5%)	12 (4.7%)	134 (7.4%)
	Severe	63 (10.6%)	51 (22.9%)	15 (7.0%)	61 (22.8%)	27 (10.3%)	52 (20.5%)	269 (14.8%)
	Missing	24 (4.0%)	13 (5.8%)	12 (5.6%)	9 (3.4%)	24 (9.1%)	4 (1.6%)	86 (4.7%)
GCS score	Mean (SD)	13.4 (3.09)	12.2 (4.38)	13.7 (2.73)	11.9 (4.13)	13.4 (3.13)	12.4 (4.51)	12.9 (3.68)
	Median [Min, Max]	15.0 [3.00, 15.0]	15.0 [3.00, 15.0]	15.0 [3.00, 15.0]	14.0 [3.00, 15.0]	15.0 [3.00, 15.0]	15.0 [3.00, 15.0]	15.0 [3.00, 15.0]
	Missing	15 (2.5%)	1 (0.4%)	8 (3.8%)	0 (0%)	2 (0.8%)	2 (0.8%)	28 (1.5%)
Total ISS	*M* (*SD*)	17.0 (12.1)	20.5 (17.4)	13.3 (9.94)	22.8 (18.0)	17.6 (14.7)	19.4 (16.3)	18.3 (14.8)
	Mdn [*Min*, *Max*]	13.0 [1.00, 75.0]	17.0 [1.00, 75.0]	9.00 [1.00, 50.0]	18.0 [1.00, 75.0]	13.0 [1.00, 75.0]	16.0 [1.00, 75.0]	13.0 [1.00, 75.0]

[a] For categorical variables and the GOSE score, absolute (N) and relative (%) frequencies are reported. Due to rounding, the values may not sum up to exactly 100%

M, mean; SD, Standard deviation; Mdn, median; Min, minimum; Max, maximum; Injury cause (Other: summarized category for relative frequencies < 5%, including non-intentional injury; suicide attempt; violence/assault, act of mass violence); ER, emergency room; ward, admission to hospital ward; ICU, intensive care unit; GOSE, Glasgow Outcome Scale – Extended; ISS, Injury Severity Score

and Italian (12.4 ± 4.37) participants had slightly fewer (but statistically significant) years of education compared to the others. The Italian sample had a lower GCS (11.93 ± 4.13) compared to the Dutch, Finnish, and Norwegian samples. At six months after TBI, Finnish participants recovered slightly better (*Mdn* = 8; range 3–8) and had less severe extracranial injuries (ISS: 13.3 ± 9.94) compared to individuals from the Netherlands, the UK, Italy, and Spain (ISS only). However, the effects were small according to the predefined cut-offs (i.e., 0.35–0.44 or 0.56–0.63). The distributions of both four (i.e., uncomplicated mild, complicated mild, moderate, and severe) and two (i.e., mild/moderate and severe) TBI severity groups differed between language samples ($p < 0.001$).

The effect sizes represented a small effect (V = 0.13 and V = 0.17, respectively). There was no significant difference in the RPQ total score across the samples. For details, see Additional file 1: Table S2 – Comparisons of language samples regarding sociodemographic and injury-related factors.

Distribution of PCS across language samples and TBI severity groups

The distribution of PCS was similar across all language samples. *Fatigue* was the most frequently reported symptom at six months after TBI with 37% (Spanish sample) to 56% (English sample), followed by *Forgetfulness* with 36% (Finnish sample) to 46% (English sample), and *Poor concentration* with 31% (Spanish sample) to 40% (Italian sample). Individuals from the English sample tended to report more intense PCS (8 out of 16 symptoms were rated as at least a mild problem) compared to participants from other language samples (see Table 3, left part—Language samples).

Similar patterns were observed when examining TBI severity groups. Items *Fatigue* (44% and 63%), *Forgetfulness* (38% and 57%), and *Poor concentration* (34% and 44%) presented the most frequently reported symptoms in both the mild/moderate and severe TBI groups, respectively. More than one-third of individuals after severe TBI also rated the following symptoms as at least mild: prolonged thinking (42%), being frustrated (39%),

irritable (36%), and depressed (35%) (see Table 3, right part—TBI severity). For visualization, see Additional file 2: Figure S1 – Distribution of the PCS ratings in (A) each language sample and (B) for the TBI severity groups.

Analyses of response pattern

The analysis of response patterns per language sample revealed an unequal distribution of the response categories across all samples. Especially higher ranked categories (i.e., 3 "a moderate problem" and 4 "a severe problem") showed a low endorsement rate in some items. One item (*Nausea*) was not rated as a severe problem in Finnish, Italian, and Norwegian samples. The frequencies of endorsements for this item in the Dutch, English, and Spanish samples were also sparse: 1 (0.4%), 2 (0.9%), 5 (0.8%). The endorsement of the category "a moderate problem" varied from 0% (English sample) to 2.2% (Dutch sample). In addition, the item *Double vision* was rated a maximum of 2 ("a mild problem") in the Norwegian sample, resulting in no endorsement in two response categories (3 "a moderate problem" and 4 "a severe problem"). The highest endorsement rate for the category 3 "a moderate problem" or 4 "a severe problem" was observed in the Italian sample (5.2%). For more details, Additional file 1: Table S3 – Analyses of response patterns by language sample.

Table 3 Proportion of PCS rated as at least mild

Symptom	Language samples						TBI severity	
	Dutch (N = 597) (%)	English (N = 223) (%)	Finnish (N = 213) (%)	Italian (N = 268) (%)	Norwegian (N = 263) (%)	Spanish (N = 254) (%)	Mild/moderate (N = 1,521) (%)	Severe (N = 269) (%)
Headaches	27	31	**33**	27	31	**33**	30	29
Dizziness	27	28	**34**	23	28	30	28	28
Nausea	7	8	8	7	7	6	8	4
Noise sensitivity	26	22	17	19	27	24	22	27
Sleep disturbance	29	**34**	**36**	26	28	27	30	26
Fatigue	**52**	**56**	**42**	**40**	**48**	**37**	**44**	**63**
Irritable	28	**33**	24	29	27	31	27	**36**
Depressed	27	**38**	21	28	22	28	26	**35**
Frustrated	30	**43**	23	29	28	24	28	**39**
Forgetful	**41**	**46**	**36**	**45**	**38**	**37**	**38**	**57**
Poor concentration	**37**	**39**	32	**40**	**34**	31	**34**	**44**
Longer to think	**42**	**41**	25	28	**34**	22	32	**42**
Blurred vision	19	18	18	17	18	17	17	22
Light sensitivity	14	13	16	18	16	15	15	18
Double vision	10	13	11	12	6	11	9	17
Restless	24	20	19	19	21	16	20	22

Symptoms rated at least as a mild problem (i.e., ≥ 2). Bold values indicate relative frequencies over 33% (i.e., 1/3 of each language sample)

Table 4 CFA results for competitive factorial structure analyses of the RPQ across the language samples (raw data)

Factor structure	Language	X^2	df	X^2/df	p	CFI	TLI	RMSEA	CI$_{90\%}$	SRMR
One-factor structure [6]	Dutch	776.54	104	7.47	<0.001	0.988	0.986	0.105	[0.098, 0.112]	**0.075**
	English	324.13	104	3.12	<0.001	0.987	0.985	0.098	[0.086, 0.111]	0.096
	Finnish	267.05	104	2.57	<0.001	0.988	0.986	0.087	**[0.074, 0.100]**	0.088
	Italian	422.31	104	4.06	<0.001	0.979	0.975	0.107	[0.097, 0.118]	0.093
	Norwegian	228.23	104	2.19	<0.001	0.992	0.991	**0.068**	**[0.056, 0.080]**	**0.078**
	Spanish	262.18	104	2.52	<0.001	0.986	0.984	**0.078**	**[0.066, 0.090]**	0.085
Two-factor model (emotional-somatic, cognitive)[a] [19]	Dutch	549.25	89	6.17	<0.001	0.992	0.990	0.094	[0.086, 0.102]	**0.065**
	English	260.63	89	2.93	<0.001	0.989	0.988	0.094	[0.081, 0.107]	0.090
	Finnish	238.98	89	2.69	<0.001	0.988	0.986	0.090	**[0.076, 0.104]**	0.085
	Italian	324.95	89	3.65	<0.001	0.983	0.980	0.100	[0.088, 0.112]	0.084
	Norwegian	188.48	89	2.12	<0.001	0.993	0.992	**0.066**	**[0.053, 0.079]**	**0.066**
	Spanish	214.95	89	2.42	<0.001	0.989	0.987	**0.075**	[0.062, 0.088]	**0.077**
Two-factor model (RPQ-3, RPQ-13) [20]	Dutch	738.87	103	7.17	<0.001	0.989	0.987	0.103	[0.096, 0.110]	**0.071**
	English	275.66	103	2.68	<0.001	0.990	0.988	0.087	**[0.075, 0.100]**	0.087
	Finnish	241.52	103	2.34	<0.001	0.990	0.988	0.080	**[0.067, 0.094]**	0.083
	Italian	379.96	103	3.69	<0.001	0.981	0.978	0.101	[0.090, 0.111]	0.086
	Norwegian	217.73	103	2.11	<0.001	0.993	0.991	**0.065**	**[0.053, 0.078]**	**0.076**
	Spanish	242.58	103	2.36	<0.001	0.987	0.985	**0.073**	**[0.062, 0.085]**	**0.079**
Three-factor model (somatic, emotional, cognitive) [21]	Dutch	249.47	101	2.47	<0.001	0.997	0.997	**0.050**	**[0.042, 0.058]**	**0.051**
	English	189.19	101	**1.87**	<0.001	0.995	0.994	0.063	**[0.049, 0.077]**	**0.077**
	Finnish	158.61	101	**1.57**	<0.001	0.996	0.995	0.052	**[0.036, 0.068]**	**0.067**
	Italian	239.43	101	2.37	<0.001	0.991	0.989	0.072	**[0.060, 0.084]**	**0.076**
	Norwegian	128.81	101	**1.28**	0.032	0.998	0.998	0.033	**[0.010, 0.048]**	**0.067**
	Spanish	139.54	101	**1.38**	0.007	0.997	0.996	0.039	**[0.021, 0.054]**	**0.064**
Three-factor model (general somatic, mood/cognition, visual somatic) [22]	Dutch	572.62	101	5.67	<0.001	0.992	0.990	0.089	[0.082, 0.096]	**0.060**
	English	214.40	101	2.12	<0.001	0.993	0.992	0.072	**[0.058, 0.085]**	0.074
	Finnish	178.44	101	**1.77**	<0.001	0.994	0.993	0.061	**[0.046, 0.075]**	**0.071**
	Italian	265.30	101	2.63	<0.001	0.989	0.987	0.078	[0.067, 0.090]	**0.072**
	Norwegian	207.86	101	2.06	<0.001	0.993	0.992	0.064	**[0.051, 0.076]**	**0.073**
	Spanish	165.25	101	**1.64**	<0.001	0.994	0.993	**0.050**	**[0.036, 0.064]**	**0.067**

Table 4 (continued)

Factor structure	Language	χ^2	df	χ^2/df	p	CFI	TLI	RMSEA	$CI_{90\%}$	SRMR
Four-factor model (vertigo, mood/somatic, cognitive, vision) [23]	Dutch	**176.40**	98	**1.80**	<0.001	**0.999**	**0.998**	**0.037**	**[0.028, 0.046]**	**0.041**
	English	122.90	98	**1.25**	0.045	**0.999**	**0.998**	**0.034**	**[0.005, 0.052]**	**0.066**
	Finnish	113.21	98	**1.16**	**0.140**	**0.999**	**0.999**	**0.027**	**[0.000, 0.047]**	**0.057**
	Italian	157.42	98	**1.61**	<0.001	**0.996**	**0.995**	**0.048**	**[0.033, 0.061]**	**0.059**
	Norwegian	117.66	98	**1.20**	**0.086**	**0.999**	**0.998**	**0.028**	**[0.000, 0.045]**	**0.062**
	Spanish	101.95	98	**1.04**	**0.372**	**1.000**	**1.000**	**0.013**	**[0.000, 0.036]**	**0.058**

Values in bold indicate good model fit according to the respective cut-offs

[a] Estimation of the two-factor model comprising emotional/somatic and cognitive domains resulted in a not positive definite covariance matrix of the latent variables. Therefore, the results should be interpreted with caution

χ^2, chi square; df, degree of freedom; χ^2/df, ratio (cut-off: ≤2); p, p-value; CFI, Comparative Fit Index (cut-off: > 0.95); TLI, Tucker-Lewis Index (cut-off: > 0.95); RMSEA, root mean square error of approximation (cut-off: < 0.08) with 90% confidence interval (CI); SRMR, standardized root mean square residual (cut-off: < 0.08)

Zeldovich *et al. Journal of Patient-Reported Outcomes*   (2023) 7:90

Page 10 of 18

Confirmatory factor analyses (CFA)

The CFA revealed that a four-factor structure [23] comprising *vertigo*, *mood/somatic*, *cognitive*, and *vision* symptoms fitted the data best across the languages closely followed by the three-factor structure [21] including *somatic*, *emotional*, and *cognitive* symptoms (see Table 4). The estimation of the two-factor model comprising emotional-somatic and cognitive domains [19] did not converge in a proper way with covariance matrix of latent variables being not positive definite. Therefore, interpretation of goodness of fit indices of this model should be carried out with caution. Correlations between latent factors were high across all models and languages (i.e., standardized coefficients exceeded 0.65; see Additional file 1: Table S4 – Correlations between latent variables (raw data)). When using trichotomized responses of the items *Nausea* and *Double Vision*, the models revealed comparable fit across the languages (i.e., difference observed on the third decimal place; see Additional file 1: Table S5 – CFA results for competitive factorial structure analyses of the RPQ across the language samples (trichotomized items Nausea and Double Vision) for the model fit indices and Table S6 – Correlations between latent variables (trichotomized items Nausea and Double Vision) for correlations between latent factors).

When "1" responses were treated as zero, some fit indices indicated slightly better model fit across all estimated factorial solutions and languages. However, the item measuring *Nausea*, which was dichotomized because of missing responses in the higher response categories, showed high correlations (approx. r = 1.00) with items *Dizziness*, *Feeling Frustrated*, *Poor Concentration*, *Taking Longer to Think*, and *Blurred Vision*. Furthermore, two model estimations resulted in not positive definite covariance matrices: the three-factor model (general somatic, mood/cognition, and visual somatic symptoms) [22], and the firstly favorized four-factor model (vertigo, mood/somatic, cognitive, and vision) [23]. For details, see Additional file 1: Table S7 – CFA results for competitive factorial structure analyses of the RPQ across the language samples (considering "1" responses as "0" and using dichotomized items Nausea and Double Vision) for the model fit indices and Table S8 – Correlations between latent variables (considering "1" responses as "0" and using and dichotomized items Nausea and Double Vision) for correlations between latent factors).

Overall, the three-factor model comprising *somatic*, *emotional*, and *cognitive* factors [21] performed best across all competing factorial solutions in all language samples, regardless of how the "1" responses were treated as "1" or "0". Therefore, this factorial solution was chosen as a baseline model for the MI analyses.

Measurement invariance (MI)

The cross-linguistic MI analyses revealed satisfactory results (see Table 5—upper part). Except for the χ^2 *p*-values, no fit indices exceed the predefined cut-off values in the baseline model as well as in the models with increased constraints (i.e., thresholds as well as threshold and loadings model). Model comparisons were not significant. Taken together, the free-factor model did not show any violation of measurement equivalence between languages. When treating "1" responses as zero, model fit slightly increased (see Additional file 1: Table S9 – Results of MI analyses between language samples and TBI severity groups and model comparison for the three-factor model comprising somatic, emotional, and cognitive factors considering "1" responses as "0"; upper part). Therefore, this model was considered suitable for measuring PCS using the six RPQ translations.

Analyses of the TBI severity groups revealed no violation of MI assumption as reflected by non-significant difference between the models with different constraints (see Table 5—lower part). Here, again, an increase of the model fit was observed when treating "1" as "0" (see Additional file 1: Table S9 – Results of MI analyses between language samples and TBI severity groups and model comparison for the three-factor model comprising somatic, emotional, and cognitive factors considering "1" responses as "0"; lower part). These findings also support the applicability of the three-factor solution for PCS assessment using the RPQ in both examined TBI severity groups.

Final model

Estimation of the final model comprising *somatic*, *emotional*, and *cognitive* factors using raw data of the total study sample revealed satisfactory results with $\chi^2(101) = 647.04$, $\chi^2/df = 6.41$, $p < 0.001$, CFI = 0.995, TLI = 0.994, RMSEA = 0.055, $CI_{90\%}$[0.051, 0.059], SRMR = 0.051. Except for significant *p*-value and χ^2/df-ratio > 2, which can be explained by the large sample size, all other fit indices showed excellent model fit. The correlation between latent factors was high (*somatic–emotional*: 0.85; *somatic–cognitive*: 0.81; *emotional–cognitive*: 0.81). For the model visualization, see Fig. 2. When treating "1" responses as "0", the results indicated a better fit with $\chi^2(101) = 377.78$, $\chi^2/df = 3.74$, p < 0.001, CFI = 0.997, TLI = 0.996, RMSEA = 0.039, $CI_{90\%}$[0.035, 0.043], SRMR = 0.049. Again, latent factors were highly correlated (*somatic–emotional*: 0.86; *somatic–cognitive*: 0.83; *emotional–cognitive*: 0.81). For the model visualization, see Additional file 3: Figure S2 – Final model somatic (soma), emotional (emo), and cognitive (cog) factors for the total study sample when treating "1" responses as "0".

Table 5 Results of MI analyses between language samples and TBI severity groups and model comparison for the three-factor model comprising *somatic, emotional,* and *cognitive* factors using raw data [21]

Groups	Constrains	Model fit							Model comparison				
		χ^2	df	p	CFI	TLI	RMSEA	$CI_{90\%}$	$\Delta \chi^2$	Δdf	ΔCFI	ΔRMSEA	p
Language samples[a]	Baseline	1633.08	606	<0.001	**0.976**	**0.971**	**0.075**	[0.071, 0.080]	–	–	–	–	–
	Thresholds	1873.67	746	<0.001	**0.973**	**0.974**	**0.071**	[0.067, 0.075]	164.50	140	**0.003**	**0.004**	**0.077**
	Thresholds and loadings	1914.19	811	<0.001	**0.974**	**0.977**	**0.068**	[0.064, 0.071]	71.504	65	**– 0.001**	**0.003**	**0.271**
TBI severity groups (mild/moderate vs. severe)	Baseline	1028.87	202	<0.001	**0.98**	**0.977**	**0.068**	[0.064, 0.072]	–	–	–	–	–
	Thresholds	1092.79	230	<0.001	**0.979**	**0.979**	**0.065**	[0.061, 0.069]	40.216	28	**0.001**	**0.003**	**0.063**
	Thresholds and loadings	1052.98	243	<0.001	**0.981**	**0.981**	**0.061**	[0.058, 0.065]	10.587	13	**– 0.002**	**0.004**	**0.645**

Values in bold indicate good model fit according to the respective cut-offs

[a] Dutch, English, Finnish, Italian, Norwegian, Spanish

χ^2, chi square; df, degree of freedom; χ^2/df, ratio (cut-off: ≤2); p, p-value; CFI, Comparative Fit Index (cut-off: >0.95); TLI, Tucker–Lewis Index (cut-off: >0.95); RMSEA, root mean square error of approximation (cut-off: <0.08) with 90% confidence interval (CI); $\Delta \chi^2$, change in chi square values between compared models; Δdf, change in degrees of freedom between compared models; ΔCFI, change in CFI between compared models (cut-off: <0.01); ΔRMSEA, change in RMSEA between compared models (cut-off: ≤0.01)
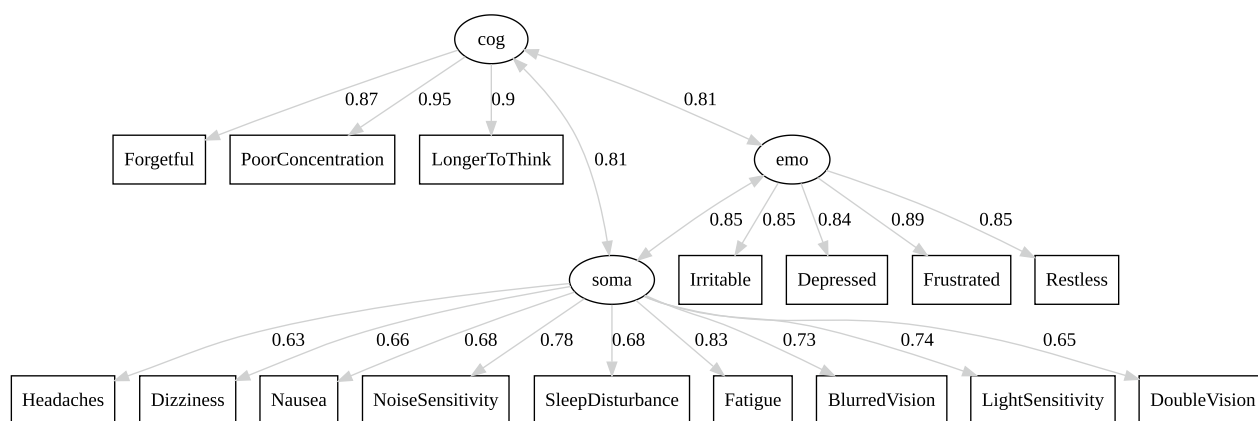
Fig. 2 Final model including somatic (soma), emotional (emo), and cognitive (cog) factors using raw data of the total study sample. The numbers depict standardized coefficients

Discussion

The present study aimed to examine the factorial validity and cross-linguistic comparability of the RPQ between six language samples. Additionally, measurement equivalence of the RPQ within TBI severity groups was investigated. The results suggest that a three-factor structure consisting of somatic, emotional, and cognitive symptom groups best captures PCS across languages. Moreover, the RPQ measures PCS equivalently across both the six language samples (i.e., Dutch, English, Finnish, Italian, Norwegian, and Spanish) and TBI severity groups (i.e., mild/moderate vs. severe). This enables national and international research on PCS and direct comparisons of outcomes across the analyzed languages within the full spectrum of TBI severity.

The RPQ has a relatively long history of attempts (2003–2018) to identify the best fitting factorial solution and thus best suitable scoring. To date, there is still no agreement as to which factorial structure would be more appropriate to assess PCS. Nevertheless, most researchers do agree on the multidimensionality of the RPQ [19–24].

Our findings show that the three-factor structure [21] including *somatic*, *emotional*, and *cognitive* scale is most appropriate for PCS assessment across six language-based samples after TBI. The favorized model is also the only one—apart from the original unidimensional factor structure—which was based on theoretical assumptions [28]. This point can also partly explain gaining problems with fitting of the models showing satisfactory results in previous studies [19, 23]. Exploratory-founded, data-driven factorial solutions may fit the data well in derivation studies but perform poorly in other datasets.

Furthermore, the scoring demonstrates clinical practicality because there are no additional constraints that may complicate the calculation of scale scores (e.g., no correlated error terms as proposed by Thomas et al. [23] to increase the model fit). In addition, two studies on factorial structure of the RPQ [19, 23], aimed in part at replicating previous scoring results, found that the three-factor model provided at least a satisfactory model fit. Potter et al. [19] found high covariance (i.e., 1.02) between the *somatic* and *emotional* latent factors, but this however was not be demonstrated in the present study (i.e., 0.85 for the total sample using raw data and 0.86 considering "1" responses as "0").

Yielding satisfactory results in one language does not provide any evidence for cross-linguistic comparability of a questionnaire. All but one study [24], which recruited Swedish-speaking participants, investigated factorial structure of the RPQ in English-speaking samples. In the present study, we observed that fit indices of the competitive factorial solutions were comparable across the languages. Since the favored three-factor structure showed empirical evidence of MI, we would recommend using this scoring in both national and international clinical and scientific investigations using the RPQ. However, from the intercorrelations between the scales, it is evident that cognitive, somatic, and emotional symptoms are not completely independent of each other. Therefore, the use of the RPQ total score can be maintained at least as a proxy for the total PCS severity rating.

In line with previous suggestions [20], we would recommend a reduction of the response categories. In particular, the response category "1—*no more of a problem than before*" contributes hardly any added value to obtain more information. The original scoring of the RPQ excludes this category from the calculation of the

total score. However, there are some pitfalls in modifying data for scoring post hoc, which is generally not recommended [44]. First, there is a difference between the number of categories presented and the number of categories used for scoring. Second, specifically in case of the RPQ, the original response scale consists of a mixture of information from the present (i.e., current symptom burden) and the past (i.e., before TBI). Although these types of scales have advantages, such as avoiding the administration of two forms of questionnaires to assess pre-TBI and post-TBI symptoms, as in the use of the Postconcussion Symptom Inventory [45], they may be particularly challenging for participants with cognitive impairments, which is likely to be the case after TBI. In addition, potential self-report or memory bias may influence response behavior in general [46] and in case of traumatic (brain) events in particular [47]. The use of this type of scale may result in inaccurate or even false information being collected. In the present study, the results of both CFAs and MI analyses using simplified response scale resulted in a better model fit. Hence, we can conclude that treating "1" as "0" may contribute to a more valid outcome assessment. However, further empirical evidence is needed before reducing the number of responses. We would suggest that future studies should address this issue by having the same group of patients complete the RPQ using different response scales (i.e., 0–4 and 0–3, where "0" could mean either "no problem at all or as before TBI" or "currently no problem"). This comparison would provide more evidence and facilitate the decision on the number of response categories of the RPQ, as has been done with other questionnaires [48].

Alternatively, future studies could address the issue of the RPQ scoring by investigating the differences between individuals choosing "0" and "1" responses, for example, using multidimensional Item Response Theory based models. Furthermore, identification of the individuals suffering from symptoms comparable to PCS prior to TBI would facilitate interpretation of the "1" responses. For example, those suffering from chronic health complains such as cancer, chronic pains, or other conditions, can suffer from fatigue, problems with concentration or sleep. This information could be considered when establishing reference or norm values for the interpretation of the results of the patients when applying the RPQ. For example, in the recent study [49] which provided reference values for the United Kingdom, the Netherlands, and Italy, one of the stratifications for the reference values was the presence the chronic health conditions which has proved important for the RPQ scores.

Strengths and limitations

The present study holds several advantages over previous investigations. First, this is the first study involving data on multiple RPQ translations which allows for a broader overview of PCS self-report in six European languages. Second, in contrast to other studies, we applied methods within the CFA framework considering the ordinal nature of the questionnaire items. Third, we additionally address the applicability of the RPQ in different TBI severity groups which had not been yet carried out.

Some limitations should be mentioned as well. Most of the sample consisted of individuals after mild TBI. Thus, those affected by moderate and severe TBI were underrepresented in this study. Therefore, the results of the MI analyses for the TBI groups should be interpreted with caution and further investigation of moderate and severe TBI regarding PCS or PC-like symptoms is highly recommended. A larger sample size of the moderate and/or severe group may result in higher test power and thus lead to more robust results.

Furthermore, there are still some difficulties in assessing PCS related to particular symptoms. The authors are aware that modification of the responses of the items *Nausea* and *Double vision* presents a potential weakness, as response behavior reflects the exhaustion of response choices and/or absence of these symptoms at six months after TBI. Interestingly, these items have already undergone some rearrangements during previous analyses of the factorial structure of the RPQ. For example, Potter et al. [19] suggested to drop the item *Double vision* from the RPQ due to severe skewness and kurtosis. Eyres et al. [20] distinguished between "acute" and "post-acute" PCS whereby the item *Double vision* was a part of the "post-acute" scale. Other authors attributed the item either to a somatic scale [21], visual somatic [22], or visual domain [23]. Lannsjö and colleagues (2011) [24] found an underrepresentation of responses in the category "severe problem" in a large mild TBI sample and the omission of this item had been suggested again.

The item *Nausea* was the one with the lowest endorsement rate across all language samples. This finding is consistent with the distinction between early and late onset PCS proposed by Ryan and Warden [50] within a mild TBI group. Moreover, Eyres et al. [20] have allocated the *Nausea* item to the "acute" symptoms using a Rasch-based approach questioning the stability of the PCS and thus the factorial structure of the RPQ over time. Since our data refer to the six-month outcome assessments, there is no information on early-onset symptoms.

Furthermore, the focus of this study was on the factorial structure and its validity, as well as the comparability of the overall PCS construct across language samples. Therefore, item-by-item comparisons using differential

Zeldovich *et al. Journal of Patient-Reported Outcomes*   (2023) 7:90

Page 14 of 18

item functioning (DIF) techniques were not conducted. Given the rigorous translation and linguistic validation process of the RPQ, which included several stages of harmonization of translations with feedback from psychologists and health professionals, translators, laypersons, and TBI patients, and item-by-item evaluation at the syntactic, cultural, idiomatic/pragmatic, and syntactic/grammatical levels, all possible linguistic issues that might arise during the translation process were addressed [10]. However, some specific problems of individual items may have been overlooked. To further strengthen the evidence for the comparability of RPQ translations, additional research involving item-level analyses is strongly encouraged.

Finally, we only took one specific point of time, i.e., six months after TBI, into account. Longitudinal analyses would provide more insight into the prevalence and persistence of PCS, and the applicability of the RPQ over time. Agtarap et al. [51] provided longitudinal analyses on PCS using a large U.S. sample of individuals after mild TBI. In Europe, a recent study using CENTER-TBI data at 3, 6, and 12 months post TBI [52] showed evidence of the applicability of the RPQ over time and the stability of the three-factor model by Smith-Seemiller et al. [21] that includes emotional, somatic, and cognitive domains.

Conclusions

Although with some limitations, the six RPQ translations were found to measure the PCS construct equally across six European languages and TBI severity groups. The three-factor model consisting of somatic, emotional, and cognitive domains showed the best fit regardless of the treatment of "1" responses. Further studies on the reduction of the RPQ response categories may provide more insight into the comparability of four- and five-point response scales. In the absence of further evidence, we recommend the use of the three-factor structure for scoring, with "1" treated as "0", in addition to the conventional total score. Finally, item-by-item comparisons between different translations of the RPQ are recommended to strengthen its aggregated applicability across languages.

Abbreviations

AIS	Abbreviated Injury Scale
CENTER-TBI	Collaborative European NeuroTrauma Effectiveness Research
CFA	Confirmatory factor analysis
CFI	Comparative fit index
CI	Confidence interval
CT	Computed tomography
DSM-IV/5	Diagnostic and the statistical manual of mental disorders
ER	Emergency room
GCS	Glasgow Coma Scale
GOSE	Glasgow outcome scale – extended
ICD-10/11	International classification of diseases
ICU	Intensive care unit
ISS	Injury severity score
MI	Measurement invariance
PROM	Patient-reported outcome measure
RMSEA	Root mean square error of approximation
RPQ	Rivermead Post-Concussion Symptoms Questionnaire
SRMR	Standardized root mean square residual
TBI	Traumatic brain injury
TLI	Tucker–Lewis index
WLSMV	Robust weighted least squares estimator

Singh[101], Toril Skandsen[123, 124], Peter Smielewski[26], Abayomi Sorinola[125], Emmanuel Stamatakis[47], Simon Stanworth[39], Robert Stevens[126], William Stewart[127], Ewout W. Steyerberg[64,128], Nino Stocchetti[129], Nina Sundström[130], Riikka Takala[131], Viktória Tamás[125], Tomas Tamosuitis[132], Mark Steven Taylor[20], Braden Te Ao[52], Olli Tenovuo[103], Alice Theadom[52], Matt Thomas[87], Dick Tibboel[133], Marjolein Timmers[74], Christos Tolias[134], Tony Trapani[28], Cristina Maria Tudora[93], Andreas Unterberg[90], Peter Vajkoczy [135], Shirley Vallance[28], Egils Valeinis[60], ZoltánVámos[50], Mathieu van der Jagt[136], Gregory Van der Steen[43], Joukje van der Naalt[71], Jeroen T.J.M. van Dijck [101], Inge A. M. van Erp[101], Thomas A. van Essen[101], Wim Van Hecke[137], Caroline van Heugten[138], Dominique Van Praag[139], Ernest van Veen[64], Thijs Vande Vyvere[137], Roel P. J. van Wijk[101], Alessia Vargiolu[32], Emmanuel Vega[83], Kimberley Velt[64], Jan Verheyden[137], Paul M. Vespa[140], Anne Vik[123, 141], Rimantas Vilcinis[132], Victor Volovici[67], Nicole von Steinbüchel[38], Daphne Voormolen[64], Petar Vulekovic[46], Kevin K. W. Wang[142], Daniel Whitehouse[47], Eveline Wiegers[64], Guy Williams[47], Lindsay Wilson[69], Stefan Winzeck[47], Stefan Wolf[143], Zhihui Yang[113], Peter Ylén[144], Alexander Younsi[90], Frederick A. Zeiler[47,145], Veronika Zelinkova[20], Agate Ziverte[60], Tommaso Zoerle[27]

[1]Department of Physiology and Pharmacology, Section of Perioperative Medicine and Intensive Care, Karolinska Institutet, Stockholm, Sweden. [2]János Szentágothai Research Centre, University of Pécs, Pécs, Hungary. [3]Division of Clinical Neuroscience, Department of Physical Medicine and Rehabilitation, Oslo University Hospital and University of Oslo, Oslo, Norway. [4]Department of Neurosurgery, University Hospital Northern Norway, Tromso, Norway. [5]Department of Physical Medicine and Rehabilitation, University Hospital Northern Norway, Tromso, Norway. [6]Trauma Surgery, Medical University Vienna, Vienna, Austria. [7]Department of Anesthesiology & Intensive Care, University Hospital Nancy, Nancy, France. [8]Raymond Poincare hospital, Assistance Publique – Hopitaux de Paris, Paris, France. [9]Department of Anesthesiology & Intensive Care, S Raffaele University Hospital, Milan, Italy. [10]Department of Neurosurgery, Radboud University Medical Center, Nijmegen, The Netherlands. [11]Department of Neurosurgery, University of Szeged, Szeged, Hungary. [12]International Projects Management, ARTTIC, Munchen, Germany. [13]Department of Neurology, Neurological Intensive Care Unit, Medical University of Innsbruck, Innsbruck, Austria. [14]Department of Neurosurgery & Anesthesia & intensive care medicine, Karolinska University Hospital, Stockholm, Sweden. [15]NIHR Surgical Reconstruction and Microbiology Research Centre, Birmingham, UK. [16]Anesthesie-Réanimation, Assistance Publique – Hopitaux de Paris, Paris, France. [17]Department of Anesthesia & ICU, AOU Città della Salute e della Scienza di Torino—Orthopedic and Trauma Center, Torino, Italy. [18]Department of Neurology, Odense University Hospital, Odense, Denmark. [19]BehaviourWorks Australia, Monash Sustainability Institute, Monash University, Victoria, Australia. [20]Department of Public Health, Faculty of Health Sciences and Social Work, Trnava University, Trnava, Slovakia. [21]Quesgen Systems Inc., Burlingame, California, USA. [22]Australian & New Zealand Intensive Care Research Centre, Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia. [23]Department of Surgery and Perioperative Science, Umeå University, Umeå, Sweden. [24]Department of Neurosurgery, Medical School, University of Pécs, Hungary and Neurotrauma Research Group, János Szentágothai Research Centre, University of Pécs, Hungary. [25]Department of Medical Psychology, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany. [26]Brain Physics Lab, Division of Neurosurgery, Dept of Clinical Neurosciences, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. [27]Neuro ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy. [28]ANZIC Research Centre, Monash University, Department of Epidemiology and Preventive Medicine, Melbourne, Victoria, Australia. [29]Department of Neurosurgery, Hospital of Cruces, Bilbao, Spain. [30]NeuroIntensive Care, Niguarda Hospital, Milan, Italy. [31]School of Medicine and Surgery, Università Milano Bicocca, Milano, Italy. [32]NeuroIntensive Care Unit, Department Neuroscience, IRCCS Fondazione San Gerardo dei Tintori, Monza, Italy. [33]Department of Neurosurgery, Medical Faculty RWTH Aachen University, Aachen, Germany. [34]Department of Anesthesiology and Intensive Care Medicine, University Hospital Bonn, Bonn, Germany. [35]Department of Anesthesia & Neurointensive Care, Cambridge University Hospital NHS Foundation Trust, Cambridge, UK. [36]School of Public Health & PM, Monash University and The Alfred Hospital, Melbourne, Victoria, Australia. [37]Radiology/MRI department, MRC Cognition and Brain Sciences Unit, Cambridge, UK. [38]Institute of Medical Psychology and Medical Sociology, Universitätsmedizin Göttingen, Göttingen, Germany. [39]Oxford University Hospitals NHS Trust, Oxford, UK. [40]Intensive Care Unit, CHU Poitiers, Potiers, France. [41]University of Manchester NIHR Biomedical Research Centre, Critical Care Directorate, Salford Royal Hospital NHS Foundation Trust, Salford, UK. [42]Movement Science Group, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK. [43]Department of Neurosurgery, Antwerp University Hospital and University of Antwerp, Edegem, Belgium. [44]Department of Anesthesia & Intensive Care, Maggiore Della Carità Hospital, Novara, Italy. [45]Department of Neurosurgery, University Hospitals Leuven, Leuven, Belgium. [46]Department of Neurosurgery, Clinical centre of Vojvodina, Faculty of Medicine, University of Novi Sad, Novi Sad, Serbia. [47]Division of Anaesthesia, University of Cambridge, Addenbrooke's Hospital, Cambridge, UK. [48]Center for Stroke Research Berlin, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany. [49]Intensive Care Unit, CHR Citadelle, Liège, Belgium. [50]Department of Anaesthesiology and Intensive Therapy, University of Pécs, Pécs, Hungary. [51]Departments of Neurology, Clinical Neurophysiology and Neuroanesthesiology, Region Hovedstaden Rigshospitalet, Copenhagen, Denmark. [52]National Institute for Stroke and Applied Neurosciences, Faculty of Health and Environmental Studies, Auckland University of Technology, Auckland, New Zealand. [53]Department of Neurology, Erasmus MC, Rotterdam, the Netherlands. [54]Department of Anesthesiology and Intensive care, University Hospital Northern Norway, Tromso, Norway. [55]Department of Neurosurgery, Hadassah-hebrew University Medical center, Jerusalem, Israel. [56]Fundación Instituto Valenciano de Neurorrehabilitación (FIVAN), Valencia, Spain. [57]Department of Neurosurgery, Shanghai Renji hospital, Shanghai Jiaotong University/school of medicine, Shanghai, China. [58]Karolinska Institutet, INCF International Neuroinformatics Coordinating Facility, Stockholm, Sweden. [59]Emergency Department, CHU, Liège, Belgium. [60]Neurosurgery clinic, Pauls Stradins Clinical University Hospital, Riga, Latvia. [61]Department of Computing, Imperial College London, London, UK. [62]Department of Neurosurgery, Hospital Universitario 12 de Octubre, Madrid, Spain. [63]Department of Anesthesia, Critical Care and Pain Medicine, Medical University of Vienna, Austria. [64]Department of Public Health, Erasmus Medical Center-University Medical Center, Rotterdam, The Netherlands. [65]College of Health and Medicine, Australian National University, Canberra, Australia. [66]Department of Neurosurgery, Neurosciences Centre & JPN Apex trauma centre, All India Institute of Medical Sciences, New Delhi-110029, India.[67]Department of Neurosurgery, Erasmus MC, Rotterdam, the Netherlands. [68]Department of Neurosurgery, Oslo University Hospital, Oslo, Norway. [69]Division of Psychology, University of Stirling, Stirling, UK. [70]Division of Neurosurgery, Department of Clinical Neurosciences, Addenbrooke's Hospital & University of Cambridge, Cambridge, UK. [71]Department of Neurology, University of Groningen, University Medical Center Groningen, Groningen, Netherlands. [72]Neurointensive Care, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK. [73]Salford Royal Hospital NHS Foundation Trust Acute Research Delivery Team, Salford, UK. [74]Department of Intensive Care and Department of Ethics and Philosophy of Medicine, Erasmus Medical Center, Rotterdam, The Netherlands. [75]Department of Clinical Neuroscience, Neurosurgery, Umeå University, Umeå, Sweden. [76]Hungarian Brain Research Program—Grant No. KTIA_13_NAP-A-II/8, University of Pécs, Pécs, Hungary. [77]Department of Anaesthesiology, University Hospital of Aachen, Aachen, Germany. [78]Cyclotron Research Center, University of Liège, Liège, Belgium. [79]Centre for Urgent and Emergency Care Research (CURE), Health Services Research Section, School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, UK. [80]Emergency Department, Salford Royal Hospital, Salford UK. [81]Institute of Research in Operative Medicine (IFOM), Witten/ Herdecke University, Cologne, Germany. [82]VP Global Project Management CNS, ICON, Paris, France. [83]Department of Anesthesiology-Intensive Care, Lille University Hospital, Lille, France. [84]Department of Neurosurgery, Rambam Medical Center, Haifa, Israel. [85]Department of Anesthesiology & Intensive Care, University Hospitals Southhampton NHS Trust, Southhampton, UK. [86]Cologne-Merheim Medical Center (CMMC), Department of Traumatology, Orthopedic Surgery and Sportmedicine, Witten/Herdecke University, Cologne, Germany. [87]Intensive Care Unit, Southmead Hospital, Bristol, Bristol, UK. [88]Department of Neurological Surgery, University of California, San Francisco, California, USA. [89]Department of Anesthesia & Intensive Care,M. Bufalini Hospital, Cesena, Italy. [90]Department of Neurosurgery, University Hospital Heidelberg, Heidelberg, Germany. [91]Department of Neurosurgery, The Walton centre NHS Foundation Trust, Liverpool, UK. [92]Department of Medical Genetics, University of Pécs, Pécs, Hungary. [93]Department of Neurosurgery, Emergency County Hospital Timisoara, Timisoara, Romania. [94]School of Medical Sciences, Örebro University, Örebro, Sweden. [95]Institute for Molecular

Zeldovich *et al. Journal of Patient-Reported Outcomes*   (2023) 7:90

Page 16 of 18

Medicine Finland, University of Helsinki, Helsinki, Finland. [96]Analytic and Translational Genetics Unit, Department of Medicine; Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry; Department of Neurology, Massachusetts General Hospital, Boston, MA, USA. [97]Program in Medical and Population Genetics; The Stanley Center for Psychiatric Research, The Broad Institute of MIT and Harvard, Cambridge, MA, USA. [98]Department of Radiology, University of Antwerp, Edegem, Belgium. [99]Department of Anesthesiology & Intensive Care, University Hospital of Grenoble, Grenoble, France. [100]Department of Anesthesia & Intensive Care, Azienda Ospedaliera Università di Padova, Padova, Italy. [101]Department of Neurosurgery, Leiden University Medical Center, Leiden, The Netherlands and Department of Neurosurgery, Medical Center Haaglanden, The Hague, The Netherlands. [102]Department of Neurosurgery, Helsinki University Central Hospital. [103]Division of Clinical Neurosciences, Department of Neurosurgery and Turku Brain Injury Centre, Turku University Hospital and University of Turku, Turku, Finland. [104]Department of Anesthesiology and Critical Care, Pitié -Salpêtrière Teaching Hospital, Assistance Publique, Hôpitaux de Paris and University Pierre et Marie Curie, Paris, France. [105]Neurotraumatology and Neurosurgery Research Unit (UNINN), Vall d'Hebron Research Institute, Barcelona, Spain. [106]Department of Neurosurgery, Kaunas University of technology and Vilnius University, Vilnius, Lithuania. [107]Department of Neurosurgery, Rezekne Hospital, Latvia. [108]Department of Anaesthesia, Critical Care and Pain Medicine NHS Lothian and University of Edinburg, Edinburgh, UK. [109]Director, MRC Biostatistics Unit, Cambridge Institute of Public Health, Cambridge, UK. [110]Department of Physical Medicine and Rehabilitation, Oslo University Hospital/University of Oslo, Oslo, Norway. [111]Division of Orthopedics, Oslo University Hospital, Oslo, Norway. [112]Institue of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [113]Broad Institute, Cambridge MA Harvard Medical School, Boston MA, Massachusetts General Hospital, Boston MA, USA. [114]National Trauma Research Institute, The Alfred Hospital, Monash University, Melbourne, Victoria, Australia. [115]Department of Neurosurgery, Odense University Hospital, Odense, Denmark. [116]International Neurotrauma Research Organisation, Vienna, Austria. [117]Klinik für Neurochirurgie, Klinikum Ludwigsburg, Ludwigsburg, Germany. [118]Division of Biostatistics and Epidemiology, Department of Preventive Medicine, University of Debrecen, Debrecen, Hungary. [119]Department Health and Prevention, University Greifswald, Greifswald, Germany. [120]Department of Anaesthesiology and Intensive Care, AUVA Trauma Hospital, Salzburg, Austria. [121]Department of Neurology, Elisabeth-TweeSteden Ziekenhuis, Tilburg, the Netherlands. [122]Department of Neuroanesthesia and Neurointensive Care, Odense University Hospital, Odense, Denmark. [123]Department of Neuromedicine and Movement Science, Norwegian University of Science and Technology, NTNU, Trondheim, Norway. [124]Department of Physical Medicine and Rehabilitation, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [125]Department of Neurosurgery, University of Pécs, Pécs, Hungary. [126]Division of Neuroscience Critical Care, John Hopkins University School of Medicine, Baltimore, USA. [127]Department of Neuropathology, Queen Elizabeth University Hospital and University of Glasgow, Glasgow, UK. [128]Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands. [129]Department of Pathophysiology and Transplantation, Milan University, and Neuroscience ICU, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milano, Italy. [130]Department of Radiation Sciences, Biomedical Engineering, Umeå University, Umeå, Sweden. [131]Perioperative Services, Intensive Care Medicine and Pain Management, Turku University Hospital and University of Turku, Turku, Finland. [132]Department of Neurosurgery, Kaunas University of Health Sciences, Kaunas, Lithuania. [133]Intensive Care and Department of Pediatric Surgery, Erasmus Medical Center, Sophia Children's Hospital, Rotterdam, The Netherlands. [134]Department of Neurosurgery, Kings college London, London, UK. [135]Neurologie, Neurochirurgie und Psychiatrie, Charité – Universitätsmedizin Berlin, Berlin, Germany. [136]Department of Intensive Care Adults, Erasmus MC–University Medical Center Rotterdam, Rotterdam, the Netherlands. [137]icoMetrix NV, Leuven, Belgium. [138]Movement Science Group, Faculty of Health and Life Sciences, Oxford Brookes University, Oxford, UK. [139]Psychology Department, Antwerp University Hospital, Edegem, Belgium. [140]Director of Neurocritical Care, University of California, Los Angeles, USA. [141]Department of Neurosurgery, St.Olavs Hospital, Trondheim University Hospital, Trondheim, Norway. [142]Department of Emergency Medicine, University of Florida, Gainesville, Florida, USA. [143]Department of Neurosurgery, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health, Berlin, Germany. [144]VTT Technical Research Centre, Tampere, Finland. [145]Section of Neurosurgery, Department of Surgery, Rady Faculty of Health Sciences, University of Manitoba, Winnipeg, MB, Canada.

Author contributions
Conceptualization: MZ, FB, and NS. Methodology: MZ. Software: MZ. Formal analysis: MZ. Data curation: MZ. Writing – original draft preparation: MZ. Writing – review and editing: FB, AC, IM, SP, AM, MV, and NS. Visualization: MZ. Supervision: NS. Funding acquisition: SP and NS.

Availability of data and materials
All relevant data are available upon request from CENTER-TBI, and the authors are not legally allowed to share it publicly. The authors confirm that they received no special access privileges to the data. CENTER-TBI is committed to data sharing and in particular to responsible further use of the data. Hereto, we have a data sharing statement in place: https://www.center-tbi.eu/data/sharing (last access on 18 April 2023). The CENTER-TBI Management Committee, in collaboration with the General Assembly, established the Data Sharing policy, and Publication and Authorship Guidelines to assure correct and appropriate use of the data as the dataset is hugely complex and requires help of experts from the Data Curation Team or Bio- Statistical Team for correct use. This means that we encourage researchers to contact the CENTER-TBI team for any research plans and the Data Curation Team for any help in appropriate use of the data, including sharing of scripts. Requests for data access can be submitted online: https://www.center-tbi.eu/data (last access on 18 April 2023). The complete Manual for data access is also available online: https://www.center-tbi.eu/files/SOP-Manual-DAPR-2402020.pdf (assessed on 18 April 2023).

Declarations

Ethics approval and consent to participate
The CENTER-TBI study (EC grant 602150) has been conducted in accordance with all relevant laws of the EU if directly applicable or of direct effect and all relevant laws of the country where the recruiting sites were located, including but not limited to, the relevant privacy and data protection laws and regulations (the "Privacy Law"), the relevant laws and regulations on the use of human materials, and all relevant guidance relating to clinical studies from time to time in force including, but not limited to, the ICH Harmonized Tripartite Guideline for Good Clinical Practice (CPMP/ICH/135/95) ("ICH GCP") and the World Medical Association Declaration of Helsinki entitled "Ethical Principles for Medical Research Involving Human Subjects". Written informed consent was obtained for all patients recruited in the Core Dataset of CENTER-TBI and documented in the e-CRF. Ethical approval was obtained for each recruiting site. The list of sites, Ethical Committees, approval numbers and approval dates can be found on the project's website https://www.center-tbi.eu/project/ethical-approval (assessed 18th on April 2023).

Consent for publication
Not applicable.

Competing interests
The authors declare no conflict of interest.

Author details
[1]Institute of Medical Psychology and Medical Sociology, University Medical Center Göttingen, Göttingen, Germany. [2]Department of Psychiatry, Columbia University Medical Center, New York, NY, USA. [3]Department of Public Health, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [4]Department of Psychology, The University of British Columbia, Vancouver, BC, Canada. [5]Rehabilitation Research Program, Centre for Aging SMART at Vancouver Coastal Health, Vancouver, BC, Canada.

Zeldovich *et al. Journal of Patient-Reported Outcomes*   (2023) 7:90

Page 17 of 18

References

1. Menon DK, Schwab K, Wright DW, Maas AI (2010) Position statement: definition of traumatic brain injury. Arch Phys Med Rehabil 91:1637–1640. https://doi.org/10.1016/j.apmr.2010.05.017
2. Maas AIR, Menon DK, Adelson PD et al (2017) Traumatic brain injury: integrated approaches to improve prevention, clinical care, and research. Lancet Neurol 16:987–1048. https://doi.org/10.1016/S1474-4422(17)30371-X
3. Powell JM, Wise EK, Brockway JA et al (2017) Characteristics and concerns of caregivers of adults with traumatic brain injury. J Head Trauma Rehabil 32:E33–E41. https://doi.org/10.1097/HTR.0000000000000219
4. Hiploylee C, Dufort PA, Davis HS et al (2017) Longitudinal study of postconcussion syndrome: not everyone recovers. J Neurotrauma 34:1511–1523. https://doi.org/10.1089/neu.2016.4677
5. World Health Organization (2004) ICD-10: international statistical classification of diseases and related health problems, 10th revision, 2nd ed.
6. King NS, Crawford S, Wenden FJ et al (1995) The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms commonly experienced after head injury and its reliability. J Neurol 242:587–592
7. Sigurdardottir S, Andelic N, Roe C et al (2009) Post-concussion symptoms after traumatic brain injury at 3 and 12 months post-injury: a prospective study. Brain Inj 23:489–497. https://doi.org/10.1080/02699050902926309
8. Zeldovich M, Wu Y-J, Gorbunova A et al (2020) Influence of sociodemographic, premorbid, and injury-related factors on post-concussion symptoms after traumatic brain injury. JCM 9:1931. https://doi.org/10.3390/jcm9061931
9. Maas AIR, Menon DK, Steyerberg EW et al (2015) Collaborative European NeuroTrauma Effectiveness Research in traumatic brain injury (CENTER-TBI): a prospective longitudinal observational study. Neurosurgery 76:67–80. https://doi.org/10.1227/NEU.0000000000000575
10. von Steinbuechel N, Rauen K, Krenz U et al (2021) Translation and linguistic validation of outcome instruments for traumatic brain injury research and clinical practice: a step-by-step approach within the observational CENTER-TBI study. J Clin Med 10:2863
11. Acquadro C (2012) Linguistic validation manual for health outcome assessments. MAPI Research Institute, Lyon
12. Meredith W, Teresi JA (2006) An essay on measurement and factorial invariance. Med Care 44:S69–S77. https://doi.org/10.1097/01.mlr.0000245438.73837.89
13. von Steinbuechel N, Rauen K, Bockhop F et al (2021) Psychometric characteristics of the patient-reported outcome measures applied in the CENTER-TBI Study. JCM 10:2396. https://doi.org/10.3390/jcm10112396
14. Steyerberg EW, Wiegers E, Sewalt C et al (2019) Case-mix, care pathways, and outcomes in patients with traumatic brain injury in CENTER-TBI: a European prospective, multicentre, longitudinal, cohort study. Lancet Neurol 18:923–934. https://doi.org/10.1016/S1474-4422(19)30232-7
15. Teasdale G, Jennett B (1974) Assessment of coma and impaired consciousness. A practical scale. Lancet 2:81–84. https://doi.org/10.1016/s0140-6736(74)91639-0
16. Williams DH, Levin HS, Eisenberg HM (1990) Mild head injury classification. Neurosurgery. https://doi.org/10.1097/00006123-199009000-00014
17. Wilson JTL, Pettigrew LEL, Teasdale G (1998) Structured interviews for the glasgow outcome scale and the extended Glasgow Outcome Scale: guidelines for their use. J Neurotrauma 15:573–585. https://doi.org/10.1089/neu.1998.15.573
18. Gennarelli TA, Wodzin E, Association for the Advancement of Automotive Medicine (2008) Abbreviated injury scale 2005: update 2008. Association for the Advancement of Automative Medicine, Barrington
19. Potter S, Leigh E, Wade D, Fleminger S (2006) The Rivermead Post Concussion Symptoms Questionnaire: a confirmatory factor analysis. J Neurol 253:1603–1614. https://doi.org/10.1007/s00415-006-0275-z
20. Eyres S, Carey A, Gilworth G et al (2005) Construct validity and reliability of the Rivermead Post-Concussion Symptoms Questionnaire. Clin Rehabil 19:878–887. https://doi.org/10.1191/0269215505cr905oa
21. Smith-Seemiller L, Fow NR, Kant R, Franzen MD (2003) Presence of postconcussion syndrome symptoms in patients with chronic pain vs mild traumatic brain injury. Brain Inj 17:199–206. https://doi.org/10.1080/0269905021000030823
22. Herrmann N, Rapoport MJ, Rajaram RD et al (2009) Factor analysis of the Rivermead Post-Concussion Symptoms Questionnaire in mild-to-moderate traumatic brain injury patients. JNP 21:181–188. https://doi.org/10.1176/jnp.2009.21.2.181
23. Thomas M, Skilbeck C, Cannan P, Slatyer M (2018) The structure of the Rivermead Post-Concussion Symptoms Questionnaire in Australian adults with traumatic brain injury. Brain Impairment 19:166–182. https://doi.org/10.1017/BrImp.2017.26
24. Lannsjö M, Borg J, Björklund G et al (2011) Internal construct validity of the Rivermead Post-Concussion Symptoms Questionnaire. J Rehabil Med 43:997–1002. https://doi.org/10.2340/16501977-0875
25. Vargha A, Delaney HD (2000) A critique and improvement of the *CL* common language effect size statistics of McGraw and Wong. J Educ Behav Stat 25:101–132. https://doi.org/10.3102/10769986025002101
26. Cohen J (1988) Statistical power analysis for the behavioural sciences, 2nd edn. L. Erlbaum Associates, Hillsdale
27. Brown TA (2015) Confirmatory factor analysis for applied research, 2nd edn. The Guilford Press, London
28. Gerber DJ, Schraa JC (1995) Mild traumatic brain injury: searching for the syndrome. J Head Trauma Rehabil 10:28–40. https://doi.org/10.1097/00001199-199508000-00004
29. Bentler PM (1990) Comparative fit indexes in structural models. Psychol Bull 107:238–246
30. Bentler PM, Bonett DG (1980) Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull 88:588–606. https://doi.org/10.1037/0033-2909.88.3.588
31. Steiger JH (1980) Statistically based tests for the number of common factors. In: The annual meeting of the Psychometric Society. Iowa City
32. Cole DA (1987) Utility of confirmatory factor analysis in test validation research. J Consult Clin Psychol 55:584–594. https://doi.org/10.1037/0022-006X.55.4.584
33. Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Model 6:1–55
34. Finch WH, French BF (2018) A simulation investigation of the performance of invariance assessment using equivalence testing procedures. Struct Equ Model 25:673–686. https://doi.org/10.1080/10705511.2018.1431781
35. Xia Y, Yang Y (2019) RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: the story they tell depends on the estimation methods. Behav Res 51:409–428. https://doi.org/10.3758/s13428-018-1055-2
36. Wu H, Estabrook R (2016) Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. Psychometrika 81:1014–1045. https://doi.org/10.1007/s11336-016-9506-0
37. Svetina D, Rutkowski L, Rutkowski D (2020) Multiple-group invariance with categorical outcomes using updated guidelines: an illustration using mplus and the Lavaan/semtools packages. Struct Equ Model 27:111–130. https://doi.org/10.1080/10705511.2019.1602776
38. Cheung GW, Rensvold RB (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. Struct Equ Model 9:233–255
39. Chen FF (2007) Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct Equ Model 14:464–504. https://doi.org/10.1080/10705510701301834
40. R Core Team (2020) R: a language and environment for statistical computing
41. Rich B (2021) Table1: tables of descriptive statistics in HTML
42. Rosseel Y (2012) lavaan: An R package for structural equation modeling. J Stat Soft 48:5. https://doi.org/10.18637/jss.v048.i02
43. Lishinski A (2021) lavaanPlot: Path Diagrams for "Lavaan" Models via "DiagrammeR". R package version 0.6.2.
44. Van Dusen B, Nissen JM (2020) Criteria for collapsing rating scale responses: a case study of the CLASS. In: 2019 physics education research conference proceedings. American Association of Physics Teachers, Provo, UT
45. Sady MD, Vaughan CG, Gioia GA (2014) Psychometric characteristics of the postconcussion symptom inventory in children and adolescents. Arch Clin Neuropsychol 29:348–363. https://doi.org/10.1093/arclin/acu014

Zeldovich *et al. Journal of Patient-Reported Outcomes*   *(2023) 7:90*

Page 18 of 18

46. Althubaiti A (2016) Information bias in health research: definition, pitfalls, and adjustment methods. JMDH. https://doi.org/10.2147/JMDH.S104807
47. Lalande KM, Bonanno GA (2011) Retrospective memory bias for the frequency of potentially traumatic events: a prospective study. Psychol Trauma Theory Res Pract Policy 3:165–170. https://doi.org/10.1037/a0020847
48. Colvin KF, Gorgun G (2020) Collapsing scale categories: comparing the psychometric properties of resulting scales. Pract Assess Res 25:6
49. Zeldovich M, Bockhop F, Covic A et al (2022) Reference values for the Rivermead Post-Concussion Symptoms Questionnaire (RPQ) from general population samples in the United Kingdom, Italy, and The Netherlands. JCM 11:4658. https://doi.org/10.3390/jcm11164658
50. Ryan LM, Warden DL (2003) Post concussion syndrome. Int Rev Psychiatry 15:310–316. https://doi.org/10.1080/09540260310001606692
51. Agtarap S, Kramer MD, Campbell-Sills L et al (2021) Invariance of the Bifactor Structure of Mild Traumatic Brain Injury (mTBI) Symptoms on the Rivermead Postconcussion Symptoms Questionnaire across time, demographic characteristics, and clinical groups: a TRACK-TBI study. Assessment 28:1656–1670. https://doi.org/10.1177/1073191120913941
52. Rivera D, Greving S, Arango-Lasprilla JC (2022) Comparability of (post-concussion) symptoms across time in individuals after traumatic brain injury: Results from the center-tbi study. J Clin Med 11:4090

Publisher's Note