ScienceDirect

Contents lists available at sciencedirect.com
Journal homepage: www.elsevier.com/locate/jval

Systematic Literature Review

Psychometric Properties of Cognition Bolt-Ons for the EQ-5D-3L and EQ-5D-5L: A Systematic Review

Fanni Rencz, DSc, Stevanus Pangestu, MBA, Brendan Mulhern, PhD, Aureliano Paolo Finch, PhD, Mathieu F. Janssen, PhD

ABSTRACT

Objectives: Cognition is the most commonly used EQ-5D bolt-on, with many different versions varying by descriptors and response levels (3L vs 5L). We aimed to systematically review the psychometric properties of cognition bolt-ons for the EQ-5D-3L and EQ-5D-5L.

Methods: A systematic review was conducted in PubMed, Web of Science, and Google Scholar following PRISMA 2020 guidelines (PROSPERO:CRD42023445567). We assessed the bolt-ons' performance both as individual items and when added to the EQ-5D. Each bolt-on version was rated as positive (+) or nonpositive (−) within each publication using a checklist, and scores were summed across publications to reflect overall performance.

Results: In total, 101 publications from 72 studies met the inclusion criteria, examining 15 3-level and 13 5-level bolt-ons. The most frequently reported psychometric properties were item-level ceiling (n = 75) and known-groups validity (n = 54). Fewer studies explored convergent or divergent validity (n = 8 for each), responsiveness (n = 3), patient-proxy agreement (n = 2), and test-retest reliability (n = 1). None reported on content validity. Five-level bolt-ons outperformed 3-level bolt-ons in terms of overall performance (3L: 55+/57−; 5L: 45+/28−). Supportive psychometric evidence varied by populations, eg, head/brain injury (3L: 11+/11−; 5L: 1+/3−) and dementia (3L: 9+/8−; 5L: 4+/4−). The most-tested bolt-ons were the Janssen 2013 (5L, cognition: 18+/15−) and Haagsma 2005 (3L, thinking ability: 8+/12−) versions, with fewer than 10 assessments for all other bolt-ons.

Conclusions: Despite several publications, the psychometric evidence remains insufficient to identify a preferred cognition descriptor. Future research should prioritize content validity testing to inform the selection of candidate items, with quantitative psychometric evaluation preferably conducted afterward.

Keywords: bolt-on, cognition, EQ-5D, measurement properties, psychometric, systematic literature review.

VALUE HEALTH. 2025; ■(■):■–■

Highlights

- No systematic reviews have been conducted to summarize all psychometric evidence on EQ-5D cognition bolt-ons.
- This systematic review synthesized evidence on the psychometric performance of cognition bolt-ons for the EQ-5D-3L and EQ-5D-5L, based on more than 100 publications. Several bolt-ons demonstrated good divergent and known-groups validity for relevant clinical groups. However, the psychometric evidence base remains limited for most bolt-on items and across the majority of psychometric properties. The lack of content validity evidence particularly limits the identi
fication of the best-performing items.
- The review provides extensive information to support users of cognition bolt-ons in understanding their psychometric performance. Further assessments are needed to compare the performance of candidate items, guide item selection, and assess the potential need for further development work.

Introduction

The EQ-5D is the most widely used generic preference-accompanied measure of health-related quality of life (HRQoL).[1,2] It is recommended by numerous pharmacoeconomic guidelines and health technology assessment bodies worldwide.[3,4] It also serves as a patient-reported outcome measure to assess HRQoL in both general populations and patient groups across diverse contexts.[5,6] The EQ-5D is the shortest of the existing generic preference-accompanied measures, consisting of 5 single-item dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Although these 5 dimensions are broadly relevant across most populations and interventions, they may not fully capture all important aspects of HRQoL in some specific populations.[7,8] To address areas not included in the EQ-5D, an increasing number of studies have proposed "bolt-ons"—additional dimensions developed to extend the EQ-5D descriptive system.[9]

Similar to any HRQoL measure, psychometric testing of bolt-ons across different target populations, health interventions, languages, modes of administration (eg, self-completion or interviewer-administered) and versions (eg, self-report or proxy-report) is necessary to ensure that the bolt-on is valid, reliable, and sensitive to changes in HRQoL for the intended future applications. Because bolt-ons supplement an existing instrument, their testing typically focuses on the improvement in measurement performance achieved by their addition.[10-23] Recent guidance on the development, selection, and testing of bolt-ons has recommended using mixed-methods research that combines both qualitative and quantitative approaches in the psychometric testing of bolt-ons.[24]

To date, bolt-ons have been proposed for more than 50 constructs. Cognition is one of the earliest and most frequently used bolt-ons,[9,25] with 181 publications and 52 different wordings for this dimension.[26] A comprehensive evaluation of the psychometric properties of each bolt-on is essential to guide decisions regarding the selection of the best-performing bolt-ons. Demonstrating strong psychometric properties is also a key

criterion for bolt-ons to be accepted by pharmacoeconomic or regulatory guidelines set by relevant bodies,[27,28] and it can also influence future adoption and uptake by different user groups. Therefore, we aimed to systematically review the literature on the psychometric properties of cognition bolt-ons for the EQ-5D-3L and EQ-5D-5L.

Methods

Study Protocol

This systematic review is the second part of a 2-part systematic review on cognition bolt-ons for the EQ-5D (PROSPERO reg. no. CRD42023445567). The first part of the review explored the development and use of cognition bolt-ons,[26] whereas this second part examines their psychometric properties using a subset of publications from the first part. Detailed descriptions of data sources, search strategy, and study selection are provided in the first part.[26] Here, we briefly summarize the methods and note any deviations from or additions to the first part.

Study Selection, Inclusion Criteria

A literature search was conducted in 2023 using PubMed, Web of Science, Google Scholar, and EuroQol conference proceedings.[26] Study selection adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines.[29] Two independent reviewers (F.R. and S.P.) performed the selection (based on inclusion and exclusion criteria), data extraction, and psychometric assessment. The inclusion criteria were as follows:

- Publication type: original articles (excluding published study protocols), short communications, research letters, conference papers (excluding posters), master's or doctoral theses, and published working articles or research reports.
- Publication language: any language.
- Study type: qualitative studies, cross-sectional studies, cohort studies, case-control studies, or clinical trials.
- Study population: Studies involving participants aged ≥16 years.
- Sample size for quantitative studies: ≥30 (based on the central limit theorem).[30]
- Intervention/exposure: any type of intervention or exposure was eligible.
- Instrument: EQ-5D-3L or EQ-5D-5L with any cognition bolt-on(s). (Note: cognition bolt-ons not following the format of the EQ-5D, such as items of other instruments were not considered).
- Psychometric properties: studies that reported on at least 1 psychometric property of the cognition bolt-on or the EQ-5D+cognition bolt-on.

Multiple publications from the same data set were included if they presented results on different psychometric properties or for different subsets of respondents. In the summary table of study characteristics, we report the results separately for "studies" and "publications" to account for multiple publications arising from the same study, with the publication that has the largest sample size considered the primary representation of the study.

Data Extraction and Strategy for Data Synthesis

The following data were extracted from eligible publications:

- Publication details: authors, year of publication, and publication language.
- Study characteristics: study design, sample size, country, and language.
- Characteristics of the study population: age and diagnosis.
- Completion: self-completed versus interviewer-administered.
- Reporting method: self-report versus proxy (including perspective: proxy-proxy or proxy-patient).
- Characteristics of the cognition bolt-on: version used (3-level [3L] or 5-level [5L]), number of bolt-on items used, precise wording of the bolt-on item(s), any modifications used (eg, recall period or pictorial version).
- Psychometric methodology: qualitative, quantitative, and mixed methods.
- Psychometric results (for qualitative studies): key findings, including relevant quotes.
- Psychometric results (for quantitative studies; bolt-on item(s)): descriptive statistics of bolt-on responses, eg, response distribution on bolt-on item(s) and also the core EQ-5D items, floor, ceiling, divergent, convergent, known-groups and structural validity, correlation with EQ VAS, responsiveness, test-retest reliability, inter-rater (eg, patient-proxy) agreement, and differential item functioning.
- Psychometric results (for quantitative studies; EQ-5D+bolt-on item(s)): floor, ceiling, absolute and relative informativity, divergent, convergent and known-groups validity, explanatory power on EQ VAS in a multivariable regression, responsiveness, test-retest reliability, and interrater agreement.

Floor, ceiling, and absolute and relative informativity at the instrument level (EQ-5D+bolt-on(s)) were included only if reported for both EQ-5D and EQ-5D+bolt-on because results are not interpretable otherwise. External measures for which correlations were reported were individually assessed based on their content and the specific bolt-on item tested, to determine whether they were appropriate for evaluating convergent or divergent validity. Correlations with all external measures were reported for completeness, regardless of their later consideration in the psychometric assessment using the checklist. Explanatory power on the EQ VAS was considered only when assessed in a multivariable model that included all five EQ-5D items and the bolt-on. Internal consistency reliability was not extracted when reported because it is deemed inappropriate for EQ-5D(+bolt-on) because of the assumption of independence across dimensions. For longitudinal studies in which a psychometric property was assessed at multiple time points, we extracted the relevant data for each time point. Patient populations were classified according to the International Classification of Diseases, 11th Revision (ICD-11).[31]

Assessment of Psychometric Properties

We extracted only numerical results from the included studies, excluding the authors' interpretations (eg, rating a certain psychometric finding as poor or good), and applied our own checklist to minimize bias. We developed a psychometric checklist based on the 8 bolt-on testing criteria proposed by Mulhern et al, 2022.[24] The following considerations prompted us

Table 1. Characteristics summary of the included publications and studies.

Characteristic	Publications (k = 101)		Studies (*n* = 72)	
	k	%	*n*	%
Publication language				
English	94	93.1%	66	91.7%
Dutch	3	3.0%	2	2.8%
Mandarin Chinese	3	3.0%	3	4.2%
Japanese	1	1.0%	1	1.4%
Publication type				
Journal article	95	94.1%	66	91.7%
Thesis	3	3.0%	3	4.2%
EuroQol conference paper	1	1.0%	1	1.4%
Other research reports	2	2.0%	2	2.8%
Country*				
The Netherlands	63	62.4%	44	61.1%
New Zealand	10	9.9%	3	4.2%
Australia	4	4.0%	2	2.8%
Belgium	4	4.0%	3	4.2%
China	4	4.0%	4	5.6%
Germany	3	3.0%	2	2.8%
India	2	2.0%	2	2.8%
Japan	2	2.0%	2	2.8%
South Africa	2	2.0%	2	2.8%
Zimbabwe	2	2.0%	2	2.8%
Other countries†	8	7.9%	8	11.1%
Language of instrument*				
Dutch	65	64.4%	46	63.9%
English	19	18.8%	10	13.9%
Mandarin Chinese	4	4.0%	4	5.6%
French	3	3.0%	3	4.2%
German	3	3.0%	2	2.8%
Japanese	2	2.0%	2	2.8%
Other languages‡	11	10.9%	9	12.5%
Not reported	2	2.0%	2	2.8%
Study design*				
Cross-sectional	38	37.6%	29	40.3%
Longitudinal cohort	53	52.5%	33	45.8%
Randomized controlled trial	6	5.9%	6	8.3%
Case control	2	2.0%	2	2.8%
Other study designs§	3	3.0%	3	4.2%
Sample size				
30-49	4	4.0%	2	2.8%
50-99	11	10.9%	8	11.1%
100-199	21	20.8%	16	22.2%
200-299	12	11.9%	9	12.5%
300-399	12	11.9%	10	13.9%
400-499	6	5.9%	4	5.6%
500-999	12	11.9%	9	12.5%
1000 and above	23	22.8%	14	19.4%

continued on next page

Table 1. Continued

Characteristic	Publications (k = 101)		Studies (*n* = 72)	
	k	%	*n*	%
Study population*				
Patient	83	82.2%	57	79.2%
General population (15+)	18	17.8%	15	20.8%
General population: solely older adults (55+)	8	7.9%	7	9.7%
Caregivers	6	5.9%	5	6.9%
Patient population (ICD-11 categorization)*				
01 Certain infectious or parasitic diseases	10	9.9%	8	11.1%
02 Neoplasms	3	3.0%	2	2.8%
04 Immune system	1	1.0%	1	1.4%
05 Endocrine, nutritional or metabolic	2	2.0%	2	2.8%
06 Mental, behavioral, or neurodevelopmental	15	14.9%	14	19.4%
08 Nervous system	6	5.9%	4	5.6%
09 Visual system	1	1.0%	1	1.4%
10 Ear or mastoid process	1	1.0%	1	1.4%
11 Circulatory system	4	4.0%	3	4.2%
12 Respiratory system	1	1.0%	1	1.4%
13 Digestive system	6	5.9%	3	4.2%
14 Skin	1	1.0%	1	1.4%
15 Musculoskeletal system or connective tissue	1	1.0%	1	1.4%
16 Genitourinary system	2	2.0%	2	2.8%
18 Pregnancy, childbirth or the puerperium	3	3.0%	3	4.2%
21 Symptoms, signs or clinical findings, not elsewhere classified	2	2.0%	1	1.4%
22 Injury, poisoning or certain other consequences of external causes	39	38.6%	20	27.8%
24 Factors influencing health status or contact with health services	3	3.0%	3	4.2%
Other diagnoses	4	4.0%	4	5.6%
Not applicable‖	18	17.8%	15	20.8%
Most common diagnoses*				
Injury -Head or brain injury	31 19	30.7% 18.8%	15 10	20.8% 13.9%
Dementia	11	10.9%	10	13.9%
Burn	6	5.9%	4	5.6%
HIV	4	4.0%	3	4.2%
Depression	3	3.0%	3	4.2%
Pregnancy complication	3	3.0%	3	4.2%
Q-fever and chronic fatigue syndrome	3	3.0%	2	2.8%
Cardiovascular disease (not specified)	2	2.0%	2	2.8%
Chronic kidney disease	2	2.0%	2	2.8%
Diabetes	2	2.0%	2	2.8%
Hypertension	2	2.0%	2	2.8%
Schistosomiasis	2	2.0%	2	2.8%
Stroke	2	2.0%	2	2.8%

continued on next page

Table 1. Continued

Characteristic	Publications (k = 101)		Studies (*n* = 72)	
	k	%	*n*	%
Others[¶]	16	15.8%	12	16.7%
Not applicable[‖]	18	17.8%	15	20.8%
Response levels of the EQ-5D+cognition				
EQ-5D-3L + 3-level cognition bolt-on	75	74.3%	50	69.4%
EQ-5D-5L + 5-level cognition bolt-on	24	23.8%	21	29.2%
EQ-5D-5L + 3-level cognition bolt-on	2	2.0%	1	1.4%
Number of cognition bolt-ons used				
1	100	99.0%	71	98.6%
4	1	1.0%	1	1.4%
Administration*				
Self-completed	69	68.3%	47	65.3%
Interviewer-administered	39	38.6%	28	38.9%
Not known	5	5.0%	2	2.8%
Self/proxy reporting*				
Self-report	85	84.2%	62	86.1%
Proxy-report	15	14.9%	13	18.1%
Proxy perspective				
Proxy-proxy	4	4.0%	3	4.2%
Proxy-patient	7	6.9%	6	8.3%
Perspective NR	4	4.0%	4	5.6%
Not known	4	4.0%	1	1.4%
Bolt-on modifications*				
Retrospective assessment[£]	17	16.8%	7	9.7%
'Last week' recall period	1	1.0%	1	1.4%
Pictorial EQ-5D-5L + pictorial 5-level cognition bolt-on	1	1.0%	1	1.4%

Note. When a study has multiple publications, the one with the largest sample size is considered the primary representation. NR indicates not reported.
*One publication or study may belong in multiple categories.
[†]Countries with 1 study each: Hungary, Indonesia, Malaysia, South Korea, Spain, Switzerland, Thailand, and United Kingdom.
[‡]A study may use more than 1 instrument language. Languages with one study each: Hungarian, Indonesian, Korean, Malay, Moroccan, Moroccan-Arabic, Sarnámi-Hindustani, Shona, Spanish, Sranan Tongo, Surinamese, Tarifit, Telugu, Thai, and Turkish.
[§]Designs with 1 study each: prospective panel, pragmatic clinical trial, and non-randomized controlled trial.
[‖]Studies that involved general populations, caregivers, or community-dwelling (older) adults.
[¶]Diagnoses with 1 study each: abdominal aorta aneurysm, dental problems, intensive care patients, liver transplant patients, mercury intoxication, schizophrenia, surgery patients, tendency to fall, trichinosis, and unspecified chronic diseases. Diagnosis with 2 studies: unspecified acute diseases.
[£]Preaccident, preburn, prehospitalization, or preinjury.

to further develop and adapt this list of criteria for the current study. First, the criteria do not include all psychometric properties for which evidence was available in the bolt-on literature (eg, patient-proxy agreement), whereas other aspects included (eg, missing rates) could not be assessed because of limited reporting. Second, the criteria do not consistently distinguish between psychometric assessment at the item level and at the level of the instrument (ie, EQ-5D with the bolt-on), and we considered it important to evaluate both. Third, although some criteria include proposed cutoff values, these are sometimes drawn from literature unrelated to bolt-on research. Moreover, although specific psychometric tests are recommended for certain properties, alternative methods are also commonly used in the literature. Lastly, many of the included publications in our systematic review were not explicitly designed for psychometric validation, with psychometric findings reported only as a secondary outcome. The authors of this review have extensive experience in psychometric testing of various bolt-ons, and their expertise also informed the development of the checklist. The checklist was iteratively refined through feedback from the research team and the EuroQol Group's Descriptive Systems Working Group and was finalized after being tested on 20 different publications (see Appendix Table 1 in Supplemental Materials.

When developing the checklist, we considered its potential broader applicability beyond the cognition bolt-on. For example, we created 2 separate checklist items (1a and 1b) to assess item-level ceiling. Item 1a uses a 70% threshold following earlier literature,[32] whereas item 1b compares the ceiling of the bolt-on item with that of the core EQ-5D items. For cognition, demonstrating a lower ceiling than any of the

Table 2. Ceiling, floor, and informativity of the EQ-5D-3L(+bolt-on) and EQ-5D-5L(+bolt-on).

Version	Wording code	Publication	Study population*	Sample size	Reporting	Ceiling	
						EQ-5D	EQ-5D+cognition
3L	A-3L-Co-03a	Dijkshoorn et al. 2022	Adult (18+) burn patients, 3 months post-burn	120	Self-report	39.0%	38.0%
3L	A-3L-Co-03a	Geraerds et al.	Adult (18+) injury patients, 1 month post-injury	2693	Self-report	Total sample: 7.9% TBI patients: 18.0% Non-TBI patients: 4.3%	Total sample: 7.3% TBI patients: 16.4% Non-TBI patients: 4.0%
3L	A-3L-Co-03a	Ophuis et al. 2019	Adolescent or adult (15+) patients with injury, 2.5 months post-injury	5346	Self-report	26.5%	25.5%
3L	A-3L-Co-21	Perneger & Courvoisier 2011	Adults (20+) from general population	349	Self-report	40.2%	32.7%
3L	Unknown	Schiffczyk et al. 2010	Adult (52+) patients with dementia	274	Self-report or proxy-report (perspective: proxy-patient)	Patient: 61.2% Proxy: 8.2%	Patient: 32.9% Proxy: 0.0%
5L	A-5L-Co-01a	Rencz and Janssen	Adults (18+) from general population	1587	Self-report	40.50%	38.5%
5L	A-5L-Co-01b	Geraerds et al.	Adolescent or adult (15+) patients presented at the emergency department due to trauma/poisoning, 6-12 months after trauma	1799	Self-report	All patients, 6 months: 26.3% TBI patients, 6 months: 36.3% PTSD patients, 6 months: 0.9% Neither TBI nor PTSD, 6 months: 23.5%	All patients, 6 months: 24.1% TBI patients, 6 months: 28.9% PTSD patients, 6 months:0.9% Neither TBI nor PTSD, 6 months: 22.2%
5L	A-5L-Co-01b	Geraerds et al.	Adult (18+) Q-fever patients, up to 10 years post-infection	432	Self-report	5.0%	2.0%
5L	A-5L-Co-02	De Graaf et al. 2020	Adult patients with stroke, 3 months post-stroke	360	Self-report	22.0%	14.0%
5L	A-5L-Co-16	Huang et al. 2022	Rural adults (16+) without any cognitive impairment or mental disease and with no serious hearing, vision or communication impairment	320	Self-report	30.9%	EQ-5D-5L+memory: 17.9% EQ-5D-5L +concentration: 26.6% EQ-5D-5L+learning: 19.4% EQ-5D-5L +understanding: 19.7%
5L	A-5L-Co-22	Kim et al. 2017	Adults (20+) from general population	600	Self-report	60.8%	58.5%

PTSD indicates posttraumatic stress disorder; TBI, traumatic brain injury.
*If the minimum age was not indicated, it was not clearly reported in the publication. However, no pediatric samples were included.
†Four cognition bolt-on wordings were assessed for this study.

core 5 dimensions may be difficult to achieve as (severe) cognitive impairments are likely to affect other dimensions, such as usual activities or anxiety/depression.[10,14,33,34] However, for other bolt-ons (eg, symptom- or functioning-specific bolt-ons, such as skin irritation or hearing problems), item 1b may well be relevant.

For 4 checklist items, we opted not to impose rigid numerical thresholds to define performance (eg, changes in ceiling or informativity for the EQ-5D+bolt-on compared with the EQ-5D alone). Instead, the adequacy of improvements in psychometric properties was assessed based on the authors' judgment, considering the study design and population. In cases in which a psychometric property could be assessed not only for the total sample but also for subsets of respondents (eg, ceiling) or in which comparisons were made across different variables or groups (eg, convergent and divergent validity, known-groups validity, or responsiveness), we followed a "75% of the respondent groups/cases/comparisons" rule. The criterion requiring a priori hypotheses and a threshold for correspondence was not included in our checklist because psychometric testing was not the primary aim of most studies, and including such a criterion would have resulted in very few studies being rated positively, even among high-quality ones.

Using this checklist, each psychometric property reported in a publication for a bolt-on version was evaluated as either positive (+) or nonpositive (−). Our team considered including an intermediate "unclear" or "doubtful" category but decided against it because of inconsistent interpretation and application across different checklists' different items, which could introduce additional subjectivity. However, we chose to use the term "nonpositive" instead of "negative" to better capture that these ratings include both negative and uncertain outcomes. For the publication reporting on multiple bolt-on items, each item was assessed individually. Any modifications to the bolt-ons (eg, pictorial versions or alternative recall periods) were assessed separately. The total number of positive and nonpositive evaluations was summed for each item. For item 1, only item 1a was included in the total score because of the reasons outlined above.

In addition to the total sample, assessments were also conducted for 2 subgroup categories: (1) bolt-on items grouped as

Table 2. Continued

Floor		Absolute informativity (H')		Relative informativity (J')		Assessed in Table 6	
EQ-5D	EQ-5D +cognition	EQ-5D	EQ-5D+cognition	EQ-5D	EQ-5D+cognition	Ceiling reduction	Informativity
-	-	3.62	3.96	0.46	0.42	✓	✓
-	-	TBI patients: 5.08 Non-TBI patients: 5.58	TBI patients: 5.88 Non-TBI patients: 6.38	TBI patients: 0.64 Non-TBI patients: 0.70	TBI patients: 0.62 Non-TBI patients: 0.67	✓	✓
0.4%	0.2%	-	-	-	-	✓	✗
-	-	-	-	-	-	✓	✗
0.0%	0.0%	-	-	-	-	✓	✗
-	-	4.81	5.49	0.41	0.39	✓	✓
-	-	All patients: 5.74	All patients: 6.37	All patients: 0.49	All patients: 0.46	✓	✓
-	-	-	-	-	-	✓	✗
0.0%	0.0%	-	-	-	-	✓	✗
0.0%	0.0%	-	-	-	-	✓[†]	✗
-	-	-	-	-	-	✓	✗

global items (eg, cognition with or without examples in the dimension title), composite items (eg, memory/concentration, similarly to pain/discomfort or anxiety/depression in the EQ-5D) and subdimension items (eg, items focusing specifically on memory); and (2) patient populations with at least 10 studies available.

Results

Inclusion of Relevant Publications

The literature search identified a total of 181 publications representing 137 individual studies eligible for inclusion in the first part of our systematic review.[26] Of these, 109 studies evaluated at least one psychometric property of a cognition bolt-on. Four publications were excluded because the psychometric results for populations aged <16 and ≥16 years could not be separated,[35-38] and a further 4 were excluded because of small sample sizes (<30).[39-42] Consequently, 101 publications representing 72 unique studies met the inclusion criteria for this second part of the systematic review (see Appendix Table 2 in Supplemental Materials.

Characteristics of Included Publications and Studies

The included studies originated from 18 countries, with The Netherlands (61%) contributing the most, followed by China (6%), and Belgium and New Zealand (4% each) (Table 1). Cognition bolt-ons were tested in 21 different languages, with Dutch (64%) being the most frequently used, followed by English (14%), Mandarin Chinese (6%) and French (4%). The majority of studies were either longitudinal cohorts (46%) or cross-sectional studies (40%). Overall, 32% of the studies had sample sizes exceeding 500 respondents. Four-fifths of the studies included patient populations, with the most common ICD-11 categories being "Injury, poisoning, or certain other consequences of external causes" (28%), "Mental, behavioral, or neurodevelopmental disorders" (19%), and "Certain infectious or parasitic diseases" (11%). The 2 most prevalent disease groups were head/brain injuries and dementia, each with 10 studies.

Bolt-On Versions

Psychometric properties were reported for 15 3L bolt-ons and 13 5L bolt-ons (see Appendix Table 3 in Supplemental Materials. For 20 studies (26 publications), the bolt-on wording was unknown.

Table 3. Convergent and divergent validity of EQ-5D-3L and EQ-5D-5L cognition bolt-ons.

Version	Wording code	Publication	Study population*	EQ-5D Reporting	Comparator	Type of correlation	Dimension-level analysis	Level sum score (LSS)-based analysis
Correlations with EQ-5D dimensions								
3L	A-3L-Co-03a	Ophuis et al. 2019	Adolescent or adult (15+) patients with injury, 2.5 months post-injury	Self-report	EQ-5D	Spearman's	Correlations between cognition bolt-on and EQ-5D-3L dimensions (P = NR): - Mobility: 0.223 - Anxiety/depression: 0.42 - Any other dimension <0.31	-
3L	Unknown	Van Dongen 2017a	Repatriated military personnels receiving treatment for wounds on the lower extremity sustained in Afghanistan, 5 years post-injury	Self-report	EQ-5D	Spearman's	Correlation between cognition bolt-on and EQ-5D pain/discomfort: 0.22 ($P \geq .05$)	-
5L	A-5L-Co-01a	Rencz and Janssen	Adults (18+) from general population	Self-report	EQ-5D	Spearman's	Correlations between cognition bolt-on and EQ-5D-5L dimensions ($P < .01$ for all): - Mobility: 0.27 - Self-care: 0.29 - Usual activities: 0.36 - Pain/discomfort: 0.36 - Anxiety/depression: 0.45	-
5L	A-5L-Co-01b	Geraerds et al.	Adult (18+) Q-fever patients, up to 10 years post-infection	Self-report	EQ-5D	Spearman's	Correlations between cognition bolt-on and EQ-5D-5L dimensions ($P < .001$ for all): - Mobility 0.291 - Self-care 0.292 - Usual activities 0.554 - Pain/discomfort 0.451 - Anxiety/depression 0.374	-
5L	A-5L-Co-02	De Graaf et al. 2020	Adult patients with stroke, 3 months post-stroke	Self-report	EQ-5D	Spearman's	Correlations between cognition bolt-on and EQ-5D-5L dimensions ($P \leq .001$ for all): - Mobility: 0.20 - Self-care: 0.25 - Usual activities: 0.42 - Pain/discomfort: 0.28 - Anxiety/depression: 0.35	- Correlation between cognition bolt-on and EQ-5D-5L LSS: 0.43 ($P \leq .001$) - Correlation between cognition bolt-on and EQ-5D-5L+cognition LSS: 0.62 ($P \leq .001$) - Correlations between EQ-5D-5L LSS and EQ-5D-5L+cognition LSS 0.93 (P = NR)
Correlations with EQ VAS[†]								
3L	A-3L-Co-03a	Dijkshoorn et al. 2022	Adult burn patients, 3 months post-burn	Self-report	EQ VAS	Spearman's	-	EQ-5D-3L LSS: −0.612 ($P < .001$) EQ-5D-3L+cognition LSS: −0.617 ($P < .001$)
3L	A-3L-Co-03a	Geraerds et al.	Adult (18+) injury patients, 1 month post-injury	Self-report	EQ VAS	Spearman's	-	Total sample, traumatic brain injury (TBI) patients, non-TBI patients (P = NR): - EQ-5D-3L LSS: −0.673, −0.719, −0.652 - EQ-5D-3L+cognition LSS: −0.690, −0.736, −0.665
5L	A-5L-Co-01a	Rencz and Janssen	Adults (18+) from general population	Self-report	EQ VAS	Spearman's	Cognition bolt-on: −0.35 ($P < .01$)	-
5L	A-5L-Co-01b	Geraerds et al.	Adolescent or adult (15+) patients presented at the emergency department due to trauma/poisoning, 6-12 months after trauma	Self-report	EQ VAS	Spearman's	-	- EQ-5D-5L LSS: −0.651 (95% CI −0.679 to −0.620) - EQ-5D-5L+cognition LSS: −0.664 (95% CI −0.693 to −0.634)

continued on next page

Table 3. Continued

Version	Wording code	Publication	Study population*	EQ-5D Reporting	Comparator	Type of correlation	Dimension-level analysis	Level sum score (LSS)-based analysis
Correlations with external measures[‡]								
3L	A-3L-Co-02	Wolfs et al. 2007	Older adult (55+) patients with cognitive impairment	Proxy-report (perspective: proxy-patient)	MMSE	Spearman's	Correlations between cognition bolt-on and MMSE[§]: - Baseline: −0.35 ($P < .01$) - 6 months: −0.52 ($P < .01$) - 12 months: −0.54 ($P < .01$)	-
3L	A-3L-Co-04	Makai et al. 2014	Older adults (55+) diagnosed with dementia living in the nursing home for longer than two months and nursing professionals	Proxy-report (perspective: proxy-patient)	ICECAP-O	Correlation not specified	Correlations between cognition bolt-on and ICECAP-O dimensions: - Attachment: 0.14 ($P \geq .05$) - Security: −0.07 ($P \geq .05$) - Role: 0.52 ($P < .01$) - Enjoyment: 0.25 ($P < .05$) - Control: 0.54 ($P < .01$) Correlation between cognition bolt-on and ICECAP-O index value: 0.48 ($P < .01$)	-
3L	A-3L-Co-10b	Hoeymans et al. 2005	Adult (18+) from general population	Self-report	SF-36	Spearman's	Correlations between cognition bolt-on and SF-36[‖] domains: Physical functioning: −0.21 ($P < .01$) Role-physical: −0.20 ($P < .01$) Bodily pain: −0.16 ($P < .01$) General health: −0.23 ($P < .01$) Vitality: −0.26 ($P < .01$) Social functioning: −0.26 ($P < .01$) Role-emotional: −0.23 ($P < .01$) Mental health: −0.26 ($P < .01$)	-
3L	Unknown	Arons & Krabbe 2011	Adult dementia patients and their caregivers	Self-report	MMSE	Spearman's	Correlations between cognition bolt-on and MMSE[§]: - Baseline: −0.61 ($P = .425$) - 6 months: −0.92 ($P = .267$) - 12 months: −0.68 ($P = .430$)	-
3L	Unknown	Arons & Krabbe 2011	Adult dementia patients and their caregivers	Proxy-report (perspective: proxy-patient)	MMSE	Spearman's	Correlations between cognition bolt-on and MMSE[§]: - Baseline: −0.118 ($P = .123$) - 6 months: −0.143 ($P = .86$) - 12 months: −0.258 ($P = .02$)	-
3L	Unknown	Sarabia-Cobo et al. 2017	Older adults (65+) living in the nursing home for >3 months	Proxy-report (perspective: proxy-patient)	ICECAP-O	Correlation not specified	Correlations between cognition bolt-on and ICECAP-O dimensions: - Attachment: 0.02 ($P \geq .05$) - Security: −0.05 ($P \geq .05$) - Role: 0.36 ($P < .01$) - Enjoyment: 0.47 ($P < .01$) - Control: 0.35 ($P < .01$) Correlation between cognition bolt-on and ICECAP-O index value: 0.54 ($P < .01$)	-

continued on next page

Table 3. Continued

Version	Wording code	Publication	Study population*	EQ-5D Reporting	Comparator	Type of correlation	Dimension-level analysis	Level sum score (LSS)-based analysis
3L	Unknown	Schölzel-Dorenbos et al. 2012	Older adult patients with mild to moderate dementia living independently	Proxy-report (perspective: NR)	DQI	Spearman's	Correlations between cognition bolt-on and DQI domains: - Memory§: 0.49 ($P < .01$) - Orientation¶: 0.36 ($P < .01$) - Independence‖: 0.39 ($P < .01$) - Social activities‖: 0.11 ($p>0.05$) - Mood‖: 0.19 ($P < .05$)	-
3L	Unknown	Schölzel-Dorenbos et al. 2012	Older adult patients with mild to moderate dementia living independently	Self-report	DQI	Spearman's	Correlations between cognition bolt-on and DQI domains: - Memory§: 0.35 ($P < .01$) - Orientation¶: 0.32 ($P < .01$) - Independence‖: 0.20 ($P < .01$) - Social activities‖: −0.09 ($p>0.05$) - Mood‖: 0.05 ($p>0.05$)	-
3L	Unknown	Van Dongen 2017a	Repatriated military personnels receiving treatment for wounds on the lower extremity sustained in Afghanistan, 5 years post-injury	Self-report	SF-36	Spearman's	Cognition bolt-on and SF-36 bodily pain‖ $r = -0.29$ ($P < .05$)	-
5L	A-5L-Co-01a	Igarashi et al.	Older adults with dementia living in nursing homes and their caregivers	Proxy-report (perspective: proxy-patient)	MMSE	Spearman's	Correlation between cognition bolt-on and MMSE§ Baseline: −0.640 (95% CI −0.730 to −0.550)	-
5L	A-5L-Co-01a	Rencz and Janssen	Adult (18+) general population	Self-report	PROMIS-29+2, PROMIS Global Health	Spearman's	Correlations between cognition bolt-on (P = NR), in absolute values: PROMIS-29+2§: - Being able to concentrate: 0.35 - Being able to remember to do things: 0.32 - Cognitive function total score (raw): 0.37 PROMIS Global Health§, Mental Health (including mood and ability to think): 0.36	-
5L	A-5L-Co-02	De Graaf et al. 2020	Adult patients with stroke, 3 months post-stroke	Self-report	mRS	Spearman's	Correlation between cognition bolt-on and mRS§: −0.42 ($P \leq .001$)	- EQ-5D-5L LSS and mRS −0.66 ($P \leq .001$) - EQ-5D-5L+cognition and mRS −0.67 ($P \leq .001$)
5L	A-5L-Co-16	Huang et al. 2022	Rural adults (16+) without any cognitive impairment or mental disease and with no serious hearing, vision or communication impairment	Self-report	MMSE	Spearman's	Correlations between memory bolt-on and MMSE dimensions ($P < .01$): - Orientation: 0.424 - Registration: 0.216 - Attention and calculation: 0.376 - Recall§: 0.444 - Language: 0.425	-
5L	A-5L-Co-19	Huang et al. 2022	Rural adults (16+) without any cognitive impairment or mental disease and with no serious hearing, vision or communication impairment	Self-report	MMSE	Spearman's	Correlations between concentration bolt-on and MMSE dimensions ($P < .01$): - Orientation: 0.225 - Registration: 0.187 - Attention and calculation§: 0.270 - Recall: 0.280 - Language: 0.255	-

continued on next page

Table 3. Continued

Version	Wording code	Publication	Study population*	EQ-5D Reporting	Comparator	Type of correlation	Dimension-level analysis	Level sum score (LSS)-based analysis
5L	A-5L-Co-27	Huang et al. 2022	Rural adults (16+) without any cognitive impairment or mental disease and with no serious hearing, vision or communication impairment	Self-report	MMSE	Spearman's	Correlations between learning bolt-on and MMSE dimensions ($P < .01$): - Orientation[§]: 0.545 - Registration[§]: 0.401 - Attention and calculation[§]: 0.521 - Recall[§]: 0.596 - Language[§]: 0.604	-
5L	A-5L-Co-29	Huang et al. 2022	Rural adults (16+) without any cognitive impairment or mental disease and with no serious hearing, vision or communication impairment	Self-report	MMSE	Spearman's	Correlations between understanding bolt-on and MMSE dimensions ($P < .01$) - Orientation[§]: 0.522 - Registration[§]: 0.347 - Attention and calculation[§]: 0.592 - Recall[§]: 0.448 - Language[§]: 0.522	-
5L	Unknown	Easton et al. 2018	Older adults living in residential care facilities	Proxy-report (perspective: NR)	DEMQOL, PAS-Cog, MBI, NPI-Q	Spearman's	Correlations between cognition bolt-on and DEMQOL-Proxy dimensions: - Positive emotion[‖]: 0.243 ($P < .01$) - Negative emotion[‖]: 0.025 ($P \geq .05$) - Memory[§]: −0.005 ($P \geq .05$) - Appearance: 0.128 ($P < .05$) Correlations between cognition bolt-on and: - DEMQOL-Proxy index value: −0.177 ($P < .01$) - PAS-Cog[§]: 0.495 ($P < .01$) - MBI[§]: −0.488 ($P < .01$) - NPI-Q[‖]: 0.062 ($P \geq .05$)	-
5L	Unknown	Easton et al. 2018	Older adults living in residential care facilities	Self-report	DEMQOL, PAS-Cog, MBI, NPI-Q	Spearman's	Correlations between cognition bolt-on and DEMQOL dimensions: - Positive emotion[‖]: 0.181 ($P < .05$) - Negative emotion[‖]: −0.158 ($P \geq .05$) - Loneliness[‖]: −0.120 ($P \geq .05$) - Cognition[§]: −0.304 ($P < .01$) - Relationships[‖]: −0.234 ($P < .01$) Correlations between cognition bolt-on and: - DEMQOL index value: −0.222 ($P < .01$) - PAS-Cog[§]: 0.025 ($P \geq .05$) - MBI[§]: −0.039 ($P \geq .05$) - NPI-Q[‖]: 0.045 ($P \geq .05$)	-

NR indicates not reported; *P*, *P* value. Outcome measures: DEMQOL, dementia quality of life-utility; DQI, dementia quality of life; ICECAP-O, ICEpop capability measure for older people; MBI, modified Barthel index; MMSE, mini mental state examination; mRS, modified Rankin scale; NPI-Q, neuropsychiatric inventory questionnaire; PAS-Cog, psychogeriatric assessment scales-cognitive impairment scale; PROMIS, patient-reported outcomes measurement information system; SF-36, 36-item short form health survey.

*(1) In instances in which multiple populations are involved (eg, adults and caregivers), we underlined the population corresponding to the information presented in the table, and (2) If the minimum age was not indicated, it was not clearly reported in the publication. However, no pediatric samples were included.

†In instances in which information was not assessed in Table 6 (indicated by ✗), it was because of (1) missing information (eg, correlations not reported for all EQ-5D dimensions), (2) EQ VAS only correlated with the cognition bolt-on but not with the EQ-5D+cognition level sum score, or (3) external measures not measuring relevant health-related quality of life construct.

‡To evaluate convergence with or divergence from external measures, we relied on our research team's judgment, which considered: (1) the wording of the bolt-on, (2) the population, and (3) the content of the external measure's items.

§External measures (ie, other than the EQ-5D descriptive system and EQ VAS) that were used to evaluate the convergent validity of the cognition bolt-ons in this systematic review).

‖External measures (ie, other than the EQ-5D descriptive system and EQ VAS) that were used to evaluate the divergent validity of the cognition bolt-ons in this systematic review).

¶Not assessed due to absence of wording.

Table 4. Known-group validity of EQ-5D-3L and EQ-5D-5L cognition bolt-ons.

Known groups*	Cognition bolt-on			EQ-5D+cognition bolt-on			In statistically significant outcomes, more cognition problems were reported by
	3L	5L	Total	3L	5L	Total	
Acute or chronic disease/condition, or multimorbidity	**9+/5−**	**2+/0−**	**16**	**1+/0−**	**2+/0−**	**3**	
Battle injury (military service members injured in battle vs not)	1+		1				Injured in battle group
Brain injury (trauma patients with brain injury vs no brain injury)	1+		1				Brain injury group
Burn (patients with severe vs minor burn injuries)		1+	1				Severe burn group
Chronic disease (trauma/poisoning/injury patients with two or more chronic diseases vs one or none)	1+		1		1+	1	Multiple or with chronic disease(s) groups
Chronic Q-fever (patients with chronic Q-fever vs QFS vs QFS-like disease)		1+	1				QFS group
Comorbidities (patients with two or more comorbidities vs one or none, or with vs without comorbidities)	3+/1−		4				Multiple or more comorbidities groups
Delirium (survivors of critical illness with delirium vs without)	1+		1				Delirium group
Dental conditions (patients with caries, fracture, denture problems, failed restoration, pulpal infection, wear, sensitivity, or aesthetics)				1+		1	Denture problems group
Disability (former injury patients with vs without disability)	1−		1				-
Disease or non-battle injuries (DNBI): repatriated military service members vs not due to DNBI	1−		1				-
Frailty (former injury patients with higher vs lower frailty score)	1+		1				Higher frailty group
HIV (HIV-positive vs negative individuals)	1−		1				-
Stroke (former patients with ischemic stroke vs without, or with stroke vs matched injury)	1+/1−		2				Stroke group
Traumatic brain injury (patients with traumatic brain injury or PTSD vs without)					1+	1	With traumatic brain injury or PTSD
Threats and perceptions	**1+/3−**	**0+/2−**	**6**	**0+/0−**	**0+/0−**	**0**	
Fall risk (older adults with high vs low fall risk)		2−	2				-
Threat of disability (former injury patients with self-perceived threat of disability vs none)	1−		1				-
Threat to life (former injury patients with self-perceived threat of disability vs none)	1+		1				With threat of life group
Environmental exposure	**0+/2−**	**0+/0−**	**2**	**0+/0−**	**0+/0−**	**0**	
Mercury contact (miners with mercury exposure vs none)	1−		1				-
Mining years (miners with more vs less than five years of exposure)	1−		1				-
Caregiving	**0+/0−**	**0+/0−**	**0+/0−**	**2+/0−**	**0+/0−**	**2**	
Clusters of dementia patients as assessed by caregivers (patients with more severe cognitive impairments vs less severe)[†]				2+		2	Cluster of patients with more severe impairments

continued on next page

Table 4. Continued

Known groups*	Cognition bolt-on			EQ-5D+cognition bolt-on			In statistically significant outcomes, more cognition problems were reported by
	3L	5L	Total	3L	5L	Total	
Injury-related characteristics	**12+/2−**	**0+/0−**	**14**	**0+/0−**	**0+/0−**	**0**	
Calcaneal compression angle (patients with negative calcaneal compression angle vs not)	1+		1				Negative calcaneal compression angle group
Fracture (patients with high-energy vs low-energy trauma classification)	1+		1				High-energy trauma group
Functional capacity index (patients with higher vs lower indices)	1+		1				Lower capacity index group
Glasgow Outcome Scale-Extended (patients with better vs worse scores)	1+		1				Worse scores group
Injury cause (caused by assaultive intent vs not, or nature of injury: fracture, sprain and strain, concussion, open wound/amputation, contusion/ superficial, other single injury)	2+		2				Caused by assaultive intent group (more cognition problems), sprain and strain group (fewer cognition problems)
Injury location (eg, lower extremity, upper extremity, head and neck, spine and back, and torso)	3+		3				Traumatic or brain, head or skull, or intracranial groups
Injury Severity Score (patients with higher vs lower scores)	1+		1				Higher score group
New Injury Severity Score (patients with higher vs lower scores)	1+		1				Higher score group
Pre-injury problems (with pre-injury problems vs not)	1+		1				With pre-injury problems group
Prehospital Trauma Triage Protocol (level of injuries)	1−		1				-
Revised Trauma Score (score level at emergency department)	1−		1				-
Inpatient/outpatient care	**3+/5−**	**0+/1−**	**9**	**1+/0−**	**0+/0−**	**1**	
Duration of medical evacuation from point of injury to arrival in hospital (within vs after 72 hours)	1−		1				-
Emergency visit (patients with dental emergency vs not)				1+		1	Emergency group
Hospital admission (admitted to the hospital vs not)	1+/1−	1−	3				Hospital-admitted group
Hospital readmission (readmitted to the hospital vs not)	1−		1				-
Hospitalization duration (shorter vs longer hospitalization, eg, 8-14 days vs 15 + days)	1+		1				Longer hospitalization group
ICU admission (patients admitted to the ICU vs not)	1+		1				ICU group
ICU duration (shorter vs longer ICU care)	1−		1				-
Trauma care (higher vs lower level of facilities of trauma care required)	1−		1				-
Lifestyle-related characteristics	**1+/3−**	**0+/0−**	**4**	**0+/0−**	**0+/0−**	**0**	
Alcohol use (hazardous alcohol use vs not)	1−		1				-
Body mass index (ie, underweight, normal, overweight, or obese)	1−		1				-
Physical activity (active vs not)	1+		1				Physically active group (in injury patients)
Smoking (regularly vs not)	1−		1				-

continued on next page

Table 4. Continued

Known groups*	Cognition bolt-on			EQ-5D+cognition bolt-on			In statistically significant outcomes, more cognition problems were reported by
	3L	5L	Total	3L	5L	Total	
Mental health/disorder	**1+/1−**	**0+/0−**	**3**	**0+/0−**	**2+/0−**	**2**	
Depression (depressive episode vs not)	1−		1				-
PROMIS Global04 item (mental health, including mood and ability to think)					1+	1	Worse mental health group
Psychiatric comorbidities (injured trauma patients with vs without any psychiatric comorbidity)		1−	1				-
PTSD or symptoms (patients with PTSD or its symptoms vs without)	1+		1		1+	1	With PTSD (symptoms) group
Overall health	**1+/0−**	**0+/0−**	**1**	**0+/0−**	**0+/0−**	**0**	
Fair/poor health vs good and very good/ excellent health	1+		1				Poor health group
Severity groups (non-injury)	0+/0−	0+/0−	0	1+/2−	2+/0−	5	
Mini Mental State Examination‡ (patients with better vs worse scores)				1+/2−		3	Worse scores group
Modified Rankin Scale (patients with better vs worse scores)					2+	2	Worse scores group
Socioeconomic characteristics	**17+/10-**	**0+/0−**	**27**	**6+/0−**	**1+/1−**	**8**	
Age	6+/2−		8	1+	1+	2	Older age (*n* = 5 publications) and younger age (*n* = 3)
Education (low vs higher groups)	2+/1−		3	1+		1	Lower educated group
Employment (employed or working vs not)	2+		2				Unemployed, or did not return to work post-burn group
Ethnicity	1+		1	1+		1	Moroccan, Surinamese, or Turkish groups (vs Dutch or Indonesian), and Pacific group (vs non-Pacific)
Financial status (sufficient vs not)	1−		1				Financially challenged group
Household size (living alone vs not)	2+/1−		3				Living alone group
Insurance compensation (compensated vs not)	1+		1				Uncompensated group
Insurance status (insured vs not)				1+		1	Uninsured group
Neighborhood (more deprived vs not)				1+		1	More deprived neighborhood group
Relationship status (married vs not)	1−		1				-
Sex/gender (male vs female)	3+/4−		7	1+	1−	2	Female (*n* = 3 publications), male (*n* = 1)
Treatment/intervention	**2+/8−**	**0+/1−**	**11**	**1+/0−**	**0+/1−**	**2**	
Expectant monitoring vs immediate delivery outcomes at postpartum, 6 weeks				1+		1	Expectant monitoring group
Frail older adult intervention (multidisciplinary comprehensive care model vs care as usual, assessment with vs without motivational interviewing)	1−	1−	2				-
Hemodialysis (recipient of 1 year and above vs less than 1 year)					1−	1	-
Number of surgeries (patients with 1 vs more surgeries)	1−		1				-
Preference for induction of labor (vs expectant management, or vs no preference) in late pregnancy	1+		1				Preference for induction of labor group
Palliative care planning vs no planning	1−		1				-
Surgery classification (eg, trauma vs non-trauma surgeries, amputation vs limb salvage)	1+/3−		4				Trauma surgery group (vs oncology, vascular, gastrointestinal, or general surgery)
Use of mechanical ventilation vs no use	2−		2				-

continued on next page

Table 4. Continued

Known groups*	Cognition bolt-on			EQ-5D+cognition bolt-on			In statistically significant outcomes, more cognition problems were reported by
	3L	5L	Total	3L	5L	Total	
Other groups	**1+/1−**	**0+/0−**	**2**	**0+/0−**	**1+/0−**	**2**	
Level of albumin in blood (hemodialysis patients with higher vs lower blood albumin level)					1+	1	Lower albumin group
Healthcare access (patients with trouble vs no trouble of access)	1+		1				Trouble of access group
One year vs two years after trauma care	1−		1				-
TOTAL	**48+/38−**	**2+/5−**	**93**	**12+/2−**	**8+/2−**	**24**	

Note. "+" indicates statistically significant outcomes ($P < .05$, $P < .01$ or $P < .001$), whereas "−" indicates insignificant outcomes. Bold indicates the total rows, representing the combined values of multiple categories.
PTSD, posttraumatic stress disorder; QFS, Q-fever fatigue syndrome.
*Assessment of 1 known-group category may be from more than 1 publication and may have different categorizations (eg, sex/gender is consistently male vs female across all publications, but number of comorbidities can be either with vs without or multiple vs one vs none).
†The 2 assessments were from 1 publication (Jonas et al., 2011) because they reported for both self-report and proxy-report.
‡The 3 assessments were from 1 publication (Schiffczyk et al. 2010) because they reported self-report by patients, self-report by caregivers, and proxy-report (perspective: proxy-patient).

Psychometric Properties

Qualitative evidence

None of the publications reported evidence on content validity or qualitative aspects of acceptability.

Bolt-on item response distribution, floor, and ceiling

Response distributions were available for 194 respondent groups (81% 3L) across 75 publications (see Appendix Table 4 in Supplemental Materials. For 3L bolt-ons, the highest floor was 53% (older adults with cognitive impairment), whereas for 5L bolt-ons, it was 10% (1-year after major trauma, with psychiatric comorbidities). In multiple populations, no responses on the worst level were reported either for the 3L or 5L. The item-level ceiling for 3L bolt-ons ranged from 3% (dementia) to 100% (teeth sensitivity), whereas for 5L bolt-ons, it ranged from 5% (Q-fever fatigue syndrome) to 84% (adult general population). Based on a ceiling threshold of <70%, 5L bolt-ons performed better (17+/7−) than 3L bolt-ons (18+/25−). In most cases, respondents reported more problems on at least 1 of the 5 core items than on the bolt-on (3L: 3+/37−; 5L: 2+/18−).

EQ-5D+bolt-on floor, ceiling, and informativity

At the instrument level, floor, ceiling, and absolute and relative informativity were reported in 4, 11, 4, and 4 publications, respectively (Table 2). For both the EQ-5D and the EQ-5D+bolt-on, the floor was (close to) 0% in all of these publications. Adding 3L bolt-ons resulted in an absolute reduction in ceiling of the EQ-5D ranging from 0.3%-point (1-month post-injury, no traumatic brain injury) to 28%-points (dementia). For the 5L, the absolute reduction in ceiling ranged from 0%-points (6-12 months post-trauma/poisoning, with posttraumatic stress disorder) to 13%-points (adults without cognitive impairment, mental illness, or serious hearing, vision, or communication impairments). Overall, there was limited evidence for adequate reduction in ceiling with the bolt-on (3L: 2+/3−; 5L: 6+/3−). In all cases, absolute informativity improved after adding the bolt-on (change 3L: 0.34-0.80; 5L: 0.63-0.68). However, relative informativity slightly deteriorated (change 3L: −0.02 to −0.04; 5L: −0.02 to −0.03).

Divergent validity

Appendix Table 5 in Supplemental Materials presents the constructs that were and were not considered for divergent and convergent validity assessments. Eight publications assessed divergent validity (Table 3). Overall, the bolt-ons demonstrated divergence from the core EQ-5D dimensions (3L: 5+/0−; 5L: 4+/1−). In all cases, the correlations with other instruments' unrelated domains (eg, those of SF-36) were weak to moderate. However, some bolt-ons exhibited moderate or strong correlations with usual activities and anxiety/depression EQ-5D dimensions.

Convergent validity

Eight publications evaluated convergent validity of cognition bolt-ons with external instruments' domains, the most common being the MMSE. Several HRQoL instruments' relevant items used for convergence testing appeared in only 1 study each, such as PROMIS Global Health (r = 0.36), PROMIS-29+2 (0.32-0.35), DEMentia Quality of Life (0.01 to 0.30), and Dementia Quality of Life Instrument (0.35-0.49) (Note: correlation coefficients are in absolute values). The performance of the bolt-ons was mixed (3L: 1+/4−; 5L: 5+/4−).

Known-Groups validity

Overall, 54 publications assessed the known-groups validity of the bolt-ons across 72 different known groups (Table 4). The 3L and 5L bolt-on items were able to differentiate between 56% and 29% of known groups, whereas the EQ-5D+bolt-on was able to differentiate between known groups in 88% (3L) and 80% (5L). The greatest amount of positive evidence was found for the following known groups: injury-related characteristics, acute and chronic conditions and comorbidities and sociodemographic factors. More evidence was available for the bolt-on items (3L: 15+/18−; 5L: 2+/4−) than for the EQ-5D+bolt-on items (3L: 6+/1−; 5L: 3+/1).

Structural validity

Two publications assessed the structural validity of the EQ-5D with bolt-on(s) using external instruments included in the item pool, such as PROMIS-29+2, SF-6D, SWLS, and OHIP-14 (see Appendix Table 6 in Supplemental Materials. These used

Table 5. Explanatory power of EQ-5D-3L and EQ-5D-5L cognition bolt-ons with EQ VAS as the dependent variable.

Version	Wording code	Publication	Study population*	Reporting	R-squared (EQ-5D)	R-squared (EQ-5D+bolt-on)	Assessed in Table 6 (associations with EQ VAS)
3L	A-3L-Co-03a	Dijkshoorn et al. 2022	Adult (18+) patients one year after ICU admission	Self-report	52.90%	56.90%	✓
3L	A-3L-Co-03a	Geraerds et al.	Adult (18+) injury patients, 1 month post-injury	Self-report	All patients: 48.0%[†] Patients with TBI: 52.8%[†] Patients with no TBI: 45.1%[†]	All patients: 49.9%[†] Patients with TBI: 56.0%[†] Patients with no TBI: 46.6%[†]	✓
3L	A-3L-Co-03a	Ophuis et al. 2019	Adolescent or adult (15+) patients with injury, 2.5 months post-injury	Self-report	All patients: 45.6%[†] TBI patients: 55.6%[†] Non-TBI patients: 44.5%[†]	All patients: 46.9%[†] TBI patients: 56.5%[†] Non-TBI patients: 45.8%[†]	✓
3L	Unknown	Arons & Krabbe 2011	Adult dementia patients and their caregivers	Proxy-report (perspective: proxy-patient)	Baseline: 21.8%[†] 6 months: 15.9%[†] 12 months: 26.0%[†]	Baseline: 23.5%[†,‡] 6 months: 20.0%[†,‡] 12 months: 28.7%[†,‡]	✓
3L	Unknown	Arons & Krabbe 2011	Adult dementia patients and their caregivers	Self-report	Baseline: 38.0%[†] 6 months: 22.8%[†] 12 months: 31.6%[†]	Baseline: 37.7%[†] 6 months: 21.9%[†] 12 months: 30.8%[†]	✓
3L	Unknown	Arons & Krabbe 2011	Adult dementia patients and their caregivers	Self-report	Baseline: 27.6%[†] 6 months: 25.0%[†] 12 months: 21.3%[†]	Baseline: 33.5%[†,‡] 6 months: 25.5%[†,‡] 12 months: 27.6%[†,‡]	✓
5L	A-5L-Co-01a	Rencz and Janssen	Adults (18+) from general population	Self-report	Anxiety disease group: 46.27%[†] Hearing problems disease group: 54.22%[†]	Anxiety disease group: 50.96%[†,‡] Hearing problems disease group: 56.42%[†,‡]	✓
5L	A-5L-Co-01b	Geraerds et al.	Adolescent or adult (15+) patients presented at the emergency department due to trauma/ poisoning, 6-12 months after trauma	Self-report	Patients with TBI or PTSD: 41.3% Patients with no TBI and no PTSD: 40.1%	Patients with TBI or PTSD: 41.5%[†] Patients with no TBI and no PTSD: 41.4%[†]	✓
5L	A-5L-Co-01b	Geraerds et al.	Adult (18+) Q-fever patients, up to 10 years post-infection	Self-report	45.6%[†]	45.7%[†]	✓

continued on next page

Table 5. Continued

Version	Wording code	Publication	Study population*	Reporting	R-squared (EQ-5D)	R-squared (EQ-5D+bolt-on)	Assessed in Table 6 (associations with EQ VAS)
5L	A-5L-Co-16, A-5L-Co-19, A-5L-Co-27, A-5L-Co-29	Huang et al. 2022	Rural adults (16+) without any cognitive impairment or mental disease and with no serious hearing, vision or communication impairment	Self-report	40.4%[†]	Four cognition bolt-ons were used in the same study: -EQ-5D-5L +memory: 42.8%[†,‡] -EQ-5D-5L +concentration: 42.1%[†,‡] -EQ-5D-5L +learning: 46.6%[†,‡] -EQ-5D-5L +understanding: 42.9%[†,‡] -EQ-5D-5L +memory, concentration, learning, understanding: 47.3%[†,‡]	✓[§]

Note. All regression models were multivariate.
ISS indicates injury severity score; PTSD, posttraumatic stress disorder; SWLS, Satisfaction with Life Scale; TBI, traumatic brain injury.
*(1) In instances in which multiple populations are involved (eg, adults and caregivers), we underlined the population corresponding to the information presented in the table, and (2) if the minimum age was not indicated, it was not clearly reported in the publication. However, no pediatric samples were included.
[†]Adjusted R-squared.
[‡]Cognition bolt-on operated as an ordinal variable (ie, not dummy-coded).
[§]Four cognition bolt-on wordings were assessed for this study.

principal component analysis, confirmatory factor analysis, and hierarchical cluster analysis, with no studies using item response theory approaches. Both publications reported that the cognition bolt-ons loaded on the same factor/cluster as the anxiety/depression item. In 1 of these studies, the cognition bolt-on did not load onto the cognition factor, which comprised PROMIS-29+2 items "being able to remember to do things" (0.976) and "being able to concentrate" (0.933).

Associations with EQ VAS

Four publications assessed the correlations between EQ-5D+bolt-ons and the EQ VAS. In 3 of these, the EQ-5D+bolt-on level sum scores demonstrated slightly stronger correlations with the EQ VAS compared with the EQ-5D alone (difference in r: 0.005 to 0.017) (Table 3). Eight publications used multivariable regressions of the EQ VAS on the EQ-5D and bolt-on items (Table 5). For the 3L, the (adjusted) R-squared increase after adding the bolt-on ranged from 0.05%-points to 6.3%-points (both in dementia). In 1 case, the adjusted R-squared decreased after adding the 3L cognition bolt-on (caregivers). For the 5L, the increase in (adjusted) R-squared ranged from 0.1%-points (Q-fever up to 10 years post-infection) to 6.9%-points (adults without any cognitive impairment, mental disease, or serious hearing, vision, or communication impairment). Overall, most bolt-on items performed well in terms of associations with EQ VAS (3L: 5+/1−; 5L: 6+/1−).

Responsiveness

Despite the large number of longitudinal studies, only 3 publications reported on the responsiveness of bolt-ons (see Appendix Table 7 in Supplemental Materials. Two of these used the MMSE as an anchor for change, and 1 publication reported on the Paretian classification of health change for the EQ-5D+bolt-on. The results were not favorable for either bolt-on item (item-level responsiveness 3L: 0+/0−; 5L: 0+/2−, instrument-level responsiveness: 0+/1−; 5L: 0+/1−).

Test-retest reliability

One publication reported unfavorable test-retest reliability evidence (3L: 0+/0−; 5L: 0+/1−) for the bolt-on using MMSE as an anchor (see Appendix Table 7 in Supplemental Materials.

Interrater agreement

Two publications assessed patient-proxy agreement for cognition bolt-ons, one in dementia and the other in pre-burn as assessed retrospectively (see Appendix Table 7 in Supplemental Materials. Both used a proxy-patient perspective. The results indicated good performance in the preburn sample.

Differential item functioning

None of the publications reported evidence on differential item functioning.

Overall Psychometric Performance by 3L Versus 5L, Bolt-On Item Wording, and Disease Groups

Overall, 5L bolt-ons performed better than 3L bolt-ons based on the positive-to-non-positive ratio (3L: 55+/57−; 5L: 45+/28−) (Table 6). The improvement achieved with the 5L was particularly noticeable for the item-level ceiling, with fewer nonpositive assessments (37− vs 18−). For the 3L, the majority of evidence available was for global items (27+/27−), with limited for composites (3+/0−) and none for subdimension items. For the 5L, similarly, most evidence was mixed for global items (26+/

Table 6. Results of psychometric assessment using the checklist.

Wording	Ceiling (bolt-on item) using the 70% cutoff*	Ceiling (bolt-on item) relative to the core EQ-5D dimensions*	Ceiling (EQ-5D +cognition)	Absolute informativity (EQ-5D +cognition)	Relative informativity (EQ-5D +cognition)	Divergent validity (bolt-on item)**	Convergent validity (bolt-on item)***
Three-level (3L) version							
A-3L-Co-02[£]	1+	1+					1−
A-3L-Co-03[††,£]	6−	5−	3−	2+	2−	1+	
A-3L-Co-03-modified[‡‡,§§]	2−	2−					
A-3L-Co-05[£]	1+	1−					
A-3L-Co-09[£]	2−	2−					
A-3L-Co-10[£]	1−	1−				1+	
A-3L-Co-11[£]	2−	3−					
A-3L-Co-12[£]	1+/1−	2+					
A-3L-Co-13-modified[‡‡]	1+	1−					
A-3L-Co-16[£]	1−	1−					
A-3L-Co-17[£]	2+	2−					
A-3L-Co-19[£]	4+/1−	3−					
A-3L-Co-20[‖‖]	1+	1−					
A-3L-Co-21[‖‖]	1+	1−	1+				
A-3L-Co-22[£]	1+	1−					
A-3L-Co-22-modified[‡‡]	2−	2−					
Unknown wording	5+/7−	11−	1+			3+	1+/3−
3L global items (not modification)	**10+/14−**	**3+19−**	**0+/3−**	**2+/0−**	**0+/2−**	**2+/0−**	**0+/1−**
3L composite items (not modification)	**2+/0−**	**0+/2−**	**1+/0−**	**0**	**0**	**0**	**0**
3L TOTAL	**18+/25−**	**3+/37−**	**2+/3−**	**2+**	**2−**	**5+**	**1+/4−**
Five-level (5L) version							
A-5L-Co-01[¶¶]	7+/2−	8−	1+/2−	2+	2−	1+/1−	1+/1−
A-5L-Co-02[¶¶]	1+		1+			1+	1+
A-5L-Co-03[††,¶¶]	1−	1−					
A-5L-Co-04[¶¶]	1+	1−					
A-5L-Co-08-modified[¶]	1−						
A-5L-Co-12[¶¶]	1+/1−	1−					
A-5L-Co-16[‡,££]	1+	1−	1+				1+
A-5L-Co-19[‡,££]	1+	1−	1+				1−
A-5L-Co-20[¶¶]	1+	1−					
A-5L-Co-22[¶¶]	1−	1−	1−				
A-5L-Co-25[¶¶]	1−	1−					
A-5L-Co-27[‡,££]	1+	1+	1+				1+
A-5L-Co-29[‡,££]	1+	1+	1+				1+
Unknown wording	2+	2−				2+	2−
5L global items (not modification)	**11+/6−**	**0+/14−**	**2+/3−**	**2+/0−**	**0+/2−**	**2+/1−**	**2+/1−**
5L subdimension items (not modification)	**4+/0−**	**2+/2−**	**4+**	**0**	**0**	**0**	**3+/1−**
5L TOTAL	**17+/7−**	**2+/18−**	**6+/3−**	**2+**	**2−**	**4+/1−**	**5+/4−**
GRAND TOTAL (3L+5L)	**35+/32−**	**5+/55−**	**8+/6−**	**4+**	**4−**	**9+/1−**	**6+/8−**

Note. Bold indicates the total rows, representing the combined values of multiple categories.
*The ceiling assessment using the 70% cutoff rule includes 7 additional assessments from 7 publications, as they did not require the core EQ-5D distributions, which were not reported in the following publications: De Graaf et al., 2020; Geraerds et al.[34]; Goei et al., 2016; Looman et al., 2018; Mahapatra et al., 2002; Verbooy et al., 2018; and Wolters et al., 2014.
†Three assessments were provided by 1 publication (Arons and Krabbe, 2011) because it reported 3 different reporting: self-report (patient), proxy-patient (caregiver), and proxy-proxy (self-report by caregiver).
‡These 4 bolt-ons were used in a single publication/study (Huang et al, 2022).
§The results of both the principal component and confirmatory factor analyses were used in our assessment because they aligned.[10]
‖Does not include the ceiling (bolt-on item) assessment relative to the core EQ-5D dimensions.
¶Pictorial version of the EQ-5D-5L + 5-level bolt-on.
£3L global items (not modification).
**Four assessments were provided by 2 publications (Easton et al, 2018 and Schölzel-Dorenbos et al, 2012) because each reported both self-report and proxy-report results.
††A-3L-Co-03 and A-5L-Co-03 have the same wording in terms of language, dimension headings, examples, and phrasing.
‡‡Modified retrospective assessment (ie, pre-accident, pre-burn, pre-hospitalization, or pre-injury).
§§Modified 'last week' recall period.
‖‖3L composite items (not modification).
¶¶5L global items (not modification).
££5L subdimension items (not modification).
***Six assessments were provided by 3 publications (Arons and Krabbe, 2011; Easton et al, 2018 and Schölzel-Dorenbos et al, 2012), where each reported both self-report and proxy-report results.

Table 6. Continued

Known-groups validity (bolt-on item)	Known-groups validity (EQ-5D +cognition)	Structural validity (bolt-on item)	Associations with EQ VAS[†] (EQ-5D +cognition)	Responsiveness (bolt-on item)	Responsiveness (EQ-5D +cognition)	Test-retest reliability (bolt-on item)	Inter-rater agreement (bolt-on item)	Total[‖]
					1−			1+/2−
1+/1−	1+		3+					8+/12−
3+/ 1−							1+	4+/3-
								1+/0−
1+/1−								1+/3−
1+								2+/1−
	2+	1−						2+/3-
1−								1+/2−
1−								1+/1−
								0+/1−
2+								4+/0−
2−	2+							6+/3-
								1+/0−
								2+/0−
								1+/0−
4+/1−								4+/3−
3+/10-	1+/1−		2+/1−				1−	16+/23-
5+/5−	5+/0−	0+/1−	3+/0−	0	0+/1−	0	0	27+/27−
0	0	0	0	0	0	0	0	3+/0−
15+/18−	6+/1−	1−	5+/1−	0	1−	0	1+/1−	55+/57−
1+/2−	2+	1−[§]	2+/1−	2−	1−	1−		17+/16−
	1+							5+/0−
1+								1+/1−
								1+/0−
								0+/1−
								1+/1−
			1+					4+/0−
			1+					3+/1−
								1+/0−
								0+/2−
1−								0+/2−
			1+					4+/0−
			1+					4+/0−
1−	1−							4+/4−
2+/3−	3+/0−	0+/1−	2+/1−	0+/2−	0+/1−	0+/1−		26+/22−
0	0	0	4+	0	0	0		15+/1−
2+/4−	3+/1−	1−	6+/1−	2−	1−	1−	0	45+/28−
17+/22−	9+/2−	2−	11+/2−	2−	2−	1−	1+/1−	100+/85−

22−); however, for subdimension items (eg, memory, concentrating, learning, and understanding), the overall evidence was mostly positive (15+/1−). The most psychometric evidence was found for cognition (remembering, understanding, concentrating, and thinking) (A-5L-Co-01: 18+/15−) and thinking ability (remembering and concentrating) bolt-ons (A-3L-Co-03: 8+/12−), with fewer than 10 assessments for all other bolt-ons.

Only 2 patient populations were included in 10 studies each, in which a separate assessment of the psychometric performance of the bolt-ons was possible (see Appendix Tables 8 and 9 in Supplemental Materials. Overall, psychometric evidence was mixed in both patient populations, eg, head/brain injury (3L: 11+/11−; 5L: 1+/3−) and dementia (3L: 9+/8−; 5L: 4+/4−). The mixed results found do not imply that these bolt-ons are unnecessary in these groups, rather, they reflect that the items have not been developed according to established standards, such as those applied to the core instrument, EQ-5D-5L.

Discussion

This systematic review aimed to provide a comprehensive summary of psychometric properties of cognition bolt-ons for the adult versions of the EQ-5D. Key findings indicated that, as expected, 5-level bolt-ons tended to outperform 3-level bolt-ons on tests associated with improved distributional properties. In cases in which evidence was available, bolt-ons demonstrated good performance in terms of divergent validity with the core EQ-5D. Known-groups validity, the measurement property for which the most evidence was available, was established for several relevant groups, including injury-related characteristics, acute and chronic conditions or comorbidities, and sociodemographic characteristics, such as age groups. In most populations, cognition bolt-ons also appeared to improve the associations with the EQ VAS. However, limited evidence was found for any bolt-ons in terms of improving the relative informativity of the EQ-5D, structural validity (eg, loading onto a separate factor from any core 5 items), responsiveness or test-retest reliability, although the small number of publications limited the interpretation of these findings. No qualitative studies assessed the content validity or acceptability of existing bolt-ons, which is concerning because many studies used bolt-ons without confirming their relevance or comprehensibility for the target population.

The demand for cognition bolt-ons is increasing across various applications, including clinical trials, population health surveys, disease burden studies, and cost-utility analyses.[26] Although we found some psychometric evidence for the 28 bolt-on items, the available data are insufficient to identify any as the best-performing candidate. Importantly, the EuroQol Group plans to pursue only 5L cognition bolt-ons in future development and testing.[43] An important direction for future research is to evaluate the psychometric performance of composite (eg, memory/concentration) and subdimension bolt-ons (eg, concentration) in comparison with global ones (eg, cognition with examples).[44] Although our review showed promising performance of subdimension items, the overall evidence was limited compared with global items. The study by Sampson et al (2020), which developed both global and subdimension items and benefited from the most extensive qualitative input, is considered a starting point for future developments, additionally because its source language is English (UK), similar to the core instrument.[44]

A few potential limitations of this systematic review need to be mentioned. First, all studies were included regardless of the relevance of the cognition bolt-on to the target population because we considered cognition to be a general HRQoL area applicable across a broad range of populations. However, the composition of some samples may not have been fully adequate for psychometric testing. Furthermore, we intended to extract data on missing values to assess feasibility, but this was not possible because of lack of reporting, unclear use of imputations and insufficient information on follow-up procedures for missing responses in many instances. The methodological quality of the included studies was not assessed before psychometric evaluation because of the absence of specific checklists tailored for bolt-ons, which are additional items rather than standalone instruments. Psychometric properties were assessed using a checklist developed based on existing bolt-on testing criteria,[24] rather than a consensus-based structured approach. Numerical thresholds were used in multiple checklist items, which are difficult to apply universally to such a heterogeneous set of studies. For example, in item 1a, the 70% ceiling threshold was applied consistently across all bolt-on wordings, versions (3L vs 5L) and populations; yet, it may be too high for certain populations (eg, moderate-to-severe dementia) and too low for others (eg, dental problems). Although adjusting thresholds for individual studies might have offered greater precision, it could have introduced bias due to the researchers' subjective judgment. We relied on subjective judgment for checklist items (eg, adequacy of improvement compared with the EQ-5D) where only a small number of assessments were needed and where our team deemed it inappropriate to set a unified threshold. For future reviews of other bolt-ons, these numerical thresholds may need to be adjusted, and not every checklist item may be equally important. For example, associations with the EQ VAS may be unnecessary for symptom-specific bolt-ons (eg, itching and breathing or gastrointestinal problems).[11,45,46] We applied "relaxed" criteria to assess what we considered to be a form of known-group validity testing—essentially analyses that compared groups, not only studies that reported effect sizes or relative efficiencies, because these were extremely rare. Furthermore, the checklist was not formally validated (beyond the piloting conducted). Publication bias may be present because only 56% of all cognition bolt-on publications reported psychometric results. We attempted to mitigate this bias by including all publications that reported at least 1 psychometric property, even if they only reported distributional properties (eg, item-level ceiling). Lastly, evidence was synthesized narratively, and no meta-analysis was conducted because of substantial heterogeneity in bolt-on wording, reporting, and study populations.

Conclusions

Despite the large number of publications with cognition bolt-ons, the psychometric evidence base remains insufficient to identify a single preferred descriptor. Future research should prioritize the qualitative testing of promising candidate wordings. This should include translatability assessment, a critical step that has been absent in the development of existing bolt-ons. Preferably, quantitative psychometric testing across diverse populations, languages, and contexts should be conducted after a well-established set of items has been identified.

Author Disclosures

Author disclosure forms can be accessed below in the Supplemental Material section. Fanni Rencz, Brendan Mulhern,

Aurelino Paolo Finch, and Mathieu F. Janssen are members of the EuroQol Group. Fanni Rencz and Aureliano Paolo Finch are employed by the EuroQol Research Foundation. Views expressed in the article are those of the authors and are not necessarily those of the EuroQol Research Foundation.

Supplemental Material

Supplementary data associated with this article can be found in the online version at https://doi.org/10.1016/j.jval.2025.09.3053.

Article and Author Information

Accepted for Publication: September 19, 2025

Published Online: xxxx

doi: https://doi.org/10.1016/j.jval.2025.09.3053

Author Affiliations: Department of Health Policy, Corvinus University of Budapest, Budapest, Hungary (Rencz, Pangestu); EuroQol Research Foundation, Rotterdam, The Netherlands (Rencz, Finch); Doctoral School of Business and Management, Corvinus University of Budapest, Budapest, Hungary (Pangestu); Centre for Health Economics Research and Evaluation, University of Technology Sydney, Sydney, Australia (Mulhern); Erasmus MC, Department of Psychiatry, Section Medical Psychology and Psychotherapy, Rotterdam, The Netherlands (Janssen).

Correspondence: Fanni Rencz, DSc, Department of Health Policy, Corvinus University of Budapest, 8 Fővám tér, Budapest 1093, Hungary. Email: fanni.rencz@uni-corvinus.hu

Author Contributions: *Concept and design:* Rencz, Pangestu, Mulhern, Finch, Janssen

Acquisition of data: Rencz, Pangestu, Janssen

Analysis and interpretation of data: Rencz, Pangestu, Mulhern, Finch, Janssen

Drafting the manuscript: Rencz

Critical revision of paper for important intellectual content: Rencz, Pangestu, Mulhern, Finch, Janssen

Obtaining funding*:* Rencz, Pangestu, Mulhern, Finch, Janssen

Dr Mulhern is an editor for *Value in Health* and had no role in the peer-review process of this article.

Funding/Support: This study was supported by the EuroQol Research Foundation (1700-RA).

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review or approval of the manuscript; and decision to submit the manuscript for publication.

Acknowledgment: The authors thank the members of the Descriptive Systems Working Group for their advice on the checklist development.

Data Availability: The data that support the results of this study are available upon reasonable request from the corresponding author.

REFERENCES

1. Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;37(1):53–72.
2. Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res*. 2011;20(10):1727–1736.
3. Kennedy-Martin M, Slaap B, Herdman M, et al. Which multi-attribute utility instruments are recommended for use in cost-utility analysis? A review of national health technology assessment (HTA) guidelines. *Eur J Health Econ*. 2020;21(8):1245–1257.
4. Rencz F, Gulácsi L, Drummond M, et al. EQ-5D in Central and Eastern Europe: 2000-2015. *Qual Life Res*. 2016;25(11):2693–2710.
5. Wang A, Rand K, Yang Z, Brooks R, Busschbach J. The remarkably frequent use of EQ-5D in non-economic research. *Eur J Health Econ*. 2022;23(6):1007–1014.
6. Devlin NJ, Brooks R. EQ-5D and the EuroQol Group: past, present and future. *Appl Health Econ Health Policy*. 2017;15(2):127–137.
7. Shah KK, Mulhern B, Longworth L, Janssen MF. Views of the UK general public on important aspects of health not captured by EQ-5D. *Patient*. 2017;10(6):701–709.
8. Efthymiadou O, Mossman J, Kanavos P. Health related quality of life aspects not captured by EQ-5D-5L: results from an international survey of patients. *Health Policy*. 2019;123(2):159–165.
9. Geraerds AJLM, Bonsel GJ, Janssen MF, Finch AP, Polinder S, Haagsma JA. Methods used to identify, test, and assess impact on preferences of bolt-ons: a systematic review. *Value Health*. 2021;24(6):901–916.
10. Rencz F, Janssen MF. Testing the psychometric properties of 9 bolt-ons for the EQ-5D-5L in a general population sample. *Value Health*. 2024;27(7):943–954.
11. Angyal MM, Janssen MF, Lakatos PL, Brodszky V, Rencz F. The added value of the cognition, dining, gastrointestinal problems, sleep and tiredness bolt-on dimensions to the EQ-5D-5L in patients with coeliac disease. *Eur J Health Econ*. 2025;26(3):473–485.
12. Liao M, Wu H, Yang Z, et al. Testing four cognition bolt-on items to the EQ-5D in a general Chinese population. *Eur J Health Econ*. 2025;26(3):403–411.
13. Igarashi A, Sakata Y, Azuma-Kasai M, et al. Linguistic and psychometric validation of the cognition bolt-on version of the Japanese EQ-5D-5L for the elderly. *J Alzheimers Dis*. 2023;91(4):1447–1458.
14. Geraerds A, Bonsel GJ, Polinder S, Panneman MJM, Janssen MF, Haagsma JA. Does the EQ-5D-5L benefit from extension with a cognitive domain: testing a multi-criteria psychometric strategy in trauma patients. *Qual Life Res*. 2020;29(9):2541–2551.
15. Bahrampour M, Kochovska S, Currow DC, Viney R, Mulhern B. Testing the psychometric characteristics of EQ-5D-5L and respiratory bolt-ons using a sample of the Australian population. *Qual Life Res*. 2025;34(2):395–403.
16. Gandhi M, Ang M, Teo K, et al. A vision 'bolt-on' increases the responsiveness of EQ-5D: preliminary evidence from a study of cataract surgery. *Eur J Health Econ*. 2020;21(4):501–511.
17. Rencz F, Mukuria C, Bató A, Poór AK, Finch AP. A qualitative investigation of the relevance of skin irritation and self-confidence bolt-ons and their conceptual overlap with the EQ-5D in patients with psoriasis. *Qual Life Res*. 2022;31(10):3049–3060.
18. Szlávicz E, Szabó Á, Kinyó Á, et al. Content validity of the EQ-5D-5L with skin irritation and self-confidence bolt-ons in patients with atopic dermatitis: a qualitative think-aloud study. *Qual Life Res*. 2024;33(1):101–111.
19. Müller EM, Nikl A, Krebs M, et al. Psychometric benefits of adding bolt-ons to the EQ-5D-5L in populations undergoing minimally invasive cosmetic procedures. *Eur J Health Econ*. 2025;26:1233–1247.
20. Sussex AK, Rencz F, Gaydon M, Lloyd A, Gallop K. Exploring the content validity of the EQ-5D-5L and four bolt-ons (skin irritation, self-confidence, sleep, social relationships) in atopic dermatitis and chronic urticaria. *Qual Life Res*. 2025;34(4):991–1002.
21. Feng J, Yin Y, Luo N, et al. Performance of the EQ-5D-5L with skin irritation and self-confidence bolt-on items in patients with urticaria. *Value Health*. 2025;28(2):915–922.
22. Hoogendoorn M, Jowett S, Dickens AP, et al. Performance of the EQ-5D-5L plus respiratory Bolt-On in the Birmingham chronic obstructive pulmonary disease cohort study. *Value Health*. 2021;24(11):1667–1675.
23. Plázár D, Metyovinyi Z, Medvecz M, Rencz F. Qualitative evidence on EQ-5D-5L skin irritation and self-confidence bolt-ons in Darier's disease and Hailey-Hailey disease. *Qual Life Res*. 2025;34(4):977–989.
24. Mulhern BJ, Sampson C, Haywood P, et al. Criteria for developing, assessing and selecting candidate EQ-5D bolt-ons. *Qual Life Res*. 2022;31(10):3041–3048.
25. Xu RH, Rencz F, Sun R, Dong D, Zhang S. Development and testing the psychometric properties of 20 bolt-on items for the EQ-5D-5L across 31 rare diseases 2025;. 2025;28(5):769–780.
26. Rencz F, Pangestu S, Mulhern B, Finch AP, Janssen MF. Development and use of cognition bolt-ons for the EQ-5D-3L and EQ-5D-5L: a systematic review. *Value Health*. 2025. https://doi.org/10.1016/j.jval.2025.05.015 [Published online June 7, 2025].
27. NICE health technology evaluations: the manual. NICE Process and Methods. National Institute for Health and Care Excellence. https://www.nice.org.uk/process/pmg36; Published 2023. Accessed November 22, 2024.
28. Gnanasakthy A, Qin S, Norcross L. FDA guidance on selecting, developing, or modifying fit-for-purpose clinical outcome assessments: old wine in a new bottle? *Patient*. 2023;16(1):3–5.
29. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
30. Fischer H. *A History of the Central Limit Theorem: From Classical to Modern Probability Theory*. New York, NY: Springer; 2011.
31. World Health Organization. ICD-11: International Classification of Diseases, 11th revision https://icd.who.int/; Published 2022. Accessed June 23, 2024.
32. Peasgood T, Mukuria C, Brazier J, et al. Developing a new generic health and wellbeing measure: psychometric survey results for the EQ-HWB. *Value Health*. 2022;25(4):525–533.
33. Geraerds A, Bonsel GJ, Janssen MF, et al. The added value of the EQ-5D with a cognition dimension in injury patients with and without traumatic brain injury. *Qual Life Res*. 2019;28(7):1931–1939.

34. Geraerds A, Polinder S, Spronk I, et al. Sensitivity of the EQ-5D-5L for fatigue and cognitive problems and their added value in Q-fever patients. *Qual Life Res*. 2022;31(7):2083–2092.
35. Jia TW, Zhou XN, Wang XH, Utzinger J, Steinmann P, Wu XH. Assessment of the age-specific disability weight of chronic schistosomiasis japonica. *Bull World Health Organ*. 2007;85(6):458–465.
36. Qian Y, Zang X, Li H, et al. Assessment of the health-related quality of life in neurocysticercosis patients in hot spot areas-China, 2017-2018. *China CDC Wkly*. 2022;4(28):618–621.
37. Van Ditshuizen JC, De Munter L, Verhofstad MHJ, et al. Comparing health status after major trauma across different levels of trauma care. *Injury*. 2023;54(3):871–879.
38. Zhang X. 云南省晚期血吸虫病经济负担及其 影响因素研究 [*Economic Burden of Advanced Schistosomiasis and Its Influencing Factors in Yunnan*] (Master's Thesis). Dali University; 2011.
39. Erkan S, Cetinarslan O, Okcu G. Traumatic spinopelvic dissociation managed with bilateral triangular osteosynthesis: functional and radiological outcomes, health related quality of life and complication rates. *Injury*. 2021;52(1):95–101.
40. Gribnau AJ, van Hensbroek PB, Haverlag R, Ponsen KJ, Been HD, Goslings JC. U-shaped sacral fractures: surgical treatment and quality of life. *Injury*. 2009;40(10):1040–1048.
41. Hartgerink D, Bruynzeel A, Eekers D, et al. Quality of life among patients with 4 to 10 brain metastases after treatment with whole-brain radiotherapy vs stereotactic radiotherapy: a phase III, randomized, Dutch multicenter trial. *Ann Palliat Med*. 2022;11(4):1197–1209.
42. Reeves KD, Shaw J, McAdam R, Lam KHS, Mulvaney SW, Rabago D. A novel somatic treatment for post-traumatic stress disorder: a case report of hydrodissection of the cervical plexus using 5% dextrose. *Cureus*. 2022;14(4):e23909.
43. Devlin NJ, Xie F, Slaap B, Stolk E. Measuring and valuing health using EuroQol instruments: new developments 2025 and beyond. *Appl Health Econ Health Policy*. published online July 24, 2025. https://doi.org/10.1007/s40258-025-00989-2.
44. Sampson C, Addo R, Haywood P, et al. Candidate bolt-ons for cognition and vision: qualitative findings from a development programme. *EuroQol Topical DSWG meeting, September 2020 (virtual)*. 2020.
45. Swinburn P, Lloyd A, Boye KS, Edson-Heredia E, Bowman L, Janssen B. Development of a disease-specific version of the EQ-5D-5L for use in patients suffering from psoriasis: lessons learned from a feasibility study in the UK. *Value Health*. 2013;16(8):1156–1162.
46. Hoogendoorn M, Oppe M, Boland MRS, Goossens LMA, Stolk EA, Rutten-van Mölken M. Exploring the impact of adding a respiratory dimension to the EQ-5D-5L. *Med Decis Making*. 2019;39(4):393–404.