Systematic Literature Review

Methods Used to Identify, Test, and Assess Impact on Preferences of Bolt-Ons: A Systematic Review

Alexandra J.L.M. Geraerds, MSc, Gouke J. Bonsel, PhD, Mathieu F. Janssen, PhD, Aureliano Paolo Finch, PhD, Suzanne Polinder, PhD, Juanita A. Haagsma, PhD

A B S T R A C T

Objectives: The question of whether additional dimensions should be added to the EQ-5D, so-called *bolt-ons*, has been researched since the 1990s. Several candidate bolt-ons have been tested. The aim of this systematic review was to provide an overview of EQ-5D bolt-on studies, including the origin of possible suitable bolt-ons, their format, and methods that were used to examine their value.

Methods: Studies were identified through database search and reference screening and assessed based on a set of inclusion criteria. All studies that investigated bolt-ons for the EQ-5D were eligible for inclusion. Two reviewers independently extracted information from all included studies on objectives, study design, EQ-5D version used, the investigated bolt-ons, methods used to achieve objectives, and outcomes.

Results: Of 308 initially identified studies, 28 studies met the inclusion criteria. Of these studies, 3 identified potentially suitable bolt-on dimensions, 13 investigated the psychometric performance of EQ-5D + bolt-on(s), and 6 investigated the impact of the bolt-on on health state preferences. In total, 26 bolt-ons were identified, of which cognition was the most frequently mentioned. A wide variety of bolt-on identification methods, psychometric performance tests, and health state valuation methods were used in the included studies.

Conclusion: A range of bolt-on dimensions has been investigated using diverse methods. Guidelines are needed to standardize the wording of the bolt-on dimension and response options, evaluate minimal important gain of the bolt-on, and facilitate quality assessment of bolt-on studies. Subsequently, guidelines will facilitate decision making on whether or not to implement a bolt-on dimension to the EQ-5D.

Keywords: bolt-on, EQ-5D, HRQoL, systematic review.

Background

Improved healthcare facilities and increased standards of living altered the profile of most diseases of the world. Survival of many diseases has improved as well as health-related quality of life (HRQoL) of patients living with disease.[1] Measuring public health outcomes in general, and the impact of specific diseases in particular, therefore benefits from compound measures of survival and HRQoL.[2] HRQoL measures quantify the impact of health on a person's quality of life.[3] A subclass of HRQoL measures are so-called generic HRQoL measures. A generic measure enables the comparison of HRQoL across health conditions or health interventions.[4] To create a generic measure that is suitable for all populations (eg, healthy/diseased, younger/older), choices have to be made for what aspects of health have to be covered, or operationally translated: what items have to be included.[5]

A widely used generic classification system is the EQ-5D. The EQ-5D was developed at the end of the 20th century with the aim to create a non–disease-specific measurement instrument for describing and valuing HRQoL with a low burden on respondents.[6] Dimensions were selected from existing HRQoL measurement instruments by the members of the EuroQol Group.[6] Initially, this resulted in 6 dimensions, covering 3 main aspects of HRQoL: physical function, social functioning, and mental functioning.[7] The eventually developed EQ-5D consists of 5 different dimensions (mobility, self-care, usual activity, pain/discomfort, and anxiety/depression), and a standardized visual analog scale (EQ VAS).[7] The 5 dimensions are covered by 1 item (question) each. An ordinal response format is offered in 2 versions: "no problems," "some problems," or "unable to/extreme problems" indicated with EQ-5D-3L, and "no problems," "slight problems," "moderate problems," "severe problems," and "unable to/extreme problems"

indicated with EQ-5D-5L.[8] The descriptive profile, which results from self-rating these 5 questions, can be translated into a *utility*, that is, a preference score that is anchored at 0 for death and 1 for perfect health (or no problems on any of the dimensions of the EQ-5D). The EQ VAS asks respondents to rate overall HRQoL on a scale from 0 (worst possible health) to 100 (best possible health), which is essentially a subjective health measure. The EQ VAS can measure a broader underlying construct of health contained in the dimensions, with the associated advantages and disadvantages.[9,10]

Although the EQ-5D is widely used and regarded as valid and responsive in most populations, psychometric shortcomings have been reported, in particular lack of responsiveness in particular populations, such as patients with hearing impairments or patients with mental health problems, which impairs validity and discriminative power.[11-13] One solution to improve accuracy and precision of the EQ-5D was extending the EQ-5D-3L to the EQ-5D-5L, where 5 labeled levels seems the maximum achievable verbal refinement.[14-17] A second solution was adding dimensions, which in this context often are referred to as *bolt-ons*.

The item(s) for a bolt-on can be extracted from existing instruments or other sources (imposing the EQ-format), or created by the researcher, either based on consultation of groups of patients and other key informants or without consultation. The default EQ-5D descriptive item format is as follows: the first 1 or 2 defining terms are used to indicate the dimension the item is about (the heading, eg, mobility or vision), and then the response phrase is stated (with or without supportive examples; self-care, eg, is explained by washing and dressing oneself). Response options can use 3 or 5 ordinal labels, which essentially are intensity rather than frequency based; that is, "some" instead of "sometimes" problems. For instance, the EQ-5D dimension mobility is described as "mobility," with the following response options: I have no problems in walking about; I have slight problems in walking about; I have moderate problems in walking about; I have severe problems in walking about; I am unable to walk about. Examples of bolt-ons that have been analyzed before are cognition, vision, sleep, and a specific dermatological dimension related to psoriasis.[18-21] However, an overview of the origin and format of the bolt-ons that have been studied is missing, as well as an overview of the methods that have been used to study their added value and their impact on health state valuations. Each of these categories are linked, because the addition of 1 or more dimensions to the EQ-5D requires a new valuation to determine the associated utility of the added dimension(s).[13] An overview of all EQ-5D bolt-on studies and their study design, methodology, and results will facilitate the creation of guidelines for the standardization of terminology of the bolt-on and preferred methods to investigate bolt-ons, which in their turn will enable comparison of results among studies.

Therefore, the aim of this systematic review was to provide an overview of EQ-5D bolt-on studies, the format and origin of bolt-ons that were studied, and methods that were used to qualify the bolt-on(s). This review included studies with 1 or multiple of the following objectives: (1) identified or developed possible bolt-ons, (2) assessed the psychometric performance of a bolt-on, and (3) evaluated the impact of the bolt-on on preferences.

Methods

Identification and Selection

A literature search was performed to identify empirical studies that identified or generated possible suitable bolt-on(s), tested their psychometric performance, or evaluated their impact on health state preferences. Four databases were screened using a search strategy developed with a librarian: Embase, PubMed (Medline Ovid), Web of Science, and Google Scholar. Search terms that were used included "bolt-on," "EQ-5D," "EQ-6D," and "additional dimension." Synonyms for the search terms and related terms were also added to the search strategy. The search strategy was adjusted for each database (see Appendix 1 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.12.011). Furthermore, there was no period of publication specified, because the number of studies that have been performed on bolt-ons is limited. The date of the search, December 1, 2020, marked the end date of the period of publication that was included.

The selection of relevant studies was performed by first screening the titles, next screening the abstracts of the articles, and at last screening the full text. The screening procedure was conducted by 2 researchers, independent from each other (A.G. and J.H.). Conflicts in the selected articles were resolved by an expert in this field (S.P.). Reference lists of eligible studies were screened to identify possible additional studies that could be relevant.

Inclusion Criteria

Full-text articles that were published in peer-reviewed journals were included if they investigated the addition of a bolt-on dimension to the EQ-5D, either with or without specification of the term *bolt-on*. There was no limitation set on the nature of the bolt-on dimension, meaning that the bolt-on dimension(s) could have been extracted from, for example, another quality-of-life measurement instrument or created by a patient focus group. However, the studies had to meet the criterion that the bolt-on study compared the EQ-5D with the EQ-5D + bolt-on. If a study solely reported the HRQoL of a population on both the EQ-5D and the EQ-5D + bolt-on, the study was excluded. To mitigate possible publishing bias, we also included conference presentations that were identified through the database search, separately from the full-text studies.

Data Extraction

After determining the relevance of articles, they were categorized according to study design. We used the following prespecified 3 categories to distinguish bolt-on studies:

- Category A: Studies that identified (A1) or developed (A2) potentially suitable bolt-on dimensions, either by literature research or interviews or by comparing the EQ-5D with other HRQoL instruments to identify missing dimensions
- Category B: Studies that investigated the psychometric performance of the EQ-5D + bolt-on(s) using patient/population data
- Category C: Studies that investigated the impact of adding a bolt-on on preferences for a subset of EQ-5D health states

Information was extracted from the included articles and tabulated to gain insight into the study populations of different studies, the version of EQ-5D that was used (3L or 5L), the bolt-on(s) that were investigated/mentioned, and the outcomes. Furthermore, for each bolt-on the frequency of being mentioned/used in a study was determined. In addition, origin of the bolt-on and the terminology that was used for the item name of the bolt-on were reported. Next, for all study categories (category A: identification or development of possible suitable bolt-ons, category B: psychometric performance studies, and category C: health state valuation studies), an overview was created of study design, statistical tests, and health state valuation methods that were used. For conference presentations, information was tabulated separately from the full-text articles because the abstracts

frequently did not report information on study details, such as specific methods used.

Quality Assessment

Currently, there is no guideline on how to assess the methodological quality of bolt-on studies because there is no standard on the psychometrics, identification, and valuation of a bolt-on. However, because it was considered relevant to do some form of quality check, we used 2 checklists. The first checklist consisted of 3 boxes from the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) Risk of Bias Checklist: box 1, patient-reported outcome measure (PROM) development; box 2, content validity; and box 6, reliability, with multiple items per box, which are each rated from "inadequate" to "very good" on a 4-point scale. Some items also provide the answer option "not applicable." The lowest score on an item within a subject (box) determines the overall quality of the subject. More specifically, if for example all items within box 1 are scored "adequate" and 1 item is scored "inadequate" then the overall score for box 1 will be "inadequate."

The second checklist was developed post hoc in this study specifically for quality assessment of bolt-on studies (see Appendix 2 in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.12.011). With this checklist, methodological quality is determined with separate items for category A1 studies (5 items), category A2 studies (4 items), category B studies (6 items), and category C studies (6 items). Each item was classified using a 2- to 4-point scale rating, ranging from "inadequate" to "very good." In addition, a "not applicable" option was added to the response options if a criterion was not applicable to a study in a category. The lowest score on an item determines the overall quality of the subject, in line with the scoring system of the COSMIN Risk of Bias Checklist. Items that were not applicable were not considered in the final score.

A.G. and J.H. assessed the risk of bias independently using the 2 checklists.

Results

Search Strategy

The search yielded 308 unique records (Fig. 1).[22] Title selection excluded 209 articles, because these were not related to the performance of the EQ-5D. Abstract selection removed 68 more articles, because these did not investigate the addition of dimensions to the EQ-5D. Full text selection of the 31 remaining articles led to the inclusion of 28 articles. One article was excluded owing to a lack of reporting on bolt-ons, and 2 articles displayed no results on bolt-ons and were therefore also excluded.

Study Characteristics

Details of the included studies are presented in Table 1. Of the 28 articles that were included, 3 articles were studies that identified possible bolt-ons (category A "identification or development of possible suitable bolt-ons"),[23-25] 13 articles were psychometric performance studies (category B "psychometric performance studies"),[8,26-37] and 6 articles were valuation studies of EQ-5D + bolt-on (category C "health state valuation studies").[19,20,38-42] Furthermore, 3 articles were categorized as both categories A and B,[43-45] and 3 articles as both categories A and C.[13,21,38] Study populations consisted of a healthy population (n = 17 articles: general public,[13,19-21,24,25,27,30,36,38,40-43] random households,[44,45] faculty members[39]) and patients (n = 11 articles: individuals with a chronic condition,[23] visitors of an outpatient eye clinic,[26]

cognitive impaired elderly patients,[18] patients with diabetes,[28] general injury patients,[29,31,35] patients listed for cataract surgery,[32,34] stroke patients,[33] and burn patients[37]). Most of the included studies were cross-sectional survey studies (N = 23),[13,19-21,23-31,33,36-42,44,45] of which 1 study also performed qualitative analysis.[21] In addition, there were 4 longitudinal survey studies[18,32,34,35] and 1 case study.[43] Furthermore, approximately half of the studies analyzed the 5-level version of the EQ-5D (N = 16).

Apart from the full-text peer-reviewed published articles, the search strategy also identified conference presentations on EQ-5D bolt-ons. Nine conference presentations on bolt-on(s) were included, of which 2 presentations were categorized as category A, 3 presentations as category B, 1 presentation as category C, and 3 presentations as both A and B. More detailed information on the conference presentations can be found in Appendix 3 Table in Supplemental Materials found at https://doi.org/10.1016/j.jval.2020.12.011.

Quality of Studies

Overall, only 1 study scored "doubtful" on the COSMIN Risk of Bias checklist, whereas all other studies scored "inadequate." Considering the quality of studies per subject, 1 study scored "adequate" on PROM development,[38] 3 studies scored doubtful,[21,28,42] and 24 studies scored inadequate. Furthermore, content validity was found to be inadequate again in 25 studies and doubtful in the remaining 3 studies.[21,28,45] The third subject that was tested with the COSMIN was reliability, which was found to be inadequate in all studies.

Based on the self-developed checklist for bolt-on studies, 3 studies in category A1 ("studies that identified possible suitable bolt-ons") scored doubtful,[23,24,44] and 3 studies adequate,[25,43,45] whereas for category A2 ("studies that developed possible suitable bolt-ons") 1 study was inadequate[40] and 2 studies were very good.[21,38] For category B, 3 studies scored inadequate,[26,34,36] 6 studies doubtful,[18,27,28,30,32,33] and 4 studies adequate.[29,31,35,37] Furthermore, 1 study in category C was judged inadequate,[19] 3 studies doubtful,[21,38,39] 1 study adequate,[41] and 4 studies very good.[13,20,40,42]

Suggested and Tested Bolt-On Dimensions

The most commonly mentioned or tested bolt-on was a bolt-on for cognition (N = 13). The description of the cognition bolt-on varied between studies, with, for example, 6 studies referring to "cognitive function,"[18,29,31,33,39,45] whereas another study referred to "memory and concentration,"[27] as can be derived from Table 2. The cognition dimension originated from various sources, namely existing HRQoL instruments (N = 4) (eg, Assessment of Quality of Life [AQoL] 8D, 15D, Health Utilities Index Mark 3 [HUI3]), an expert group (N = 1,) and suggested by the general population (N = 1). The main rationale for adding a bolt-on for cognitive functioning was to reduce the ceiling effect of the EQ-5D, both in specific patient populations (eg, injury patients, chronically ill patients, and patients with dementia/mental retardation) and in the general population.[18,23,25,27,29,31,39,42-45] Cognitive functioning was followed by vision (N = 10), relationships (N = 9) and sleep (N = 9), hearing (N = 8), energy (N = 7), tiredness (N = 5), satisfaction (N = 2), and activities related to knee bending (N = 2). Vision and hearing bolt-ons were studied because of the outcomes of a series of systematic reviews, which indicated that performance of the EQ-5D was poor in hearing-related conditions and some specific vision disorders,[20] and were also retrieved from existing HRQoL instruments (eg, HUI3, 15D, AQoL 8D). The next bolt-on is a bolt-on for relationships (N = 9), again with variation

Figure 1. PRISMA flow diagram.
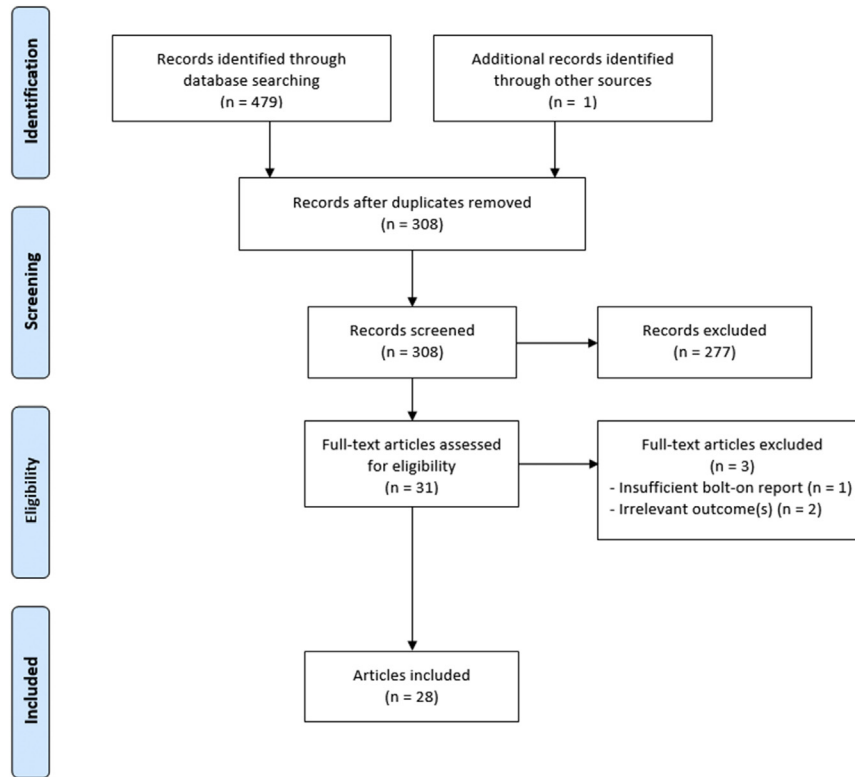
PRISMA indicates Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

in terminology. This bolt-on originated from a focus group, systematic review, in-depth interviews, and existing HRQoL instruments (eg, AQoL 8D, ICEpop CAPability measure, Short Form 6 Dimensions, Personal Wellbeing Index). The main reason for adding this bolt-on was that relationships were expected to be an important part of health that is not measured with the EQ-5D, but is relevant in, for example, chronically ill patients or in some specific cultures.[23,27,28,30,45] Studies that investigated a sleep bolt-on were more consistent in terminology.[19,23,25,27,41,44,45] Furthermore, the sleep bolt-on originated from a literature review and existing HRQoL instruments (eg, AQoL 8D, 15D), with the concern that health is not fully covered within the EQ-5D as main reason for adding a sleep bolt-on. Moreover, previous studies showed that sleep problems affected health.[19] Bolt-ons related to energy were formulated both positively (energy/vitality, N = 7) and negatively (tiredness, N = 5), with some variation in terminology. The origin of these bolt-ons was again existing HRQoL instruments (eg, AQoL 8D) and a systematic review. The rationale for adding an energy/vitality or tiredness bolt-on was again that it was a potentially relevant aspect of health not measured by the EQ-5D.[27,43,44] Furthermore, the lack of an energy dimension was problematic for the face validity of the EQ-5D.[13] Other bolt-ons that were suggested only once were mental health,[24] skin irritation,[21] self-confidence,[21] limitations in physical activities owing to shortness of breath,[38] breathing problems,[38] work limitation,[23] medication side effects,[23] presence of comorbid/long-term conditions,[23] issues with clinicians and social care received,[23] worries/fear about future,[23] financial problems,[23] sensory deprivation,[24] sexual activity,[44] bodily appearance,[44] communication,[45] itching,[37] and community connectedness.[36] Some of the suggested dimensions were not directly health-related (ie, not related to

physical or mental health), but were directing more toward quality of life in general. The format of the bolt-on was not always clearly provided. However, studies that did provide their bolt-on dimension as presented in their questionnaire all complied with the format of the EQ-5D (indicated in Table 2), except for Chen and Olsen, who used full questions.[36]

Psychometric Criteria

Identification of possible bolt-on(s)

A variety of statistical tests was used to determine what bolt-on dimensions might be suitable and should be investigated further, including principal component analysis (PCA) (n = 2 articles) and regression analysis (n = 2 articles). Other tests were used in 1 study only and can be found in Table 3. The study by Finch et al[25] also investigated the potential of a range of techniques in identifying bolt-on dimensions, and concluded that PCA and confirmatory factor analysis appeared useful if the tests used were complementary.

Psychometric performance studies

Table 4 comprises an overview of statistical analyses that were performed to determine the psychometric properties of bolt-on dimensions. All but 5 studies[18,26,32-34] that determined psychometric properties reported the explanatory power of the EQ-5D and the EQ-5D + bolt-on. The second most used psychometric outcome was the distributional effect, which was expressed as the ceiling effect of the EQ-5D and EQ-5D + bolt-on. The studies by Ophuis et al[31] and Spronk et al[37] were unique, because these were the only studies that assessed dimension dependency.

Table 1. Study characteristics and findings of included articles on bolt-on(s) for the EQ-5D-3L/5L.

Author, year, country	Study population	Study design	EQ-5D-3L/-5L	Study category (A/B/C*)	Bolt-on(s) evaluated	Findings
Efthymiadou, 2019, 38 (mainly European) countries[† 23]	Individuals with chronic condition (breast cancer; blood cancer; rheumatoid arthritis; asthma; rare diseases) (N = 67)	Cross-sectional survey study	5L	A1	Fatigue; medication side-effects; presence of other comorbid/long-term conditions; maintenance of relationships and social life; issues with clinicians and social care received; cognitive impairment; sleep deprivation; maintenance of family relationships; worries/fear about future; work limitation; financial problems	51% burden from disease not captured by EQ-5D; 19% fatigue; 12% medication side effects; 9.5% presence of other comorbid/long-term conditions; 6.5% maintenance of relationships and social life; 6.2% issues with clinicians and social care received; 4.3% cognitive impairment; 4.3% sleep deprivation; 3.7% maintenance of family relationships; 3.7% worries/fear about future; 3.7% work limitation; 2% financial problems; 1.9% loss of confidence/self-esteem; 1.8% sexual dysfunction; 1.5% inability to exercise; 1.2% emotional distress; <1% inability to travel; <1% loss of senses (eyesight/hearing)
Finch, 2017, Australia, Canada, Germany, Norway, United Kingdom, United States[25]	Members of the general public (N = 8022)	Exploratory study; cross-sectional survey study	5L	A1	Satisfaction; speech/cognition; relationships; hearing; vision; energy/sleep	Suggested bolt-ons: satisfaction; speech/cognition; relationships; hearing; vision; energy/sleep 40 items loaded on these 6 dimensions with loadings ≥0.45
Shah, 2016, United Kingdom[24]	Broadly representative sample of the UK general public (N = 436)	Valuation study; cross-sectional survey study	5L	A1	Sensory deprivation; mental health	50.7% 11111 on EQ-5D; 83.3% of them <100 on EQ VAS; bolt-on themes identified: absence of illness/unspecified illness; balance; cancer; communication; cardiovascular disease; dexterity; epilepsy; gastroenterological and urological; immune; independence; infertility; lifestyle and fitness; mental health; non-health outcomes; other; relationships; respiratory illness; sensory; sexual function; skin; spirituality; tiredness

Table 1. Continued

Author, year, country	Study population	Study design	EQ-5D-3L/-5L	Study category (A/B/C*)	Bolt-on(s) evaluated	Findings
Finch, 2019, Australia, Canada, Germany, Norway, United Kingdom, United States[43]	Members of the general public (MIC study) (N = 8022)	Case study	5L	A1, B	Energy/vitality; satisfaction; relationships; hearing; vision; speech	Linear regression fitted to determine whether bolt-on explains variation HRQoL: useful for bolt-on selection; linear regression fitted to investigate factors/items help explain negative effect of chronic conditions on HRQoL: further research needed; energy/vitality, relationships, and satisfaction: larger coefficients than speech, vision, hearing
Jelsma, 2015, South Africa[44]	Randomized cluster sample households (N = 310)	Analytical descriptive study; cross-sectional survey study	3L	A1, B	Concentration; energy; sleep; sex life; bodily appearance	Adjusted R^2 EQ-5D: 0.52 Adjusted R^2 EQ-5D + concentration + energy + sleep + Sex life + bodily appearance: 0.57 Adjusted R^2 mobility + pain/discomfort + Anxiety/depression + concentration + sleep: 0.57
Kim, 2017, South Korea[45]	Households from 15 Korean regions (N = 600)	Exploratory study; cross-sectional survey study	5L	A1, B	Vision; hearing; communication; cognitive function; social relationships; vitality; sleep	R^2 EQ-5D: 0.228 R^2 bolt-ons: 0.200 R^2 EQ-5D + bolt-ons: 0.263 Vitality and sleep significantly associated with EQ VAS
Longworth, 2014, United Kingdom[13]	Two representative samples of the UK general population from Yorkshire (N = 600)	Systematic review; modeling study; cross-sectional survey study	3L	A2, C	Hearing; vision; tiredness	Significant impact of hearing, vision, and tiredness on values of health states; direction and magnitude of differences depended on severity of health state; vision statistically significant impact on valuation of EQ-5D health states
Swinburn, 2013, United Kingdom[21]	Members of the UK general public (N = 300)	Review; qualitative analysis; cross-sectional survey study	5L	A2, C	Skin irritation; self-confidence	EQ-5D: R^2 (DLQI): 0.422; R^2 (SAPASI): 0.182 EQ-5D + bolt-ons: R^2 (DLQI): 0.646; R^2 (SAPASI): 0.445
Hoogendoorn, 2019, The Netherlands[38]	Representative sample of Dutch general public (N = 430)	Cross-sectional survey	5L	A2, C	Limitations in physical activities owing to shortness of breath; breathing problems	Limitations in physical activities: utility decrements significant for levels 3 (-0.055), 4 (-0.087), and 5 (-0.135); breathing problems: utility decrements for same levels greater (-0.086; -0.219; -0.327)

Table 1. Continued

Author, year, country	Study population	Study design	EQ-5D-3L/-5L	Study category (A/B/C*)	Bolt-on(s) evaluated	Findings
Breheny, 2020, United Kingdom[32]	Cataract surgery patients (N = 1315)	Longitudinal survey study	3L; 5L	B	Vision	Ceiling effect: EQ-5D-3L: 27.1% EQ-5D-5L: 15.7% EQ-5D-3L + vision: 17.0% Convergent validity with Cat-PROM5: EQ-5D-3L: -0.23 EQ-5D-5L-VSE: -0.31 EQ-5D-5L-CW: -0.29 EQ-5D-3L + vision: -0.42
Chen, 2020, Australia, Canada, Germany, Norway, United Kingdom, United States[36]	Members of the general public (MIC study; NHMS) (N = 8002)	Cross-sectional survey study	5L	B	Vitality, sleep, social relationships, community connectedness	Spearman correlation with VAS: Vitality: 0.600 Sleep: 0.440 Social relationships: 0.368 Community connectedness: 0.407 R^2 for VAS: EQ-5D: 0.434; EQ-5D + bolt-ons: 0.503 R^2 for SWLS: EQ-5D + bolt-ons: 0.447 Shorrocks-Shapely decomposition R^2 for VAS: vitality: 23.0; sleep: 7.8; social relationships: 7.3; community connectedness: 7.4 Shorrocks-Shapely decomposition R^2 for SWLS: vitality: 16.4; sleep: 7.9; social relationships: 24.0; community connectedness: 16.9
Gandhi, 2020, Singapore[34]	Patients due for cataract surgery (N = 63)	Longitudinal survey study	3L, 5L	B	Vision	SRM: EQ-5D-3L: 0.098 EQ-5D-5L(C): 0.249 EQ-5D-5L: 0.207 EQ-5D-3L + V: 0.458 SES: EQ-5D-3L: 0.122 EQ-5D-5L(C): 0.270 EQ-5D-5L: 0.224 EQ-5D-3L + V: 0.509
Geraerds, 2019, The Netherlands[29]	Hospitalized injury patients (N = 2693)	Cross-sectional survey study	3L	B	Cognition	Convergent validity: EQ-5D: all: -0.673; TBI: -0.719; Non-TBI: -0.652 EQ-5D + C: all: -0.690; TBI: -0.736; non-TBI: -0.665 Adjusted R^2: EQ-5D: all: 0.480; TBI: 0.528; Non-TBI: 0.451 EQ-5D + C: all: 0.499; TBI: 0.560; non-TBI: 0.466 Shannon indices: EQ-5D: TBI: H'= 5.08; J'= 0.64; non-TBI: H'= 5.58; J'= 0.70 EQ-5D + C: TBI: H'= 5.88; J'= 0.62; non-TBI: H'= 6.38; J'= 0.67

Table 1. Continued

Author, year, country	Study population	Study design	EQ-5D-3L/-5L	Study category (A/B/C*)	Bolt-on(s) evaluated	Findings
Geraerds, 2020, The Netherlands[35]	Trauma/poisoning patients at ED (N = 1799)	Longitudinal survey study	5L	B	Cognition	Ceiling effect: EQ-5D: 26.3% EQ-5D + cognition: 24.1% Shannon indices: EQ-5D: H': 6.61; J': 0.49 EQ-5D + C: H': 6.37; J': 0.46 Convergent validity: EQ-5D: -0.651 EQ-5D + C: -.664 Explanatory power: EQ-5D: TBI/PTSD: 0.413; none: 0.401 EQ-5D + C: TBI/PTSD: 0.415; none: 0.414 Level sum score: EQ-5D: 8.4 EQ-5D + C: 8.2
De Graaf, 2020, The Netherlands[33]	Stroke patients (N = 360)	Cross-sectional survey study	5L	B	Cognition	Cronbach's alpha: EQ-5D: 0.75 EQ-5D + C: 0.77 Ceiling effect: EQ-5D: 22% EQ-5D + C: 14% Total score: EQ-5D: 78.0 (19.6) EQ-5D + C: 81.8 (16.2) Association with modified Rankin Scale: EQ-5D: -0.66 EQ-5D + C: -0.67
Kangwanrattanakul, 2018, Thailand[28]	Patients with diabetes (N = 200)	Review; cross-sectional survey study	5L	B	Interpersonal relationships (IR); activities related to bending knees (AK)	R^2: EQ-5D: 0.156 R^2: EQ-5D + AK: 0.182 (significant improvement) R^2: EQ-5D + IR: 0.157 (not significant) Ceiling effect: EQ-5D: 36.5%; EQ-5D + bolt-ons: 24.0%
Kangwanrattanakul, 2019, Thailand[30]	Thai people (N = 600)	Cross-sectional survey study	5L	B	Interpersonal relationships (IR); activities related to bending knees (AK)	Ceiling effect: EQ-5D: 50.5%; EQ-5D + bolt-ons: decrease 5% R^2: EQ-5D: 0.306; EQ-5D + AK: 0.329 (sign); EQ-5D + IR: 0.307 (not sign)
Luo, 2015, Singapore[26]	Data from burden of illness study for visual impairment; Patients attending specialist outpatient clinics in Singapore National Eye Centre (N = 836)	Cross-sectional survey study	3L	B	Vision	Index score EQ-5D + vision: 0.93 (0.12); vision: 0.91 (0.15); EQ-5D: 0.96(0.06) Index score 2: EQ-5D + vision: 0.93 (0.12); vision: 0.90 (0.15); EQ-5D: 0.97 (0.05) Vision bolt-on more discriminative than EQ-5D in measurement vision problems
Ophuis, 2019, The Netherlands[31]	Adult injury patients (N = 5346)	Cross-sectional survey study	3L	B	Cognition	R^2: EQ-5D: 0.456; EQ-5D + C: 0.469

Table 1. Continued

Author, year, country	Study population	Study design	EQ-5D-3L/-5L	Study category (A/B/C*)	Bolt-on(s) evaluated	Findings
Perneger, 2011, Switzerland[27]	General resident population of French-speaking Switzerland (N = 1952)	Cross-sectional survey study	3L	B	Sleep; memory/concentration; energy/fatigue; sight/hearing; contact with others	Adjusted R^2: EQ-5D: 0.65 Adjusted R^2: Bolt-ons: 0.65 Adjusted R^2: EQ-5D + bolt-ons: 0.78 Most frequent complaints: fatigue/energy (52.5%); sleep (36.8%); memory and concentration (35.6%); seeing/hearing (28.5%); contact with others (7.0%) 44% of respondents with perfect health on EQ-5D: report problem on bolt-on Correlation bolt-on and EQ-5D: weak-moderate (0.11-0.48)
Spronk, 2020, The Netherlands[37]	Burn patients (N = 243)	Cross-sectional survey	5L	B	Itching	Ceiling effect: EQ-5D: 46.5%; EQ-5D + Itching: 32.1% Shannon indices: EQ-5D: H'=3.64; J'=0.31 EQ-5D + Itching: H'=4.76; J'=0.34 Convergent validity: EQ-5D: -0.587; EQ-5D + itching: -0.507 R^2: EQ-5D 0.490; EQ-5D + Itching: 0.493
Wolfs, 2007, The Netherlands[18]	Data from MEDICIE (N = 196)	Exploratory study; longitudinal survey study	3L	B	Cognition	Correlation EQ-5D(+C) MMSE: - Mobility: -0.02 (not sign.) - Self-care: -0.28 (sign at 0.01 level) - Usual activities: -0.34 (sign. at 0.01 level) - Pain/discomfort: -0.01 (not sign.) - Anxiety/depression: -0.05 (not sign.) - Cognition: -0.35 J (sign at 0.01 level) Responsiveness: Correlation between changes self-care and change score MMSE; correlation between changes in cognition and change score MMSE

Table 1. Continued

Author, year, country	Study population	Study design	EQ-5D-3L/-5L	Study category (A/B/C*)	Bolt-on(s) evaluated	Findings
Finch, 2020, United Kingdom[42]	General population (N = 1040)	Cross-sectional survey study	5L	C	Hearing; sleep; cognition; energy; relationships	Level 1: Nonstatistically significant impact Level 3: Largest to smallest impact: cognition, hearing and sleep, energy, relationships Level 5: Largest to smallest impact: hearing, cognition, relationships, energy, sleep
Kharroubi, 2020, United Kingdom[41]	General population South Yorkshire (N = 160)	Cross-sectional survey	3L	C	Sleep	Random effects with covariates model best; RMSE: EQ-5D: 0.037 EQ-5D + sleep: 0.019 DIC: EQ-5D: 637.5 EQ-5D + sleep: 416.4
Krabbe, 1999, The Netherlands[39]	Faculty members (N = 87)	Cross-sectional survey	3L	C	Cognition	Percentage of variance uniquely attributable to valuation of health states: EQ-5D: 78% EQ-5D + C: 74%
Longworth, 2014, United Kingdom[40]	Members of the general public (N = 300)	Exploratory study; cross-sectional survey study	3L	C	Hearing; vision; tiredness	Extent and direction of impact bolt-ons varied according to level of severity of bolt-ons and severity of state to which added Level 1 bolt-on to mild state → no impact More severe level to mild state → lower TTO value
Yang, 2014, United Kingdom[19]	Members of the public in South Yorkshire (N = 161)	Exploratory study; cross-sectional survey	3L	C	Sleep	Adjusted R^2 EQ-5D: 0.174 Adjusted R^2 EQ-5D with sociodemographic variables: 0.280 Adjusted R^2 EQ-5D + sleep: 0.256 Adjusted R^2 EQ-5D + sleep and sociodemographic variables: 0.344
Yang, 2015, United Kingdom[20]	Members of the UK public (N = 300)	Exploratory study; cross-sectional survey	3L	C	Vision, hearing, tiredness	Significant effect of bolt-ons on health state values of EQ-5D, size, direction, and significance depend on severity of core EQ-5D state and of bolt-on dimension

AK indicates activities related to bending knees; CW, crosswalk; DIC, deviance information criterion; DLQI, Dermatology Life Quality Index; IR, interpersonal relationships; MEDICIE, Maastricht Evaluation of a Diagnostic Intervention for Cognitively Impaired Elderly; MIC, Multi-Instrument Comparison; MMSE, Mini Mental State Examination; NHMS, National Health Measurement Study; PROM, patient-reported outcome measure; PTSD, post-traumatic stress disorder; RMSE, root mean squared error; SAPASI, Self-Administered Psoriasis Area and Severity Index; SRM, standardized response mean; SWLS, satisfaction with life scale; TBI, traumatic brain injury; TTO, time-trade-off; UK, United Kingdom; US, United States; VSE, value set for England.
*A: Study that identified (A1) or developed (A2) possible suitable bolt-on(s). B: Psychometric performance study. C: Health state valuation study of EQ-5D + bolt-on(s).
†Armenia, Australia, Austria, Belgium, Bulgaria, Canada, Croatia, Cyprus, Czech Republic, Denmark, Estonia, France, Germany, Gibraltar, Greece, Hungary, Ireland, Israel, Italy, Japan, Lithuania, F.Y.R.O.M., Malaysia, Malta, Moldova, The Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom, United States.

Table 2. Format and origin of the bolt-on dimension.

Bolt-on	Format	Origin	Study references
Cognition (N = 13)	Cognitive function(ing)	Expert group,[39] existing HRQoL instruments[45]	Geraerds 2019,[29] Ophuis 2019,[31] De Graaf 2020,[33] Kim 2017,[45],* Wolfs 2007,[18] Krabbe 1999[39]
	Cognitive impairment	International patient population with at least 1 chronic condition[23]	Efthymiadou, 2019[23],*
	Speech/cognition	AQoL 8D, 15D, HUI3[25]	Finch 2017,[25],* Finch 2019[43],*
	Concentration	WHOQOL-BREF[44]	Jelsma 2015[44],*
	Memory/concentration	Existing health-utility instruments and general health status questionnaires[27]	Perneger 2011[27]
	Cognition	NA	Geraerds 2020,[35],* Finch 2020[42]
Vision (N = 10)	Vision	HUI3, 15D, AQoL 8D[25]; existing HRQoL instruments[45]; Systematic review of performance EQ-5D[13,20]	Finch 2017,[25],* Finch 2019,[43],* Kim 2017,[45],* Longworth 2014,[13] Gandhi 2020,[34] Luo 2015,[26] Longworth 2014,[40] Yang 2015[20]
	Seeing	NA	Breheny 2020[32]
	Seeing and hearing	Existing health-utility instruments and general health status questionnaires[27]	Perneger, 2011[27]
Relationships (N = 9)	Maintenance of (family) relationships and social life	NA	Efthymiadou 2019[23],*
	Relationships	AQoL 8D, ICECAP, SF-6D, PWI[25]	Finch 2017,[25],* Finch 2019,[43],* Finch 2020[42]
	Social relationships	Existing HRQoL instruments[45] AQoL 8D (Chen)	Kim 2017,[45],* Chen 2020[36]
	Interpersonal relationships	In-depth interviews and literature review[28]; systematic review, in-depth interview, focus group[30]	Kangwanrattanakul 2018,[28] Kangwanrattanakul 2019[30]
	Contacts with others	Existing health-utility instruments and general health status questionnaires[27]	Perneger 2011[27]
Sleep (N = 9)	Sleep deprivation	NA	Efthymiadou 2019[23],*
	Energy/sleep	AQoL 8D, 15D[25]	Finch 2017[25],*
	Sleep	Literature review[19] AQoL 8D (Chen)	Jelsma 2015,[44],* Kim 2017,[45],* Perneger 2011,[27] Kharroubi 2020,[41] Yang 2014,[19] Chen 2020,[36] Finch 2020[42]
Hearing (N = 8)	Hearing	HUI3, 15D, AQoL 8D[25]; existing HRQoL instruments[45]; systematic review of performance EQ-5D[13,20]	Finch 2017,[25],* Finch 2019,[43],* Kim 2017,[45],* Longworth 2014,[13] Longworth 2014,[40] Yang 2015,[20] Finch 2020[42]
	Seeing and hearing	Existing health-utility instruments and general health status questionnaires[27]	Perneger 2011[27]
Energy (N = 7)	Energy/sleep	AQoL 8D[25]	Finch 2017[25],*
	Energy/vitality	NA	Finch 2019[43],*
	Energy	NA	Jelsma 2015,[44],* Finch 2020,[42] Chen 2020[36]
	Fatigue/energy	Existing health-utility instruments and general health status questionnaires[27]	Perneger 2011[27]
Tiredness (N = 5)	Fatigue	NA	Efthymiadou 2019[23],*
	Tiredness	Systematic review of performance EQ-5D[13,20]	Longworth 2014,[13] Longworth 2014,[40] Yang 2015[20]
	Fatigue/energy	Existing health-utility instruments and general health status questionnaires[27]	Perneger 2011[27]
Satisfaction (N = 2)	Satisfaction	HUI3, AQoL 8D, PWI, SWLS[25]	Finch 2017,[25],* Finch 2019[43],*
Activities related to knee bending (N = 2)	Activities related to knee bending	In-depth interviews and literature review[28]; systematic review, in-depth interview, focus group[30]	Kangwanrattanakul 2018,[28] Kangwanrattanakul 2019[30]

AQoL 8D indicates Assessment of Quality of Life; NA, not applicable; ICECAP, ICEpop CAPability measure; PWI, Personal Wellbeing Index; SWLS, Satisfaction With Life Scale; WHOQOL-BREF, World Health Organization Quality of Life short version.
*Exact wording not available, referred to bolt-on term (eg, vision bolt-on, no specific question provided).

Table 3. Statistical tests to identify potential bolt-on dimensions.

	Exploratory factor analysis (EFA)	Principal component analysis (PCA)	Confirmatory factor analysis (CFA)	Content analysis	Multivariate statistical analyses	Regression analysis	Discrete choice experiment (DCE)
Finch, 2017[25]	X	X	X	X	X	-	-
Finch, 2019[40]	-	-	-	-	-	X	-
Hoogendoorn, 2019[36]		X	-	-	-	X	X

Health state valuation studies

The studies that investigated the impact of the bolt-on dimension on preferences (category C) comprised a large variety of measures for valuation of the EQ-5D + bolt-on and measures to capture health. Table 5 provides an overview of measures that were used per study, including a specification of the study population and the number of respondents. The 2 most commonly used measures were the EQ VAS and the time trade-off (TTO). The study population of most of the studies (14/24) consisted of a general population. Wolfs et al[18] compared the EQ-5D and EQ-5D + bolt-on with the Mini Mental State Examination, which is used to detect cognitive impairment, to determine correlation between the 2 instruments. Breheny et al[32] compared the EQ-5D-3L and -5L and EQ-5D-3L + bolt-on with the Cat-PROM5, which is used for self-evaluation for cataract surgery, and Gandhi et al[34] compared the EQ-5D and EQ-5D + bolt-on to the Short Form 6 Dimensions and the HUI3.

Discussion

This systematic review aimed to provide an overview of EQ-5D bolt-on studies, the format and origin of bolt-ons that were studied, and methods that were used to study the bolt-on(s). Of the included studies, 3 studies identified or developed potentially suitable bolt-on dimensions, 13 studies investigated the psychometric performance of EQ-5D + bolt-on(s), and 6 studies assessed the impact of the bolt-on on preferences for a subset of health states. Most frequently tested or mentioned bolt-ons were cognitive function and vision. The bolt-ons originated from diverse sources, such as existing instruments, literature reviews, or interviews with patient groups or the general population. Furthermore, we found that a wide variety of bolt-on identification methods, psychometric performance tests, and health state valuation methods were used in the studies that were included in our review. An overview was provided of all methods used per study type, which may guide future studies to identify commonly used methods to test bolt-ons. According to the COSMIN risk of bias checklist, the quality of the included studies was "inadequate" for most included studies, whereas according to the self-developed checklist for bolt-on studies the quality of the studies ranged from doubtful to very good. However, classifying existing studies with a set of criteria is difficult because bolt-on research is at an early stage and the included studies focused on different aspects of the steps that are required in the process of identification, development, and evaluation of possible suitable bolt-ons.

The frequency that a certain possible suitable bolt-on dimension is investigated gives an indication of its degree of importance. However, this does not mean that it is the most important bolt-on for the measurement of HRQoL. A bolt-on can be mentioned or investigated more frequently than other bolt-ons because it was the first bolt-on that was developed, as is the case with cognition. The first bolt-on study investigated the cognition bolt-on and was published in 1999,[39] and the cognition bolt-on has since been

further investigated in many other studies. Regarding the other identified bolt-ons that were studied at least twice, all bolt-ons were also reported to be present in at least one other HRQoL measurement instrument, except for "activities related to knee bending." The bolt-ons that were studied only once also included aspects of quality of life in general, rather than HRQoL. These bolt-ons are in line with the E-QALY initiative, a newly developed QoL instrument with the same core items as the EQ-5D, but expanded with more non-health-related bolt-ons, such as relationships.[46]

Our findings showed that particularly regarding the cognitive function bolt-on, there was a diversity in terminology of the wording of the bolt-on. The cognitive bolt-on was referred to as "cognition," but also as "cognitive function" or "memory/concentration." A possible explanation for this diversity in wording may be the different sources from which the cognitive bolt-on originated. Moreover, unlike single construct concepts such as hearing, cognition is a multi-construct domain and possibly the included studies may have focused on different constructs of cognition, resulting in diversity in terminology. This diversity in terminology may hamper comparability among studies that investigated the cognitive functioning bolt-on. To prevent variation in wording, we suggest that future guidelines on bolt-on studies include a listing of the descriptors of formerly investigated bolt-on dimensions and provide guidelines regarding standardization of the terminology of the descriptors of bolt-ons.

In the 3 studies that aimed to identify possible bolt-on dimensions, explanatory power and PCA were used in 2 studies and therefore were the methods that were used most often. For both psychometric performance and valuation studies, the findings of this review showed that there is large variation regarding the methods that were used to investigate the bolt-on dimension(s), which impedes a comparison of the performance of bolt-ons. Uniformity in methods may be achieved by developing guidelines that advise on the preferred methods to investigate bolt-ons for each category of bolt-on study. This will also facilitate the comparison of results of bolt-on studies and understanding the relevance of the bolt-on.

Another aspect that is important to take in to account when evaluating the relevance of a bolt-on is the context in which it is tested. Some bolt-ons may be only relevant in a specific patient population (eg, the psoriasis bolt-on), whereas other bolt-ons are more relevant in certain cultures (eg, activities related to knee bending).

Furthermore, most studies refrained from drawing a conclusion on whether or not to implement the investigated bolt-on and advised further research before implementation. The study by Jelsma and Maart[44] stressed that additional explanatory power was found with the inclusion of bolt-ons, but that it is up to the user to weigh the advantage of adding the additional dimensions versus using the EQ-5D for which a summary index can be calculated. So far, only 1 bolt-on, for psoriasis, has reached the "beta status," whereas all other bolt-ons are "experimental versions."[21] Beta status means that the EQ-5D version, with bolt-on in this case, has not been accepted yet as an official version of the

Table 4. Psychometric tests in comparison studies.

	Bolt-on	Distributional effect	Discriminatory power	Dimension dependency	Regression analysis	Convergent validity	Construct validity	Responsiveness
Breheny, 2020[32]	Vision	X	-	-	-	X	X	X
Chen, 2020[36]	Vitality; sleep; relationships; energy; cognition	-	-	-	X	X	-	-
Finch, 2019[43]	Energy/vitality; satisfaction; relationships; hearing; vision; speech	-	-	-	X	-	-	-
Gandhi, 2020[34]	Vision	X	-	-	-	-	-	X
Geraerds, 2019[29]	Cognition	X	X	-	X	X	-	-
Geraerds, 2020[35]	Cognition	X	X	-	X	X	X	X
De Graaf, 2020[33]	Cognition	X	X	-	-	-	-	-
Jelsma, 2015[44]	Concentration; energy; sleep; sex life; bodily appearance	-	-	-	X	X	-	-
Kangwanrat-tanakul, 2018[28]	Interpersonal relationships (IR); activities related to bending knees (AK)	X	-	-	X	-	-	-
Kangwanrat-tanakul, 2019[30]	Interpersonal relationships (IR); activities related to bending knees (AK)	X	-	-	X	X	-	-
Kim, 2017[45]	Vision; hearing; communication; cognitive function; social relationships; vitality; sleep	-	-	-	X	-	-	-
Luo, 2015[26]	Vision	-	X	-	-	-	-	-
Ophuis, 2019[31]	Cognition	X	-	X	X	-	-	-
Perneger, 2011[27]	Sleep; memory/concentration; energy/fatigue; sight/hearing; contact with others	X	-	-	X	-	-	-
Spronk, 2020[37]	Itching	X	X	X	X	X	X	-
Swinburn, 2013[21]	Skin irritation; self-confidence	-	-	-	X	-	-	-
Wolfs, 2007[18]	Cognition	-	-	-	-	-	X	X
Yang, 2014[19]	Sleep	-	-	-	X	-	-	-

Table 5. Health state valuation study characteristics.

	Bolt-on	Discrete choice experiment	EQ VAS	Time trade-off (TTO)	Multiple classification analysis (MCA)	HRQoL instrument used for comparison	General population	N
Breheny, 2020[32]	Vision	-	-	-	-	Cat-PROM5	Cataract surgery patients	1315
Chen, 2020[36]	Vitality; sleep; social relationships; community connectedness	-	X	-	-	-	X	8002
Finch, 2019[43]	Energy/vitality; satisfaction; relationships; hearing; vision; speech	-	X	-	-	-	X	8022
Finch, 2020[42]	Hearing; sleep; relationships; energy; cognition	X	-	X	-	-	X	1040
Gandhi, 2020[34]	Vision	-	-	-	-	HUI3, SF-6D	Cataract surgery patients	63
Geraerds, 2019[29]	Cognition	-	X	-	-	-	Injury patients	2693
Geraerds, 2020[35]	Cognition	-	X	-	-	-	Injury patients	1799
Hoogendoorn, 2019[38]	Limitations in physical activities owing to shortness of breath; Breathing problems	X	-	X	-	-	X	430
Jelsma, 2015[44]	Concentration; energy; sleep; sex life; bodily appearance	-	X	-	-	-	X	310
Kangwanrattanakul, 2018[28]	Interpersonal relationships (IR); activities related to bending knees (AK)	-	X	-	-	-	Thai patients with diabetes	200
Kangwanrattanakul, 2019[30]	Interpersonal relationships (IR); activities related to bending knees (AK)	-	X	-	-	-	X	600
Kharroubi, 2020[41]	Sleep	-	-	X	-	-	X	160
Krabbe, 1999[39]	Cognition	-	-	-	X	-	Faculty members	87
Kim, 2017[45]	Vision; hearing; communication; cognitive function; social relationships; vitality; sleep	-	X	-	-	-	X	600
Longworth, 2014[40]	Hearing; vision; tiredness	-	-	X	-	-	X	300
Longworth, 2014[13]	Hearing; vision; tiredness	-	-	X	-	-	X	600
Luo, 2015[26]	Vision	-	-	-	-	-	Individuals with and without visual impairment	836
Ophuis, 2019[31]	Cognition	-	X	-	-	-	Injury patients	5346
Perneger, 2011[27]	Sleep; memory/concentration; energy/fatigue; sight/hearing; contact with others	-	X	-	-	-	X	1952
Spronk, 2020[37]	Itching	-	X	-	-	-	Burn patients	243
Swinburn, 2013[21]	Skin irritation; self-confidence	-	-	X	-	-	X	300
Yang, 2014[19]	Sleep	-	-	X	-	-	X	161
Yang, 2015[20]	Vision, hearing, tiredness	-	-	X	-	-	X	300
Wolfs, 2007[18]	Cognition	-	X	-	-	MMSE	Cognitively impaired elderly	196

EQ VAS indicates EuroQol visual analog scale; HUI3, Health Utilities Index 3; MMSE, Mini Mental State Examination; PROM, patient-reported outcome measure; SF-6D, Short Form 6 Dimensions.

EQ-5D because the body of evidence was not sufficient but is a work in progress.[47] Furthermore, beta status is one step ahead of the experimental versions, which are developed to for the purpose of methodological testing. Possibly, the lack of consensus or guidelines on what the minimal gains of the bolt-on in psychometric performance should be or the minimal effect of the bolt-on on health state preferences in accepting a bolt-on dimension results in indistinctness in determining when a bolt-on dimension should be included. We therefore recommend conducting research on this minimal gain or effect. Furthermore,

we recommend developing guidelines on methods to identify, test, and evaluate possible bolt-ons. Such guidelines would also support the creation of a minimal gain and improve comparability among studies.

There is a work in progress study that developed a set of criteria for the development, testing, and selection of candidate descriptors for bolt-ons for vision and cognition for the EQ-5D-5L.[48] This could be a first step in the creation of general guidelines for bolt-on development. The work in progress study does not include criteria for valuation of the candidate bolt-on items. When evaluating a bolt-on, one should take aspects such as the purpose the bolt-on serves, criteria for the development, testing and selection, usefulness, and minimal gain into consideration, but also the impact of the bolt-on on preferences and the relevance of valuation studies of the EQ-5D + bolt-on(s). To calculate a summary index for each of the bolt-ons that is considered to be implemented, value sets for the EQ-5D + bolt-on(s) need to be created. This is most likely not feasible considering the costs of performing a valuation study according to the EuroQol guidelines. Possibly, more cost-efficient options to derive value sets for EQ-5D + bolt-ons may be considered.

Overall, we found that all 3 categories of bolt-on studies are relevant in the development of bolt-ons, and that one study type should be inevitably connected to the 2 other study types. After identifying/developing a bolt-on dimension, it is crucial to assess the impact of the bolt-on on the psychometric performance and assess the impact on preferences. Therefore, it would be helpful, in addition to guidelines on bolt-on research, to create a database where bolt-ons can be registered and studies can be linked. This will facilitate researchers with a baseline of what is already known about a specific bolt-on and assure consistency in terminology, and can help to connect identification studies with psychometric performance studies and studies on preferences.

The quality assessment of the selected studies indicated that the COSMIN Risk of Bias checklist, which is developed for the quality assessment of studies that developed a PROM, appeared unsuitable for bolt-on studies. As bolt-on studies are usually not about developing a new PROM, but more on extending an existing PROM, many quality criteria of the Risk of Bias checklist did not apply to the studies of interest. This resulted in a score of "inadequate" for most studies, making it impossible to distinguish in quality. Our self-developed quality assessment checklist for bolt-on studies seemed to be able to better distinguish bolt-on studies with lower quality and those with high quality. Guidelines on preferred methods to identify, test, and evaluate possible bolt-ons may allow for an even more tailored quality assessment of bolt-on studies.

To our knowledge, this is the first systematic review on bolt-on studies for the EQ-5D-3L and -5L that covers all types of bolt-on studies and also provides an overview of statistical tests that were used in past studies. An important limitation of this systematic review is that important papers on bolt-on research might be missing, because only peer-reviewed published work was included in the review. Possibly, there are unpublished working papers on bolt-on research that were not retrieved by the search in the selected databases. Moreover, possibly bolt-on studies that showed no added value for the implementation of a bolt-on might not get published, which may lead to publication bias. However, we have attempted to mitigate for this bias by reporting conference presentations that were identified through the database search. Findings of all but 1 conference abstract showed additional value for adding bolt-on dimensions to the EQ-5D, which could indicate that there is no severe publication bias.

In conclusion, multiple bolt-on studies have been performed on a range of possible bolt-on dimensions. Regarding bolt-on

dimensions that have been tested, some studies showed that possible bolt-on dimensions had added value in a descriptive or valuation sense. However, guidelines are needed to standardize the wording of bolt-on descriptors; facilitate uniformity in methods that are used to identify, test, and evaluate bolt-ons; inform on a minimal gain or effect; and facilitate quality assessment of bolt-on studies. This will enable decision making on whether or not to implement a bolt-on dimension to the EQ-5D.

Supplemental Material

Article and Author Information

Author Affiliations: Erasmus MC, University Medical Center Rotterdam, Department of Public Health, The Netherlands (Geraerds, Bonsel, Polinder, Haagsma); Section Medical Psychology and Psychotherapy, Department of Psychiatry, Erasmus MC, Rotterdam, The Netherlands (Janssen); EuroQol Research Foundation, Rotterdam, The Netherlands (Bonsel, Janssen, Finch).

Correspondence: Alexandra J.L.M. Geraerds, MSc, Erasmus University Medical Center, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands. Email: a.geraerds@erasmusmc.nl

REFERENCES

1. van Oostrom SH, Gijsen R, Stirbu I, et al. Time trends in prevalence of chronic diseases and multimorbidity not only due to aging: data from General Practices and Health Surveys. *PLoS One*. 2016;11(8):e0160264.
2. Polinder S, Haagsma JA, Belt E, et al. A systematic review of studies measuring health-related quality of life of general injury populations. *BMC Public Health*. 2010;10:783.
3. Brazier J, Ratcliffe J, Salomon JA, Tsuchiya A. *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford: Oxford University Press; 2007.

4. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med*. 1993;118(8):622–629.

5. Lerner M. Conceptualization of health and social well-being. *Health Serv Res*. 1973;8(1):6–12.

6. EuroQol G. EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16(3):199–208.

7. Brooks R. EuroQol: the current state of play. *Health Policy*. 1996;37(1):53–72.

8. Janssen MF, Bonsel GJ, Luo N. Is EQ-5D-5L better than EQ-5D-3L? A head-to-head comparison of descriptive systems and value sets from seven countries. *Pharmacoeconomics*. 2018;36(6):675–697.

9. de Boer AG, van Lanschot JJ, Stalmeier PF, et al. Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Qual Life Res*. 2004;13(2):311–320.

10. Feng Y, Parkin D, Devlin NJ. Assessing the performance of the EQ-VAS in the NHS PROMs programme. *Qual Life Res*. 2014;23(3):977–989.

11. Qian X, Tan RL, Chuang LH, Luo N. Measurement properties of commonly used generic preference-based measures in East and South-East Asia: a systematic review. *Pharmacoeconomics*. 2020;38(2):159–170.

12. Finch AP, Brazier JE, Mukuria C. What is the evidence for the performance of generic preference-based measures? A systematic overview of reviews. *Eur J Health Econ*. 2018;19(4):557–570.

13. Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. *Health Technol Assess*. 2014;18(9):1–224.

14. Buchholz I, Thielker K, Feng YS, Kupatz P, Kohlmann T. Measuring changes in health over time using the EQ-5D 3L and 5L: a head-to-head comparison of measurement properties and sensitivity to change in a German inpatient rehabilitation sample. *Qual Life Res*. 2015;24(4):829–835.

15. Janssen MF, Buchholz I, Golicki D, Bonsel G. Is EQ-5D-5L better over time? A head-to-head comparison of responsiveness of the EQ-5D-3L and EQ-5D-5L. Presented at: 36th EuroQol Plenary Meeting 2019; 2019; Brussels, Belgium.

16. Golicki D, Niewada M, Karlinska A, et al. Comparing responsiveness of the EQ-5D-5L, EQ-5D-3L and EQ VAS in stroke patients. *Qual Life Res*. 2015;24(6):1555–1563.

17. Jia YX, Cui FQ, Li L, et al. Comparison between the EQ-5D-5L and the EQ-5D-3L in patients with hepatitis B. *Qual Life Res*. 2014;23(8):2355–2363.

18. Wolfs CA, Dirksen CD, Kessels A, Willems DC, Verhey FR, Severens JL. Performance of the EQ-5D and the EQ-5D+C in elderly patients with cognitive impairments. *Health Qual Life Outcomes*. 2007;5:33.

19. Yang Y, Brazier J, Tsuchiya A. Effect of adding a sleep dimension to the EQ-5D descriptive system: a "bolt-on" experiment. *Med Decis Making*. 2014;34(1):42–53.

20. Yang Y, Rowen D, Brazier J, Tsuchiya A, Young T, Longworth L. An exploratory study to test the impact on three "bolt-on" items to the EQ-5D. *Value Health*. 2015;18(1):52–60.

21. 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.

22. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.

23. 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.

24. Shah K, Mulhern B, Longworth L, Janssen B. *Important Aspects of Health Not Captured by EQ-5D: Views of the UK General Public*. London: 2016.

25. Finch AP, Brazier JE, Mukuria C, Bjorner JB. An exploratory study on using principal-component analysis and confirmatory factor analysis to identify bolt-on dimensions: the EQ-5D case study. *Value Health*. 2017;20(10):1362–1375.

26. Luo N, Wang X, Ang M, et al. A vision "bolt-on" item could increase the discriminatory power of the EQ-5D Index Score. *Value Health*. 2015;18(8):1037–1042.

27. Perneger TV, Courvoisier DS. Exploration of health dimensions to be included in multi-attribute health-utility assessment. *Int J Qual Health Care*. 2011;23(1):52–59.

28. Kangwanrattanakul K, Gross CR, Sunantiwat M, Thavorncharoensap M. Adding two culture-specific 'bolt-on' dimensions on the Thai version of EQ-5D-5L: an exploratory study in patients with diabetes. *Expert Rev Pharmacoecon Outcomes Res*. 2019;19(3):321–329.

29. 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.

30. Kangwanrattanakul K, Gross CR, Sunantiwat M, Thavorncharoensap M. Exploration of a cultural-adaptation of the EQ-5D for Thai population: a "bolt-on" experiment. *Qual Life Res*. 2019;28(5):1207–1215.

31. Ophuis RH, Janssen MF, Bonsel GJ, Panneman MJ, Polinder S, Haagsma JA. Health-related quality of life in injury patients: the added value of extending the EQ-5D-3L with a cognitive dimension. *Qual Life Res*. 2019;28(7):1941–1949.

32. Breheny K, Hollingworth W, Kandiyali R, et al. Assessing the construct validity and responsiveness of Preference-Based Measures (PBMs) in cataract surgery patients. *Qual Life Res*. 2020;29(7):1935–1946.

33. de Graaf JA, Kuijpers M, Visser-Meily J, Kappelle LJ, Post M. Validity of an enhanced EQ-5D-5L measure with an added cognitive dimension in patients with stroke. *Clin Rehabil*. 2020;34(4):545–550.

34. 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.

35. 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.

36. Chen G, Olsen JA. Filling the psycho-social gap in the EQ-5D: the empirical support for four bolt-on dimensions. *Qual Life Res*. 2020;29(11):3119–3129.

37. Spronk I, Bonsel GJ, Polinder S, van Baar ME, Janssen MF, Haagsma JA. The added value of extending the EQ-5D-5L with an itching item for the assessment of health-related quality of life of burn patients: an explorative study [e-pub ahead of print]. *Burns*. 2020 Sep 12. https://doi.org/10.1016/j.burns.2020.08.015.

38. Hoogendoorn M, Oppe M, Boland MRS, Goossens LMA, Stolk EA, Rutten-van Molken M. Exploring the impact of adding a respiratory dimension to the EQ-5D-5L. *Med Decis Making*. 2020;40(3):393–404.

39. Krabbe PF, Stouthard ME, Essink-Bot ML, Bonsel GJ. The effect of adding a cognitive dimension to the EuroQol multiattribute health-status classification system. *J Clin Epidemiol*. 1999;52(4):293–301.

40. Longworth L, Yang Y, Young T, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: a systematic review, statistical modelling and survey. Developing 'bolt-on' items to EQ-5D. *Health Technol Assess*. 2014.

41. Kharroubi SA, Beyh YS, Brazier J, Yang Y. Modelling a preference-based index for EQ-5D-3L and EQ-5D-3L + Sleep using a Bayesian framework. *Qual Life Res*. 2020;29(6):1495–1507.

42. Finch AP, Brazier J, Mukuria C. Selecting bolt-on dimensions for the EQ-5D: testing the impact of hearing, sleep, cognition, energy, and relationships on preferences using pairwise choices. *Med Decis Making*. 2021;41(1):89–99.

43. Finch AP, Brazier JE, Mukuria C. Selecting bolt-on dimensions for the EQ-5D: examining their contribution to health-related quality of life. *Value Health*. 2019;22(1):50–61.

44. Jelsma J, Maart S. Should additional domains be added to the EQ-5D health-related quality of life instrument for community-based studies? An analytical descriptive study. *Popul Health Metr*. 2015;13:13.

45. Kim SH, Jo MW, Ock M, Lee SI. Exploratory study of dimensions of health-related quality of life in the general population of South Korea. *J Prev Med Public Health*. 2017;50(6):361–368.

46. Extending the QALY. https://scharr.dept.shef.ac.uk/e-qaly/about-the-project/. The University of Sheffield. Accessed September 25, 2020.

47. EuroQol Group. Intellectual property (IP) protection. https://euroqol.org/eq-5d-instruments/ip-protection/. Accessed September 25, 2020.

48. Haywood P, Addo R, Reardon O, et al. Generating criteria for the development and selection of candidate bolt-on dimensions. Paper presented at: EuroQol Early Career Researcher Meeting. March 2020; Prague, Czech Republic.