The Patient - Patient-Centered Outcomes Research (2021) 14:223–240
https://doi.org/10.1007/s40271-020-00451-6

ORIGINAL RESEARCH ARTICLE

Empirical Validity of a Generic, Preference-Based Capability Wellbeing Instrument (ICECAP-A) in the Context of Spinal Cord Injury

Cassandra Mah[1,2] · Vanessa K. Noonan[3,4] · Stirling Bryan[2,5] · David G. T. Whitehurst[1,2,6]

Published online: 27 September 2020

Abstract

Background Assessing the validity of generic instruments across different clinical contexts is an important area of methodological research in economic evaluation and outcomes measurement.

Objective Our objective was to examine the empirical validity of a generic, preference-based capability wellbeing instrument (ICECAP-A) in the context of spinal cord injury.

Methods This study consisted of a secondary analysis of data collected using an online cross-sectional survey. The survey included questions regarding demographics, injury classifications and characteristics, secondary health conditions, quality of life and wellbeing, and functioning in activities of daily living. Analysis comprised the descriptive assessment of Spearman's rank correlations between item-/dimension-level data for the ICECAP-A and four preference-based health-related quality of life (HRQoL) instruments, and discriminant and convergent validity approaches to examine 21 evidence-informed or theoretically derived constructs. Constructs were defined using participant and injury characteristics and responses to a range of health, wellbeing and functioning outcomes.

Results Three hundred sixty-four individuals completed the survey. Mean index score for the ICECAP-A was 0.761; 12 (3%) individuals reported full capability (upper anchor; score = 1), and there were no reports of zero capabilities (lower anchor; score = 0). The strongest correlations were dominated by items and dimensions on the comparator (HRQoL) instruments that are non-health aspects of quality of life, such as happiness and control over one's life (including self-care). Of 21 hypothesised constructs, 19 were confirmed in statistical tests, the exceptions being the exploratory hypotheses regarding education and age at injury.

Conclusion The ICECAP-A is an empirically valid outcome measure for assessing capability wellbeing in people with spinal cord injury living in a community setting. The extent to which the ICECAP-A provides complementary information to preference-based HRQoL instruments is dependent on the comparator.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40271-020-00451-6) contains supplementary material, which is available to authorized users.

✉ David G. T. Whitehurst
david_whitehurst@sfu.ca

1 Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

2 Centre for Clinical Epidemiology and Evaluation, Vancouver Coastal Health Research Institute, Vancouver, BC, Canada

3 Praxis Spinal Cord Institute, Vancouver, BC, Canada

4 Blusson Spinal Cord Centre, Vancouver, BC, Canada

5 School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

6 International Collaboration on Repair Discoveries (ICORD), Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada

Key Points for Decision Makers

Supportive evidence is accumulating for the validity of a preference-based capability wellbeing instrument (ICECAP-A) in different clinical contexts. This study found positive evidence among individuals living with a spinal cord injury (70% injured for 10 years or more).

Relationships between the ICECAP-A and different health-related measures indicate uncertainty about what underlying constructs (capabilities and/or functionings) are being measured.

Preference-based outcomes for economic evaluation is an increasingly crowded market. Further debate and dialogue are needed, with multiple stakeholders, regarding the evaluative space of economic evaluation.

1 Introduction

Generic outcome measures play an important role in health research, providing a means of assessing outcomes in a way that allows comparisons to be drawn across different contexts. One area where generic outcome measurement is commonplace is economic evaluation, where multiple health technology assessment (HTA) agencies have published guidelines that recommend units of benefit be expressed in quality-adjusted life years (QALYs), within a cost-utility framework [1–3]. Although there is nothing inherently health-related or generic about the QALY,[1] its widely practised operationalisation affirms these two characteristics through the use of standardised, generic health-related quality of life (HRQoL) instruments [2, 11]. Such instruments—known as multi-attribute utility instruments or preference-based HRQoL instruments—comprise a fixed set of items and response options (the descriptive classification system), ensuring that an individual's set of responses correspond to one of a finite number of health states. For example, an instrument with six questions, each with three levels of response, defines 729 (3^6) unique health state descriptions. Each instrument has its own scoring procedure (the valuation system)—sometimes multiple, country-specific scoring procedures [12, 13]—which is a set of preference weights (also known as 'index scores') that represent the relative value that society places on living in each of the instrument-defined states [14, 15].

Several preference-based HRQoL instruments have been developed since the 1990s, including the Health Utilities Index (HUI) [16], instruments developed by the EuroQol Group (e.g., EQ-5D-3L [17] and EQ-5D-5L [18]), and the Assessment of Quality of Life (AQoL) instruments [19–22]. Concurrent to such developments, there have been challenges as to whether the evaluative space of outcome measurement in economic evaluation should be restricted to health. Early discussion (mid-1990s onward) focused on circumstances where individuals may be disadvantaged when adopting a decision-making framework based on health maximisation, highlighting sources of benefit such as autonomy and process of care that would be overlooked (to the extent that these benefits would not improve health outcome) [23, 24]. In recent years, several research programmes have developed (or will develop) preference-based instruments for use in economic evaluation that quantify aspects of non-health outcomes. Examples include the Adult Social Care Outcomes Toolkit (ASCOT), which is a set of instruments that focus on aspects of quality of life that are affected by social care services [25, 26]; the Investigating Choice Experiments for the Preferences of Older People (ICEPOP) project, from which the ICEpop CAPability (ICECAP) measures evolved (discussed in the following) [27]; and the 'Extending the QALY' (E-QALY) project, which seeks to develop a broad measure of quality of life that captures the impacts of treatments and interventions across health and social care sectors [28, 29].

The ICECAP instruments are measures of wellbeing that are conceptually linked to Amartya Sen's capability approach, which advocates for an evaluative focus on the extent to which people have the capability to function, regardless of whether they choose to do so [5, 30]. To distinguish between capabilities (what people are able to do) and achieved functionings (what they actually do), the ICECAP measures adopt phrasing such as 'I can/cannot' and 'I am able/unable to', with evidence suggesting that such wording can elicit differences in individuals' self-reported capabilities and functionings [31]. The first ICECAP instrument was developed for older adults (ICECAP-O) [32], with subsequent versions designed for adults (ICECAP-A) [33],

[1] (1) A literal, non-technical interpretation of the QALY is that of a metric that weights periods of survival by a quality-adjustment factor. Within an extra-welfarist approach, which underpins much of current practice in economic evaluation, the quality-adjustment factor is often health related [4, 5]. Indeed, a popular school of thought is that extra-welfarism replaces 'utility' (the evaluative space under welfarist economics) with 'health' [6]. Brouwer and colleagues [4, 7] reject the idea that extra-welfarism imposes a restrictive conceptualisation of outcomes, instead proposing that it provides a more general framework for analysis, allowing elements other than utility—such as indicators of health and wellbeing (i.e., the 'extras' in extra-welfarism)—to be incorporated.
(2) The statement, "there is nothing inherently … generic about the QALY" is recognition of the availability of condition-specific outcome measures that can be used to construct (condition-specific) QALYs [8–10].

individuals at the end of life (the ICECAP Supportive Care Measure [ICECAP-SCM]) [34] and those 'close to the dying' in the context of end-of-life care (ICECAP Close Person Measure [ICECAP-CPM] [35]). Further research aims to extend this line of work to the measurement of capability in children and explore a life-course approach to capability assessment [36].

Focusing specifically on the ICECAP-A, the instrument comprises five dimensions: Stability (an ability to feel settled and secure), Attachment (an ability to have love, friendship, and support), Autonomy (an ability to be independent), Achievement (an ability to achieve and progress in life), and Enjoyment (an ability to experience enjoyment and pleasure) [33]. Each dimension comprises one question with four levels of response, defining 1024 unique capability wellbeing states. While there are similarities in the appearance and components of ICECAP instruments and preference-based HRQoL instruments—for example, a descriptive classification system and a valuation system—a key difference is in the range and interpretation of the societal valuations. For the health-related measures, values are interpreted on a scale from zero (dead) to one (full health), where it is possible to have negative values that represent health states worse than dead. Alternatively, the anchors of the scoring ranges for ICECAP instruments are 'no capability' (zero) and 'full capability' (one) [37, 38].

Debate will persist about the merits of different evaluative spaces, as well as other methodological and operational issues related to the development and use of different preference-based instruments [39–41]. However, once an instrument is available to use in economic evaluation (and recognised in the guidelines produced by HTA agencies and national decision-making bodies [42, 43]), it is important to assess its performance across different clinical contexts. A recent (2020) systematic review identified 16 studies as having explored the measurement properties of the ICECAP-A in a specific health context or a general population sample [44]. At a population level,[2] Afentou and Kinghorn [44] reported supportive evidence for the ICECAP-A regarding content validity [45, 46], construct validity [47–49], and reliability [50]. Evidence from patient populations included women with irritative lower urinary tract symptoms (construct validity and responsiveness) [51], people with knee pain attributable to osteoarthritis (construct validity and responsiveness) [52, 53], depression (construct validity) [54, 55], opiate dependence (construct validity) [56], individuals living with spinal cord injury (SCI) (construct validity) [57] and family members of people with meningitis (content validity) [31]. Goranitis et al. also explored 'sensitivity to change' in the opiate dependency study [56], but this was not classified as assessing responsiveness in the systematic review [44].

Developments in the treatment and care of individuals with SCI have resulted in longer life expectancy post-injury [58], and complications after (or secondary to) SCI are common, with pain, spasticity, sexual dysfunction and urinary tract infections being among the most prevalent [59, 60]. Enabling people to attain an acceptable quality of life is considered by many to be the primary goal of health care providers following SCI [61]. To pursue such an objective in a rigorous and quantifiable manner, the evaluation of interventions and/or rehabilitation services requires generic instruments that are capable of adequately measuring outcomes such as quality of life and wellbeing in this population. Although the capability approach has been the focus of some attention within disability research [62–64], including implications for policy and HTA [65], there is currently only one published study in the area of SCI that includes an ICECAP instrument [57]. Using a regression-based path analysis, Engel et al. [57] showed that the presence of health conditions secondary to SCI has different impacts on HRQoL, capability wellbeing (measured by the ICECAP-A) and subjective wellbeing. Despite not being described as an assessment of construct validity (or any other measurement property) by Engel et al., this study met the inclusion criteria for the above-mentioned review by Afentou and Kinghorn [44]. To date, validation research for preference-based measures in SCI has focused primarily on health-related instruments [66–69]. This paper provides the first comprehensive report of the empirical validity of a preference-based capability wellbeing instrument (ICECAP-A) in the context of SCI. In addition to a standalone assessment of psychometric criteria, an explicit objective is to provide further insights regarding the complementarity of the ICECAP-A with health-related preference-based instruments.

2 Methods

2.1 Data Source and Outcome Measures

This study comprises a secondary analysis of data from an online cross-sectional survey. Comprehensive details of the online survey (study design, recruitment method, description of participants, survey administration and survey composition, etc.) are reported elsewhere [69]. Briefly, participants were a subsample of individuals who had previously

[2] Some of the studies described as 'general population' in the review by Afentou and Kinghorn [44] used data from the Multi Instrument Comparison (MIC) study (http://www.aqol.com.au/index.php/aqol-current). Quota sampling was used in the MIC study to obtain a target number of respondents in each of seven disease areas (arthritis, asthma, cancer, depression, diabetes, hearing loss and heart disease) and a demographically representative sample of 'healthy' participants.

participated in a Canada-wide SCI project—the SCI Community Survey [70]—and provided consent to be contacted for further research. The survey contained five sections: demographics, SCI classifications and characteristics, secondary health conditions, quality of life and wellbeing, and SCI-specific functioning in activities of daily living. A forced choice format was used, except for some demographic questions, where participants could select 'don't know' or 'prefer not to answer'. Although the survey could be completed across multiple sessions, all questions in Section 4 (quality of life and wellbeing) were required to be answered in a single session. The order in which standardised instruments appeared in Section 4 was randomised to prevent order-effect bias [71].

Details of the five generic, preference-based instruments used in this study (ICECAP-A, AQoL 8-dimension questionnaire [AQoL-8D], EQ-5D-5L, HUI-3 and SF-6D) are provided in Table 1. Other standardised measures used in the analyses capture aspects of general health, life satisfaction, secondary health conditions, and SCI-specific functioning in activities of daily living. Item 1 of the SF-36v2 was used as a measure of general health [76]. Life satisfaction was assessed by a single, 11-point scale that asked, "Using a scale of 0–10, where 0 means "very dissatisfied" and 10 means "very satisfied", how do you feel about your life as a whole right now?" [77]. The presence of health conditions secondary to SCI was captured using a modified version of the Spinal Cord Injury Secondary Conditions Scale (SCI-SCS) [78]. The modified version captures information about 16 secondary health conditions (muscle spasms [spasticity]; depression/mood problems; pressure sore(s); bladder dysfunction; trouble sleeping; joint and muscle pain; neuropathic pain; sexual dysfunction; autonomic dysreflexia; bowel dysfunction; fatigue; fractures; urinary tract infection; respiratory problems; osteoarthritis/degenerative arthritis; and cerebrovascular disease, stroke, or transient ischaemic attack) using a four-point ordinal scale (0: not experienced in the last three months/insignificant problem; 1: mild or infrequent problem; 2: moderate or occasional problem; 3: significant or chronic problem) [57]. Using the same scoring procedure as the SCI-SCS, which has been shown to have 'moderate to high' evidence of validity and reliability [78], total scores for the modified scale range from zero (no problems/insignificant problems only) to 48 (significant/chronic problems in all 16 secondary health conditions). The self-report version of the Spinal Cord Independence Measure (SCIM-III), SCIM-SR, was used to measure SCI-specific functioning in activities of daily living [79, 80]. Items for 19 daily tasks are weighted according to clinical relevance, with the scoring procedure producing a total score (0 to 100) and subscale scores across three domains (self-care [0–20], respiration and sphincter management [0–40], and mobility [0–40]), with higher scores reflecting higher levels of performance or independence.

2.2 Statistical Analysis

Prior to the consideration of empirical validity, a descriptive assessment of ICECAP-A index scores (descriptive statistics and frequency distribution) and dimension-level response patterns was performed. Adopting the classification system described by Brazier et al. [15], empirical validity was examined through a series of stated preference and hypothetical preference tests. The distinction lies in the variable(s) used to define the test. Stated preference tests comprise the use of preference-based measures administered alongside each other to investigate the extent of convergence in the respective index scores. Hypothetical preference tests require the researcher(s) to assume and define a difference in preferences and examine whether the expected difference is reflected in the index scores of the preference-based instrument being evaluated. In this study, the empirical tests used for the stated preferences and hypothetical preferences all fall under the psychometric property of construct validity, which, broadly, concerns the extent to which a measure agrees with another measure "in a manner that is consistent with theoretically derived hypotheses" [81]. More specifically, discriminant and convergent validity approaches were used.

2.2.1 Discriminant (Known-Groups) Validity

Discriminant validity is the ability of an instrument to differentiate between groups that are hypothesised to differ [81]. Given the paucity of research regarding capabilities and capability wellbeing in the context of SCI, the formation of hypothesised constructs was informed primarily by evidence pertaining to subjective wellbeing, life satisfaction, quality of life and/or HRQoL. Table 2 provides details of the 16 hypothetical preference tests that were assessed using the discriminant validity approach; these hypothesised constructs were classified as 'strong' (consistent supportive evidence) or 'exploratory' (mixed or little evidence) in recognition of the varying degree of evidence. Constructs with two groups were assessed using a one-sided, independent samples *t* test (equal variances assumed) or Welch *t* test (equal variances not assumed). Linear trend analysis was used in the assessment of the constructs comprising three or more groups. For constructs that did not meet the parametric assumption of homogeneity of variances, heteroscedasticity-consistent standard errors were applied using the 'sandwich' package in R [114]. The 'age at injury' construct was examined using multiple regression, controlling for the time (number of days) between the date of injury and the date of survey completion.

Table 1 Key properties of the preference-based instruments included in the survey

Instrument	Dimensions within the descriptive system	Number of items used to derive an index score	Number of unique states	Valuation study participants[a]	Elicitation method	Range of values[b]
ICECAP-A [37]	5: Stability, Attachment, Autonomy, Achievement, Enjoyment	5 questions, each with 4 levels of response	1024	A random sample of the UK adult population ($n = 413$)	BWS	0.149 to 1.000
AQoL-8D [22]	8: Independent Living, Mental Health, Coping, Relationships, Pain, Senses, Happiness, Self-worth	35 questions, with between 4 and 6 levels of response	2.37×10^{23}	Mental health patients ($n = 323$) and a random sample of the general public ($n = 347$) in Australia	VAS and TTO	0.090 to 1.000[c]
EQ-5D-5L [72]	5: Mobility, Self-care, Usual Activities, Pain/Discomfort, Anxiety/Depression	5 questions, each with 5 levels of response	3125	A random sample of the adult population in four Canadian cities ($n = 1073$)	Composite TTO	− 0.148 to 0.974
HUI-3 [73]	8: Vision, Hearing, Speech, Ambulation, Dexterity, Emotion, Cognition, Pain	12 questions, with between 4 and 6 levels of response[d]	972,000	A random sample of the adult population in Hamilton, Canada ($n = 256$)[e]	SG and VAS	− 0.359 to 1.000
SF-6D (SF-36v2) [74, 75]	6: Physical Functioning, Role Limitations, Social Functioning, Pain, Mental Health, Vitality	11 (of 36) questions, with between 3 and 6 levels of response	18,000	A random sample of the UK adult population ($n = 611$)	SG	0.301 to 1.000

BWS best-worst scaling, *SG* standard gamble, *TTO* time trade-off, *VAS* visual analogue scale

[a]Participants included in the derivation of the respective scoring algorithm

[b]Interpretation of the anchors differs between the health-related instruments and the ICECAP-A (see Sect. 1)

[c]In addition to the overall index score, the AQoL-8D also provides scores on a 0–1 scale for each dimension, and for ‘super-dimensions’ of physical health (Independent Living, Pain and Senses) and psychosocial health (Mental Health, Coping, Relationships, Happiness, and Self-worth)

[d]The HUI self-administered questionnaire consists of 15 questions

[e]A total of 504 interviews were conducted for the HUI-3 valuation study: 256 in a modelling survey and 248 in a validation survey. Authors explicitly state that the regional sample was representative of Canada

Table 2 Details of the 16 hypothetical constructs used to explore the discriminant validity of the ICECAP-A in the context of spinal cord injury

Construct[a]	Expected relationship with ICECAP-A scores and evidence upon which the construct was hypothesised[b]
Strong hypotheses	
General health (5; SF-36v2 item 1)	*Linear trend*—better self-perceived general health, as measured by item 1 of the SF-36/SF-12 instruments, is predictive of increased life satisfaction [82] and HRQoL [66]. Health status has also been shown to be positively associated with ICECAP-A dimensions in the general population (EQ-5D-3L) [47] and family members of people with meningitis (EQ-5D-5L) [31]
Mental health (5; SF-6D Mental Health)	*Linear trend*—reports of mental health, positive mood, control, and self-worth are positively associated with subjective quality of life [83–85]. ICECAP-A dimensions are positively associated with the Anxiety/Depression dimension of the EQ-5D-3L in the general population [47], and ICECAP-A scores satisfy hypothesised constructs in adults with depression [55]
Social functioning (5; SF-6D Social Functioning)	*Linear trend*—social integration and social support [86], activity participation [87], and community integration [88] are positively associated with subjective quality of life. Social health is positively associated with life satisfaction [85, 89]
Role/activity limitations (4; SF-6D Role Limitation)	*Linear trend*—participation in activities such as work, school and social activities is associated with higher subjective quality of life [86, 87] and life satisfaction [90]
Independence[c] (3; SCIM-SR mobility)	*Linear trend*—there are positive associations between: SCIM-III scores and life satisfaction [91]; functional independence and life satisfaction [60]; functional independence and quality of life [60, 66, 87, 92]; and, among (non-SCI) older adults, independence in activities of daily living and ICECAP-O scores [93]
Independence[c] (3; SCIM-SR self-care)	*Linear trend*—as above (for the SCIM-SR mobility construct). In (non-SCI) older adults, self-care (EQ-5D-3L) is positively associated with several ICECAP-O dimensions [94, 95]
Life satisfaction (11)	*Linear trend*—life satisfaction is positively associated with HRQoL [66] and, in (non-SCI) older adults, with ICECAP-O scores [93] and dimensions [94]
Time since injury (2)	*Higher ICECAP-A scores for those 10 years or more post-injury* – time since injury is positively associated with life satisfaction [90, 96] and quality of life [84, 97, 98]
Secondary health conditions[c] (3; modified SCI-SCS)	*Linear trend*—secondary health conditions are associated with lower HRQoL [60, 66, 92, 98–101], life satisfaction [99, 102] and subjective quality of life [88, 97]
Paid employment (2)	*Higher ICECAP-A scores for those in paid employment*—being employed is associated with higher life satisfaction [91, 96, 102], HRQoL [98, 103] and subjective quality of life [86]. Paid employment is also positively associated with ICECAP-A dimensions in the general population [47]
Marital status (2)	*Higher ICECAP-A scores if married or common law*—presence of a spouse is associated with higher life satisfaction [60, 88–90, 102, 104, 105] and quality of life [87, 98, 106]; presence of a spouse is also associated with higher reported capability on several ICECAP-A dimensions in the general population [47]
Exploratory hypotheses	
Education (3)	*Linear trend*—inconsistent evidence as to whether education is positively associated with life satisfaction [89–91, 96] or subjective quality of life [84, 107, 108]. Higher education is positively associated with ICECAP-A dimensions in the general population [47] and family members of people with meningitis [31]
Happiness (5; AQoL-8D Happiness)	*Linear trend*—positive affect is associated with increased subjective quality of life [83], and happiness is positively associated with ICECAP-A dimensions in the general population [47]
Household income (6)	*Linear trend*—limited evidence suggests that lower household income and financial hardship is associated with lower subjective quality of life [107, 108]. Income is positively associated with ICECAP-A dimensions in the general population [47]
Neurological level and completeness of injury (3)	*Linear trend*—although evidence is mixed, this hypothesis combines assertions that individuals living with paraplegia (regardless of the completeness of injury) will have higher scores than individuals living with tetraplegia [66, 84, 89, 90, 92, 96, 98, 100, 107, 109] and, for individuals living with tetraplegia, those with incomplete injuries will have higher scores than those with complete injuries [66, 89, 98]

Table 2 (continued)

Construct[a]	Expected relationship with ICECAP-A scores and evidence upon which the construct was hypothesised[b]
Age at injury	*Injury occurring at a younger age is predictive of higher ICECAP-A scores* (after controlling for the number of days between participants' date of injury and date of survey completion)—evidence suggests that SCI onset at an older age may be associated with lower quality of life [66, 92, 98, 110–113]

HRQoL health-related quality of life, *SCI* spinal cord injury

[a]Information in parentheses provides details of the number of categories and, where appropriate, the item/dimension of the standardised instrument used to define the respective construct

[b]No explicit distinctions are made between 'positive' and 'negative' linear trends to prevent confusion due to the different scoring procedures of the construct-defining variables. The directions of the expected relationships can be inferred from the evidence described. Unless stated otherwise, the evidence described is specific to the context of SCI

[c]The primary analysis of the construct involved splitting the distribution of scores into groups using tertiles, with category labels (for groups 1–3) reflecting the scoring procedure of the instrument. To test how robust the findings were to the decision to use tertiles, the analysis was repeated using quartiles and quintiles

2.2.2 Convergent Validity

Convergent validity explores the extent to which a measure correlates with other measures of the same (or similar) concept [115]. Analysis began with a descriptive, omnibus assessment of correlation, generating Spearman's rank correlations between responses to the five ICECAP-A dimensions and (1) item-level data for the AQoL-8D (35 items) and (2) dimension-level data for the EQ-5D-5L (five dimensions), HUI-3 (eight dimensions) and SF-6D (six dimensions). Different approaches were required because the AQoL-8D does not define a number of levels per dimension, unlike the EQ-5D-5L, HUI-3 and SF-6D.

Hypotheses for the convergent validity analysis examined the relationships between preference-based index scores (stated preference tests), the relative strength of association between ICECAP-A Autonomy and the SCIM-SR, and the respective approaches to the inclusion of pain in the HRQoL instruments. Regarding the latter, items focusing on interference or limitations due to pain were expected to correlate more strongly with ICECAP-A index scores compared with items referring to the severity of pain based on the idea that interference and limitation is conceptually closer to the notion of capabilities. The wording of the items and respective response options of the pain-related items included in this hypothesis are reported in the Electronic Supplementary Material (ESM1); specifically, AQoL-8D item 24 and SF-36v2 item 22 were expected to have stronger correlations with ICECAP-A index scores compared with AQoL-8D item 22, EQ-5D-5L Pain/Discomfort and SF-36v2 item 21. For the stated preference tests, it was hypothesised that the AQoL-8D would have a stronger correlation with the ICECAP-A compared with the EQ-5D-5L, HUI-3 or SF-6D because of the greater coverage of psychosocial dimensions in the AQoL-8D [46, 116, 117]. Similarly, the AQoL-8D psychosocial super-dimension was expected to have a stronger correlation with the ICECAP-A compared with the AQoL-8D physical super-dimension. Finally, ICECAP-A Autonomy was hypothesised to have a stronger correlation with the SCIM-SR self-care and mobility subscale scores when compared with the other four ICECAP-A dimensions.

The direction of recommended item-/dimension-level coding for the five preference-based instruments is not consistent. For example, higher levels of response on an ICECAP-A dimension (e.g., level 3 vs. level 2) indicate a better outcome, whereas level 1 is coded as the 'top level' for the four health-related instruments. To provide an intuitive interpretation of the direction of associations explored in this paper, item-/dimension-level coding was reversed for the health-related instruments when performing correlation analyses. This ensures that a positive correlation coefficient indicates a relationship between two variables (e.g., ICECAP-A Enjoyment and HUI-3 Emotion) where 'better' or 'worse' outcomes tend to occur together.

Where applicable, findings were interpreted using a statistical significance level of 0.05. Analyses were conducted using R (version 3.6.0) [118].

3 Results

Three hundred sixty-four individuals living with SCI completed the survey and were included in the analysis. Participant characteristics and ICECAP-A descriptive statistics are provided in Table 3. The majority of the sample was male (63%), with 90% self-identifying as Caucasian and 70% being at least 10 years post-injury. Mean age was 50.4 years, and there was representation across categories of household income, education, neurological level and injury classification. The mean ICECAP-A index score

Table 3 Demographic characteristics, clinical characteristics and descriptive statistics for the SCIM-SR and ICECAP-A[a]

Characteristic/outcome	*n* = 364
Mean age	50.40 ± 13.2
Sex—female	135 (37.1)
Ethnicity	
Caucasian	328 (90.1)
Other	33 (9.1)
Relationship status	
Married or common law	186 (51.1)
Single, never married	105 (28.8)
Divorced, separated or widowed	69 (19.0)
Living arrangement	
Own home	267 (73.4)
Rental or other	94 (25.8)
Education	
High school diploma/GED and below	102 (28.0)
Diploma/Certificate	140 (38.5)
Bachelor's degree or higher	119 (32.7)
Household income	
< $20,000	56 (15.4)
$20,000–39,999	49 (13.5)
$40,000–59,999	57 (15.7)
$60,000–79,999	47 (12.9)
$80,000–99,999	47 (12.9)
≥ $100,000	53 (14.6)
Time since injury	
1–10 years	109 (29.9)
≥ 10 years	255 (70.1)
Neurological level	
Tetraplegia	175 (48.1)
Paraplegia	189 (51.9)
ASIA Impairment Scale (neurological classification of injury)[b]	
A	115 (31.6)
B	31 (8.5)
C	119 (32.7)
D	69 (19.0)
E	8 (2.2)
Uncertain (A or C) or unclassifiable	22 (6.0)
Cause of injury	
Traumatic	272 (74.7)
Non-traumatic	92 (25.3)
SCIM-SR[c]	
Self-care subscale score	14.23 ± 6.4
Mobility subscale score	16.67 ± 10.1
Number of moderate or significant secondary conditions[d]	
Zero	9 (2.5)
1–3	90 (24.7)
4–6	97 (26.6)
7–9	111 (30.5)
≥ 10	57 (15.7)
ICECAP-A	
Mean index score	0.761 ± 0.18
Median (interquartile range) index score	0.828 (0.26)

Table 3 (continued)

Characteristic/outcome	$n = 364$
Highest reported index score	1.000
Lowest reported index score	0.149
Full capability (state 44444)	12 (3.3)
Complete absence of capability (state 11111)	0 (0.0)

Values are the number (percentage) or the mean ± standard deviation unless stated otherwise

ASIA American Spinal Injury Association, *GED* General Educational Development

[a]Numbers for some characteristics do not always sum to 364 because of 'prefer not to answer' responses

[b]ASIA Impairment Scale classifications were determined using an algorithm developed by the SCI Community Survey research team [70]

[c]Due to a programming error for one item of the SCIM-SR, SCIM-SR total scores and SCIM-SR respiration and sphincter management subscale scores are not reported

[d]Based on responses to the modified SCI-SCS

was 0.761; 12 (3%) individuals reported 'full capability' (i.e., state 44444, no suboptimal level of capability on any dimension), and no individual reported the absence of capability on all five dimensions (state 11111). As illustrated in Fig. 1, the distribution of ICECAP-A index scores was negatively skewed. The ICECAP-A dimension-level response pattern is reported in Table 4. The absence of capability (level 1) was reported by less than 5% of individuals in each of the five dimensions. The highest two levels of response (levels 3 and 4) accounted for between 67% (Achievement) and 82% (Attachment) of responses; the highest proportions of 'full capability' (level 4) responses were for Attachment (39%) and Enjoyment (24%).

Table 4 Distribution of responses across the levels of ICECAP-A dimensions[a]

Attribute	Level 1 (no capability)	Level 2	Level 3	Level 4 (full capability)
Stability	17 (4.7)	94 (25.8)	206 (56.6)	47 (12.9)
Attachment	4 (1.1)	62 (17.0)	156 (42.9)	142 (39.0)
Autonomy	18 (4.6)	83 (22.8)	204 (56.0)	59 (16.2)
Achievement	14 (3.8)	107 (29.4)	189 (51.9)	54 (14.8)
Enjoyment	5 (1.4)	88 (24.2)	182 (50.0)	89 (24.5)

Values are numbers (percentages)

[a]This coding format is in accordance with the guidance documentation for using the ICECAP-A: https://www.birmingham.ac.uk/Documents/college-mds/haps/projects/icecap/Icecap-A/ICECAP-A-Scoring.docx (Accessed 6 Aug 2020)

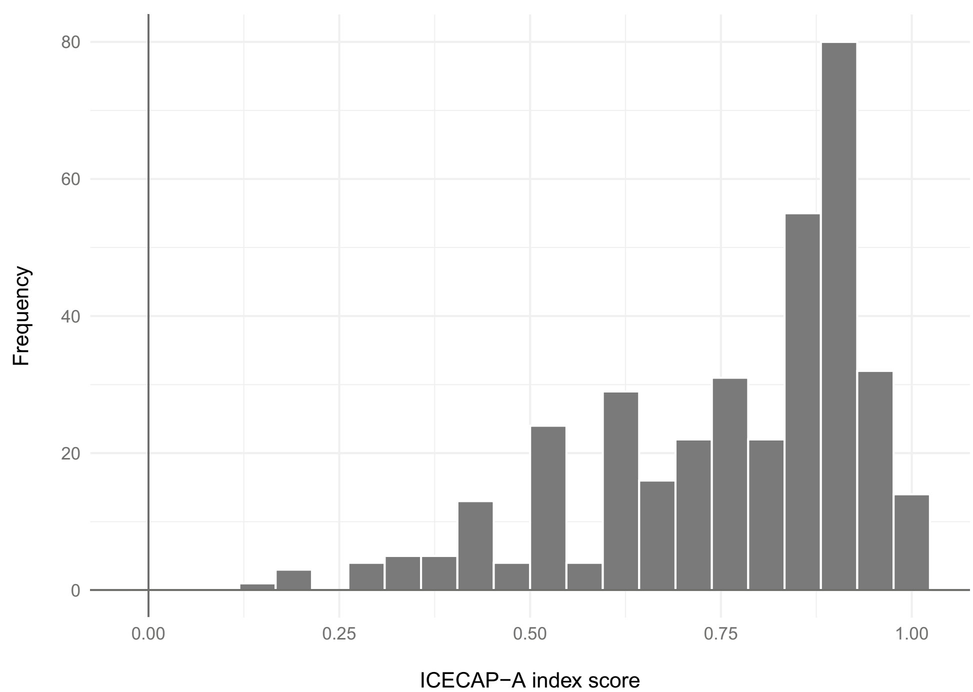

Fig. 1 Frequency distribution of ICECAP-A index scores ($n = 364$)

3.1 Discriminant Validity

Table 5 reports category-level mean ICECAP-A scores and the *p* value for the respective statistical test for the discriminant validity hypothetical preference tests (except for the 'age at injury' construct). All 11 strong hypotheses were confirmed ($p < 0.05$), as were the exploratory constructs for happiness, household income, and neurological level and completeness. No statistically significant linear trend was observed for the education construct ($p = 0.067$). After controlling for the time between injury onset and survey completion, 'age at injury' was not a statistically significant predictor of ICECAP-A index scores ($\beta = 7.60 \times 10^{-5}$, $p = 0.920$).

3.2 Convergent Validity

Coefficients from the omnibus assessment of correlation are presented as a correlation matrix heatmap in Fig. 2, with the corresponding values reported in the Electronic Supplementary Material (ESM2). Observations from this descriptive analysis include (1) ICECAP-A Autonomy having stronger correlations for items and dimensions capturing physical aspects of quality of life (and zero correlations > 0.24 with any of the 11 mental health items/dimensions) compared with the other four ICECAP-A dimensions, (2) the absence of any correlations stronger than 0.40 for items/dimensions related to pain or senses (hearing, vision or speech/communication), and (3) the relatively high proportion of low correlations (16 of 40 coefficients with an absolute value ≤ 0.10) for the HUI-3 (i.e., Fig. 2 illustrates a high proportion of pale shaded boxes for the HUI-3 relative to the other health-related instruments). At the instrument level, comparison of the 10 strongest and 10 weakest correlations (Table 6) indicated that the ICECAP-A has a greater conceptual overlap with the AQoL-8D when compared with the EQ-5D-5L, HUI-3 and SF-6D. The 'top 10' pairs are dominated by items and dimensions on the comparator (HRQoL) instruments that could be described as non-health aspects of quality of life, such as happiness (×4) and control over one's life (including self-care) (×4). Two mental health items/dimensions complete the list of the 10 strongest correlations (HUI-3 Emotion and AQoL-8D sadness). This contrasts with the 10 weakest correlations, which comprise HUI-3 Vision (×4), HUI-3 Dexterity (×2) and the mobility-related dimensions of the EQ-5D-5L (×3) and HUI-3 (×1).

In line with expectations, correlation coefficients between ICECAP-A index scores and pain-related item-/dimension-level data were stronger when the pain item was framed in the context of interference or limitations (AQoL-8D item 24 and SF-36v2 item 22) rather than severity (AQoL-8D item 22, EQ-5D-5L Pain/Discomfort, and SF-36v2 item 21) (Table 7). The hypotheses for preference-based index scores were also confirmed, with the ICECAP-A more strongly correlated with (1) AQoL-8D than with EQ-5D-5L, HUI-3 and SF-6D, and (2) the AQoL-8D psychosocial super-dimension than with the AQoL-8D physical super-dimension (Table 7). Hypotheses were also confirmed for the associations between the SCIM-SR and ICECAP-A Autonomy, i.e., Spearman's rank correlations between the SCIM-SR self-care and mobility subscale scores and ICECAP-A dimensions were strongest for ICECAP-A Autonomy [see Electronic Supplementary Material (ESM2)].

4 Discussion

This paper reports the first comprehensive examination of the empirical validity of the ICECAP-A in the context of SCI. Using construct validity techniques and correlation analysis, the findings showed that the ICECAP-A performed in line with expectations regarding relationships with health, injury, functioning and demographic indicators as well as previous research demonstrating greater overlap between the ICECAP-A and non-health items/dimensions present in preference-based HRQoL instruments [46, 48].

The observation that the ICECAP-A has the strongest associations with the AQoL-8D (when compared with other preference-based HRQoL instruments) confirms findings from studies that have used data from the Multi Instrument Comparison study (http://www.aqol.com.au/index.php/aqol-current) [46, 48, 49]. The weakest dimension-to-dimension correlations most often included EQ-5D-5L and HUI-3 dimensions (e.g., see Table 6), supporting the idea that the ICECAP-A is a complement to and not a substitute for these HRQoL measures [46, 48, 53]. Whether the ICECAP-A/AQoL-8D overlap would be observed with other AQoL instruments (AQoL-4D, AQoL-6D or AQoL-7D) is unknown. In their systematic review, Afentou and Kinghorn [44] categorised some studies that looked at the overlap between the ICECAP-A and HRQoL instruments as assessments of content validity, concluding that, "(such) results showed a clear distinction between capabilities and functioning because the ICECAP-A provided different information than that elicited from other measures." This interpretation warrants closer inspection because it assumes that the ICECAP-A and the HRQoL instruments are measuring capabilities and functionings, respectively (as opposed to measuring different functionings, different capabilities or different combinations of the two). In the context of preference-based instruments, only one study has been designed to examine the degree to which individuals perceive functionings and capabilities to differ. In the novel study by Al-Janabi [31], approximately one-third of participants from a UK convenience sample reported a difference in at least one dimension when completing a modified version of the ICECAP-A that

Table 5 Results of the discriminant validity analyses for the 11 strong hypotheses and four (of five) exploratory hypotheses[a]

Construct	Categories	*n*	ICECAP-A score (mean ± SD)	*p* value
Strong hypotheses				
General health	Excellent	20	0.909 ± 0.10	< 0.001[b]
	Very Good	106	0.830 ± 0.13	
	Good	148	0.769 ± 0.16	
	Fair	66	0.668 ± 0.19	
	Poor	24	0.540 ± 0.21	
Mental health	Level 1 (highest)	82	0.872 ± 0.11	< 0.001[b]
	Level 2	129	0.811 ± 0.13	
	Level 3	102	0.713 ± 0.17	
	Level 4	41	0.586 ± 0.18	
	Level 5 (lowest)	10	0.420 ± 0.21	
Social functioning	Level 1 (highest)	92	0.871 ± 0.12	< 0.001[b]
	Level 2	79	0.813 ± 0.13	
	Level 3	120	0.723 ± 0.17	
	Level 4	56	0.689 ± 0.18	
	Level 5 (lowest)	17	0.430 ± 0.20	
Role/activity limitation	Level 1 (highest)	38	0.900 ± 0.12	< 0.001[b]
	Level 2	88	0.837 ± 0.12	
	Level 3	13	0.784 ± 0.21	
	Level 4 (lowest)	225	0.706 ± 0.19	
Independence (self-care)[c]	Group 1 (lowest)	122	0.695 ± 0.19	< 0.001[b]
	Group 2	121	0.763 ± 0.17	
	Group 3 (highest)	121	0.825 ± 0.16	
Independence (mobility)[c]	Group 1 (lowest)	122	0.719 ± 0.18	< 0.001
	Group 2	121	0.769 ± 0.18	
	Group 3 (highest)	121	0.796 ± 0.18	
Life satisfaction	0 (lowest)	5	0.299 ± 0.18	< 0.001[b]
	1	6	0.459 ± 0.21	
	2	12	0.495 ± 0.11	
	3	30	0.528 ± 0.15	
	4	22	0.609 ± 0.16	
	5	40	0.688 ± 0.13	
	6	45	0.759 ± 0.14	
	7	78	0.828 ± 0.09	
	8	73	0.871 ± 0.09	
	9	38	0.889 ± 0.09	
	10 (highest)	15	0.933 ± 0.11	
Secondary health conditions[c]	Group 1 (highest)	122	0.851 ± 0.14	< 0.001[b]
	Group 2	121	0.777 ± 0.14	
	Group 3 (lowest)	121	0.654 ± 0.20	
Time since injury	1–10 years	109	0.732 ± 0.19	0.024
	≥10 years	255	0.773 ± 0.18	
Paid employment	No	238	0.732 ± 0.18	< 0.001[b]
	Yes	126	0.816 ± 0.17	
Marital status	Single, never married	105	0.744 ± 0.18	0.016
	Married or common law	186	0.790 ± 0.17	

Table 5 (continued)

Construct	Categories	n	ICECAP-A score (mean ± SD)	p value
Exploratory hypotheses				
Education	High school diploma/GED and below	102	0.725 ± 0.21	0.067[b]
	Diploma/Certificate	140	0.777 ± 0.16	
	Bachelor's degree or higher	119	0.774 ± 0.18	
Happiness[d]	All the time	23	0.905 ± 0.09	< 0.001[b]
	Mostly	210	0.833 ± 0.12	
	Sometimes	111	0.649 ± 0.18	
	Almost never or never	20	0.463 ± 0.17	
Household income	Under $20,000	56	0.661 ± 0.21	< 0.001
	$20,000–39,999	49	0.725 ± 0.16	
	$40,000–59,999	57	0.764 ± 0.17	
	$60,000–79,999	47	0.807 ± 0.17	
	$80,000–99,999	47	0.796 ± 0.18	
	≥$100,000	53	0.810 ± 0.18	
Neurological level and completeness	Paraplegia (complete or incomplete)	182	0.780 ± 0.18	0.021
	Tetraplegia incomplete	113	0.758 ± 0.18	
	Tetraplegia complete	45	0.709 ± 0.17	

GED General Educational Development

[a]The category labels 'highest' and 'lowest' are used to illustrate the direction of the expected relationship with ICECAP-A index scores for the constructs with non-intuitive category labels, reflecting the scoring procedures of the construct-defining variables

[b]Assumption of homogeneity of variances was not met

[c]A statistically significant linear trend was also observed when the construct was analysed using quartiles and quintiles ($p < 0.001$)

[d]The response categories for level 4 ('Almost never'; $n = 18$) and level 5 ('Never'; $n = 2$) were combined

asked about capabilities ('I can'/'I am able to') and functionings ('I do'/'I am') [31]. While it is evident that the ICECAP-A provides information beyond preference-based HRQoL instruments, the extent of the additional information is dependent on the comparator [48], and there remains uncertainty about what underlying constructs (capabilities and/or functionings) ICECAP and preference-based HRQoL instruments are measuring.

The emergence of consistent, supportive evidence for the empirical validity of the ICECAP-A and the availability/development of other preference-based instruments (e.g., ASCOT, E-QALY and other ICECAP measures) emphasises the need for further empirical and conceptual research regarding the evaluative space for economic evaluation. This is not the exclusive territory of academic debate—the needs and values of decision makers and the public are important if economic evaluation is to serve its purpose in supporting efficient and equitable decision making. Such research will need to extend beyond (important) 'within-space' issues, such as the development of decision rules and monetary thresholds for ICECAP instruments [41, 120], and the estimation of value sets in non-UK settings.

The strengths of this study lie in the clinical context (i.e., the paucity of psychometric assessment of capability instruments in the context of SCI), the use of multiple comparator HRQoL instruments, and the incorporation of SCI and non-SCI evidence to support the definition of explored constructs. As with all studies, there are limitations. First, as a secondary analysis of cross-sectional data, it was not possible to assess psychometric criteria that are best explored with repeated measurements, such as responsiveness to change or test–retest reliability. Given that it is the change in outcome that is a key factor in intervention-based research, this remains an important area for further work. Second, it is important to acknowledge the inevitable subjectivity in this field of research, which could manifest in the choice of methods (e.g., the definition of constructs and classification of 'strong' and 'exploratory' hypotheses) or the interpretation of results (e.g., highlighted observations from the heatmap). The comprehensive and transparent reporting of procedures and results mitigates the implications of this limitation for the study findings. Finally, caution is required when generalising study findings to other SCI settings. Participants were a subsample of individuals who had taken

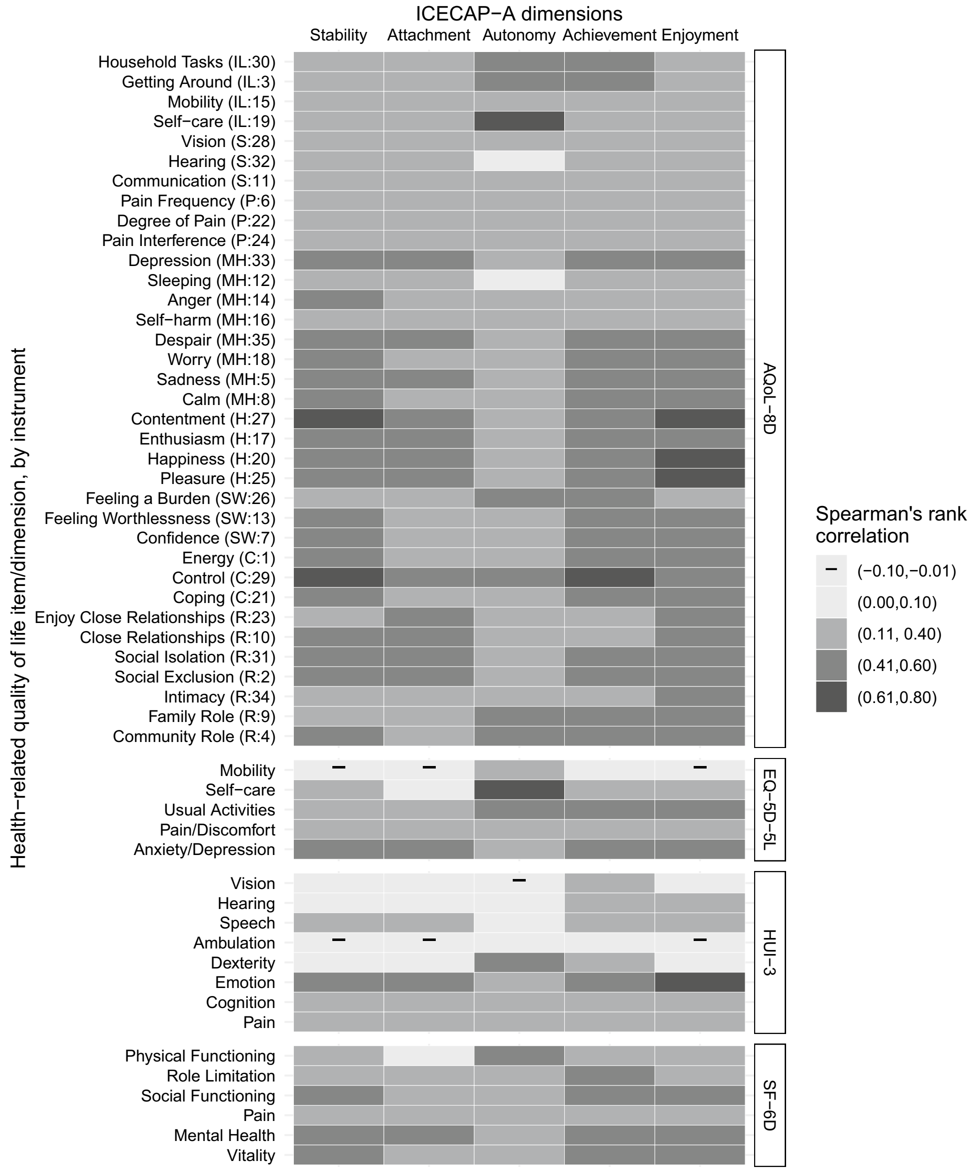

Fig. 2 Correlation matrix heatmap for the Spearman's rank correlations between ICECAP-A dimension levels and item-/dimension-level data for the AQoL-8D, EQ-5D-5L, HUI-3 and SF-6D. Item-/dimension-level coding was reversed for the health-related instruments. AQoL-8D items are grouped by dimension; for each AQoL-8D item, the corresponding dimension (*C* Coping, *H* Happiness, *IL* Independent Living, *MH* Mental Health, *P* Pain, *R* Relationships, *S* Senses, *SW* Self-worth) and item number (i.e., 1–35) are highlighted. The different shading reflects the categorical divisions of correlation coefficients reported by Shrout [119], used for illustrative purposes only

part in the SCI Community Survey [70], which was the first survey in Canada (and among the first worldwide) to draw a comprehensive picture of major aspects of the lives of people with SCI living in a community setting. Over 70% of participants in the current study had been injured for more than 10 years. It is unlikely that the relatively small proportions ($< 5\%$) of 'no capability' responses to the ICECAP-A dimensions would be replicated in a sample of individuals in the early stages of rehabilitation. Although the mean ICECAP-A index score observed in this study (0.761) is lower than the mean values reported by Mitchell et al. [121] for individuals living with arthritis, asthma, cancer, diabetes and heart disease (all between 0.80 and 0.82), the validity of the ICECAP-A in populations with significant capability impairments (SCI and non-SCI) is an area for further research.

5 Conclusions

This paper provides evidence to support the empirical validity of the ICECAP-A in the context of people with SCI living in a community setting, adding to the ever-growing literature

Table 6 Ten strongest and 10 weakest Spearman's rank correlations between ICECAP-A dimensions and items/dimensions of the EQ-5D-5L, HUI-3, SF-6D and AQoL-8D[a]

Rank	Comparator instrument: item/dimension[b]	ICECAP-A dimension	Correlation coefficient
1 (strongest)	AQoL-8D: self-care (IL:19)	Autonomy	0.636
2	AQoL-8D: pleasure (H:25)	Enjoyment	0.635
3	AQoL-8D: happiness (H:20)	Enjoyment	0.626
4	AQoL-8D: contentment (H:27)	Stability	0.625
5	AQoL-8D: contentment (H:27)	Enjoyment	0.622
6	HUI-3: Emotion	Enjoyment	0.612
7	AQoL-8D: control (C:29)	Achievement	0.609
8	AQoL-8D: control (C:29)	Stability	0.604
9	EQ-5D-5L: Self-care	Autonomy	0.601
10 (10th strongest)	AQoL-8D: sadness (MH:5)	Enjoyment	0.593
1 (weakest)	EQ-5D-5L: Mobility	Attachment	− 0.011
2	HUI-3: Ambulation	Achievement	0.015
3	HUI-3: Vision	Autonomy	− 0.019
4	HUI-3: Dexterity	Attachment	0.023
5	EQ-5D-5L: Mobility	Achievement	0.024
6	HUI-3: Vision	Attachment	0.034
7	HUI-3: Dexterity	Stability	0.042
8	HUI-3: Vision	Stability	0.043
9	EQ-5D-5L: Mobility	Enjoyment	− 0.045
10 (10th weakest)	HUI-3: Vision	Enjoyment	0.051

[a]The absolute values of the coefficients were used to identify the 10 strongest and 10 weakest correlations. Item-/dimension-level coding was reversed for the health-related instruments

[b]The corresponding AQoL-8D dimension and item number is highlighted for each AQoL-8D item reported in the table: *C* Coping, *H* Happiness, *IL* Independent Living, *MH* Mental Health

Table 7 Correlation coefficients for the pain-related and preference-weight convergent validity hypotheses

	Correlation coefficient
Pain-related hypothesis[a]	
Spearman's rank correlation between the ICECAP-A index score and…	
AQoL-8D item 22	0.315
AQoL-8D item 24	0.377
EQ-5D-5L Pain/Discomfort dimension (EQ-5D-5L item 4)	0.283
SF-36v2 item 21	0.294
SF-36v2 item 22	0.339
Preference-weight hypotheses	
Pearson's correlation between the ICECAP-A index score and…	
AQoL-8D index score	0.739
AQoL-8D psychosocial super-dimension score	0.699
AQoL-8D physical super-dimension score	0.482
EQ-5D-5L index score	0.568
HUI-3 index score	0.499
SF-6D index score	0.577

[a]Item-/dimension-level coding was reversed for the health-related instruments. The wording of the AQoL-8D, EQ-5D-5L, HUI and SF-6D items and response options included in the pain-related hypothesis is provided in the Electronic Supplementary Material (ESM1)

that explores whether instrument *X* is suitable for use in context *Y*. The availability of different preference-based instruments, overlapping in their coverage of evaluative space (e.g., health, capabilities and social care), offers a rich armoury for analysts to quantify the benefits of treatments and interventions. The collective role of such instruments to support efficient and equitable decision making (and, indeed, having a better understanding of the needs of decision makers) is an area that requires further attention.

Declarations

Funding Financial support for this study was through a grant from the Rick Hansen Institute (Rick Hansen Institute Translational Research Program, grant #2012-29: Spinal Cord Injury & Secondary Complications: A Mixed-Methods Evaluation of Preference-Based Instruments). The funding agreement ensured the research team's independence in designing the study, interpreting the data, and writing and publishing manuscripts. We would also like to thank the Office of the Vice-President, Research at Simon Fraser University for the Undergraduate Student Research Award for CM, and Health Utilities Inc. for the New Investigator Grant awarded to DGTW.

Conflicts of Interest DGTW and SB are members of the EuroQol Group Association. CM, VKN, SB, and DGTW have no further conflicts of interest that are directly relevant to the content of this article.

Ethics Approval The study was approved by the University of British Columbia Behavioural Research Ethics Board (H12-01138) and Vancouver Coastal Health Authority (#V12-01138).

Availability of Data and Material No open access availability. Individuals interested in gaining access to the survey data should contact the corresponding author.

Consent to Participate All study participants provided informed consent to participate (covered in the ethics approval).

Consent for Publication The authors' intent to publish findings from the study was explicit in the informed consent procedure.

Code Availability Not available.

Author Contributions VKN, SB and DGTW were investigators for the primary study. CM and DGTW designed the secondary analysis plan, performed data analysis and drafted the original manuscript. All authors were involved in the interpretation of results, review of the draft manuscript, and read and approved the final version.

References

1. Canadian Agency for Drugs and Technologies in Health. Guidelines for the economic evaluation of health technologies: Canada (4th ed.). Ottawa: Canadian Agency for Drugs and Technologies in Health. 2017. https://www.cadth.ca/sites/default/files/pdf/guidelines_for_the_economic_evaluation_of_health_technologies_canada_4th_ed.pdf. Accessed 6 Aug 2020.
2. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. London: National Institute for Health and Care Excellence. 2013. https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781. Accessed 6 Aug 2020.
3. Pharmaceutical Benefits Advisory Committee. Guidelines for preparing submissions to the Pharmaceutical Benefits Advisory Committee (Version 5.0). Canberra: Pharmaceutical Benefits Advisory Committee. 2016. https://pbac.pbs.gov.au/content/information/files/pbac-guidelines-version-5.pdf. Accessed 6 Aug 2020.
4. Brouwer WBF, Culyer AJ, van Exel NJA, Rutten FFH. Welfarism vs. extra-welfarism. J Health Econ. 2008;27(2):325–38.
5. Coast J, Smith R, Lorgelly P. Should the capability approach be applied in Health Economics? Health Econ. 2008;17(6):667–70.
6. Hurley J. Welfarism, extra-welfarism and evaluative economic analysis in the health care sector. In: Barer ML, Getzen TE, Stoddard GL, editors. Health, health care and health economics: perspectives on distribution. Chichester: Wiley; 1998.
7. Brouwer WB, Koopmanschap MA. On the economic foundations of CEA. Ladies and gentlemen, take your positions! J Health Econ. 2000;19(4):439–59.
8. Brazier JE, Rowen D, Mavranezouli I, Tsuchiya A, Young T, Yang Y, et al. Developing and testing methods for deriving preference-based measures of health from condition-specific measures (and other patient-based measures of outcome). Health Technol Assess. 2012;16(32):1–114.
9. Brazier J, Tsuchiya A. Preference-based condition-specific measures of health: what happens to cross programme comparability? Health Econ. 2010;19(2):125–9.
10. Goodwin E, Green C. A systematic review of the literature on the development of condition-specific preference-based measures of health. Appl Health Econ Health Policy. 2016;14(2):161–83.
11. Richardson J, McKie J, Bariola E. Multiattribute utility instruments and their use. In: Culyer AJ, editor. Encyclopedia of health economics. San Diego: Elsevier; 2014. p. 341–57.
12. Devlin NJ, Brooks R. EQ-5D and the EuroQol group: past, present and future. Appl Health Econ Health Policy. 2017;15(2):127–37.
13. Poder TG, Gandji EW. SF6D value sets: a systematic review. Value in Health. 2016;19(3):A282.
14. Neuman PJ, Goldie SJ, Weinstein MC. Preference-based measures in economic evaluation in health care. Annu Rev Public Health. 2000;21:587–611.
15. Brazier JE, Ratcliffe J, Saloman J, Tsuchiya A. Measuring and valuing health benefits for economic evaluation. 2nd ed. Oxford: Oxford University Press; 2017.
16. Furlong WJ, Feeny DH, Torrance GW, Barr RD. The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies. Ann Med. 2001;33(5):375–84.
17. Brooks R. EuroQol: the current state of play. Health Policy. 1996;37(1):53–72.
18. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727–36.
19. Hawthorne G, Richardson J, Osborne R. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health-related quality of life. Qual Life Res. 1999;8(3):209–24.
20. Misajon RA, Hawthorne G, Richardson J, Barton J, Peacock S, Iezzi A, et al. Vision and quality of life: the development of a utility measure. Invest Ophthalmol Vis Sci. 2005;46(11):4007–15.
21. Richardson JRJ, Peacock SJ, Hawthorne G, Iezzi A, Elsworth G, Day NA. Construction of the descriptive system for the assessment of quality of life AQoL-6D utility instrument. Health Qual Life Outcomes. 2012;10:38.

22. Richardson J, Iezzi A, Khan MA, Maxwell A. Validity and reliability of the assessment of quality of life (AQoL)-8D multi-attribute utility instrument. Patient. 2014;7(1):85–96.
23. Mooney G. Beyond health outcomes: the benefits of health care. Health Care Anal. 1998;6(2):99–105.
24. Ryan M, Shackley P. Assessing the benefits of health care: how far should we go? Quality in Health Care. 1995;4(3):207–13.
25. Netten A, Burge P, Malley J, Potoglou D, Towers AM, Brazier J, et al. Outcomes of social care for adults: developing a preference-weighted measure. Health Technol Assess. 2012;16(16):1–165.
26. Personal Social Services Research Unit. Which ASCOT tool should I use? https://www.pssru.ac.uk/ascot/tools/. Accessed 6 Aug 2020.
27. Coast J, Flynn T, Sutton E, Al-Janabi H, Vosper J, Lavender S, et al. Investigating Choice Experiments for Preferences of Older People (ICEPOP): evaluative spaces in health economics. J Health Serv Res Policy. 2008;13(Suppl 3):31–7.
28. Connell J, Carlton J, Peasgood T, Mukuria C, Brazier J, Augustovski F, et al. Extending the QALY: developing and testing the proposed items for a new generic measure—results from qualitative review and face validity with patients, social care users and carers. Qual Life Res. 2018;27(Suppl 1):S189.
29. School of Health and Related Research, University of Sheffield (UK). Extending the QALY. https://scharr.dept.shef.ac.uk/e-qaly/. Accessed 6 Aug 2020.
30. Sen A. Capability and well-being. In: Nussbaumn M, Sen A, editors. The quality of life. Oxford: Oxford University Press; 1993. p. 30–53.
31. Al-Janabi H. Do capability and functioning differ? A study of U.K. survey responses. Health Econ. 2018;27(3):465–79.
32. Grewal I, Lewis J, Flynn T, Brown J, Bond J, Coast J. Developing attributes for a generic quality of life measure for older people: preferences or capabilities? Soc Sci Med. 2006;62(8):1891–901.
33. Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. Qual Life Res. 2012;21(1):167–76.
34. Sutton EJ, Coast J. Development of a supportive care measure for economic evaluation of end-of-life care using qualitative methods. Palliat Med. 2014;28(2):151–7.
35. Canaway A, Al-Janabi H, Kinghorn P, Bailey C, Coast J. Development of a measure (ICECAP-Close Person Measure) through qualitative methods to capture the benefits of end-of-life care to those close to the dying for use in economic evaluation. Palliat Med. 2017;31(1):53–62.
36. Coast J. Assessing capability in economic evaluation: a life course approach? Eur J Health Econ. 2019;20(6):779–84.
37. Flynn T, Huynh E, Peters TJ, Al-Janabi H, Clemens S, Moody A, et al. Scoring the ICECAP-A capability instrument. Estimation of a UK general population tariff. Health Econ. 2015;24(3):258–69.
38. Coast J, Flynn TN, Natarajan L, Sproston K, Lewis J, Louviere JJ, et al. Valuing the ICECAP capability index for older people. Soc Sci Med. 2008;67(5):874–82.
39. Coast J, Smith RD, Lorgelly P. Welfarism, extra-welfarism and capability: the spread of ideas in health economics. Soc Sci Med. 2008;67(7):1190–8.
40. Krucien N, Watson V, Ryan M. Is Best-Worst Scaling Suitable for Health State Valuation? A Comparison with Discrete Choice Experiments. Health Econ. 2017;26(12):e1–16.
41. Mitchell PM, Roberts TE, Barton PM, Coast J. Assessing sufficient capability: a new approach to economic evaluation. Soc Sci Med. 2015;139:71–9.
42. National Institute for Health and Care Excellence. *The social care guidance manual*. London: National Institute for Health and Care Excellence. 2013. https://www.nice.org.uk/process/pmg10/resources/the-social-care-guidance-manual-pdf-72286648234693. Accessed 6 Aug 2020.
43. Zorginstituut Nederland. *Guideline for economic evaluations in healthcare*. Diemen: Zorginstituut Nederland. 2016. https://english.zorginstituutnederland.nl/publications/reports/2016/06/16/guideline-for-economic-evaluations-in-healthcare. Accessed 6 Aug 2020.
44. Afentou N, Kinghorn P. A systematic review of the feasibility and psychometric properties of the ICEpop CAPability measure for adults and its use so far in economic evaluation. Value Health. 2020;23(4):515–26.
45. Al-Janabi H, Keeley T, Mitchell P, Coast J. Can capabilities be self-reported? A think aloud study. Soc Sci Med. 2013;87:116–22.
46. Engel L, Mortimer D, Bryan S, Lear SA, Whitehurst DGT. An investigation of the overlap between the ICECAP-A and five preference-based health-related quality of life instruments. PharmacoEconomics. 2017;35(7):741–53.
47. Al-Janabi H, Peters TJ, Brazier J, Bryan S, Flynn TN, Clemens S, et al. An investigation of the construct validity of the ICECAP-A capability measure. Qual Life Res. 2013;22(7):1831–40.
48. Chen G, Ratcliffe J, Kaambwa B, McCaffrey N, Richardson J. Empirical comparison between capability and two health-related quality of life measures. Soc Indic Res. 2018;140(1):175–90.
49. Richardson J, Iezzi A, Khan MA, Chen G, Maxwell A. Measuring the sensitivity and construct validity of 6 utility instruments in 7 disease areas. Med Decis Mak. 2016;36(2):147–59.
50. Al-Janabi H, Flynn T, Peters T, Bryan S, Coast J. Test-retest reliability of capability measurement in the UK population. Health Econ. 2015;19(11):1300–17.
51. Goranitis I, Coast J, Al-Janabi H, Latthe P, Roberts TE. The validity and responsiveness of the ICECAP-A capability-well-being measure in women with irritative lower urinary tract symptoms. Qual Life Res. 2016;25(8):2063–75.
52. Keeley T, Al-Janabi H, Nicholls E, Foster NE, Jowett S, Coast J. A longitudinal assessment of the responsiveness of the ICECAP-A in a randomised controlled trial of a knee pain intervention. Qual Life Res. 2015;24(10):2319–31.
53. Keeley T, Coast J, Nicholls E, Foster NE, Jowett S, Al-Janabi H. An analysis of the complementarity of ICECAP-A and EQ-5D-3 L in an adult population of patients with knee pain. Health Qual Life Outcomes. 2016;14:36.
54. Engel L, Chen G, Richardson J, Mihalopoulos C. The impact of depression on health-related quality of life and wellbeing: identifying important dimensions and assessing their inclusion in multi-attribute utility instruments. Qual Life Res. 2018;27(11):2873–84.
55. Mitchell PM, Al-Janabi H, Byford S, Kuyken W, Richardson J, Iezzi A, et al. Assessing the validity of the ICECAP-A capability measure for adults with depression. BMC Psychiatry. 2017;17(1):1–13.
56. Goranitis I, Coast J, Day E, Copello A, Freemantle N, Seddon J, et al. Measuring health and broader well-being benefits in the context of opiate dependence: the Psychometric Performance of the ICECAP-A and the EQ-5D-5L. Value in Health. 2016;19(6):820–8.
57. Engel L, Bryan S, Noonan VK, Whitehurst DGT. Using path analysis to investigate the relationships between standardized instruments that measure health-related quality of life, capability wellbeing and subjective wellbeing: an application in the context of spinal cord injury. Soc Sci Med. 2018;213:154–64.
58. Strauss DJ, DeVivo MJ, Paculdo DR, Shavelle RM. Trends in life expectancy after spinal cord injury. Arch Phys Med Rehabil. 2006;87(8):1079–85.
59. Cobb J, Dumont FS, Leblond J, Park SE, Noonan VK, Noreau L. An exploratory analysis of the potential association between SCI

secondary health conditions and daily activities. Topics Spinal Cord lnj Rehabil. 2014;20(4):277–88.
60. Rivers CS, Fallah N, Noonan VK, Whitehurst DG, Schwartz CE, Finkelstein JA, et al. Health conditions: effect on function, health-related quality of life, and life satisfaction after traumatic spinal cord injury. A prospective observational Registry Cohort Study. Arch Phys Med Rehabil. 2018;99(3):443–51.
61. Jang Y, Hsieh CL, Wang YH, Wu YH. A validity study of the WHOQOL-BREF assessment in persons with traumatic spinal cord injury. Arch Phys Med Rehabil. 2004;85(11):1890–5.
62. Bussière C, Sicsic J, Pelletier-Fleury N. Simultaneous effect of disabling conditions on primary health care use through a capability approach. Soc Sci Med. 2016;154:70–84.
63. Mitra S. The capability approach and disability. J Disabil Policy Stud. 2006;16(4):236–47.
64. Trani J-F, Bakhshi P, Brown D, Lopez D, Gall F. Disability as deprivation of capabilities: estimation using a large-scale survey in Morocco and Tunisia and an instrumental variable approach. Soc Sci Med. 2018;211:48–60.
65. Anand P, Dolan P. Equity, capabilities and health. Soc Sci Med. 2005;60(2):219–22.
66. Engel L, Bryan S, Evers SMAA, Dirksen CD, Noonan VK, Whitehurst DGT. Exploring psychometric properties of the SF-6D, a preference-based health-related quality of life measure, in the context of spinal cord injury. Qual Life Res. 2014;23(8):2383–93.
67. Michel YA, Engel L, Rand-Hendriksen K, Augestad LA, Whitehurst DGT. "When I saw walking I just kind of took it as wheeling": interpretations of mobility-related items in generic, preference-based health state instruments in the context of spinal cord injury. Health Qual Life Outcomes. 2016;14:164.
68. Whitehurst DGT, Suryaprakash N, Engel L, Mittmann N, Noonan VK, Dvorak MFS, et al. Perceptions of individuals living with spinal cord injury toward preference-based quality of life instruments: a qualitative exploration. Health Qual Life Outcomes. 2014;12:50.
69. Whitehurst DGT, Mittmann N, Noonan VK, Dvorak MF, Bryan S. Health state descriptions, valuations and individuals' capacity to walk: a comparative evaluation of preference-based instruments in the context of spinal cord injury. Qual Life Res. 2016;25(10):2481–96.
70. Noreau L, Leblond J, Dumont FS. Spinal cord injury community survey: a National, Comprehensive Study to Portray the Lives of Canadians with Spinal Cord Injury. Topics Spinal Cord lnj Rehabil. 2014;20(4):249–64.
71. Perreault WD. Controlling Order-Effect Bias. Public Opinion Quarterly. 1975;39(4):544–51.
72. Xie F, Pullenayegum E, Gaebel K, Bansback N, Bryan S, Ohinmaa A, et al. A time trade-off-derived value set of the EQ-5D-5L for Canada. Med Care. 2016;54(1):98–105.
73. Feeny D, Furlong W, Torrance GW, Goldsmith CH, Zhu Z, DePauw Sonja, Denton M, et al. Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. Med Care. 2002;40(2):113–28.
74. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure from the SF-36. J Health Econ. 2002;21(2):271–92.
75. Brazier J, Rowen D, Hanmer J. Revised SF-6D scoring programmes: a summary of improvements. PRO Newslett. 2008;40:14–5.
76. Ware JE, Kosinski M, Bjorner JB, Turner-Bowker DM, Gandek B, Maruish ME. User's manual for the SF-36v2TM health survey. Lincoln, RI. 2007.
77. Statistics Canada. Canadian Community Health Survey (CCHS). Annual Component—2010 Questionnaire. 2010.
78. Kalpakjian CZ, Scelza WM, Forchheimer MB, Toussaint LL. Preliminary reliability and validity of a spinal cord injury secondary conditions scale. J Spinal Cord Med. 2007;30(2):131–9.
79. Catz A, Itzkovich M, Tesio L, Biering-Sørensen F, Weeks C, Laramee MT, et al. A multicenter international study on the spinal cord independence measure, version III: rasch psychometric validation. Spinal Cord. 2007;45(4):275–91.
80. Fekete C, Eriks-Hoogland I, Baumberger M, Catz A, Itzkovich M, Lüthi H, et al. Development and validation of a self-report version of the Spinal Cord Independence Measure (SCIM III). Spinal Cord. 2013;51(1):40–7.
81. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
82. Putzke JD, Richards JS, Hicken BL, DeVivo MJ. Predictors of life satisfaction: a spinal cord injury cohort study. Arch Phys Med Rehabil. 2002;83(4):555–61.
83. van Leeuwen CMC, Kraaijeveld S, Lindeman E, Post MWM. Associations between psychological factors and quality of life ratings in persons with spinal cord injury: a systematic review. Spinal Cord. 2012;50(3):174–87.
84. Kreuter M, Siösteen A, Erkholm B, Byström U, Brown DJ. Health and quality of life of persons with spinal cord lesion in Australia and Sweden. Spinal Cord. 2005;43(2):123–9.
85. van Leeuwen CMC, Post MW, Westers P, Van Der Woude LH, De Groot S, Sluis T, et al. Relationships between activities, participation, personal factors, mental health, and life satisfaction in persons with spinal cord injury. Arch Phys Med Rehabil. 2012;93(1):82–9.
86. Dijkers M. Quality of life after spinal cord injury: a meta analysis of the effects of disablement components. Spinal Cord. 1997;35(12):829–40.
87. Chang FH, Wang YH, Jang Y, Wang CW. Factors associated with quality of life among people with spinal cord injury: application of the international classification of functioning, disability and health model. Arch Phys Med Rehabil. 2012;93(12):2264–70.
88. Barker RN, Kendall MD, Amsters DI, Pershouse KJ, Haines TP, Kuipers P. The relationship between quality of life and disability across the lifespan for people with spinal cord injury. Spinal Cord. 2009;47(2):149–55.
89. Post MWM, De Witte LP, Van Asbeck FWA, Van Dijk AJ, Schrijvers AJP. Predictors of health status and life satisfaction in spinal cord injury. Arch Phys Med Rehabil. 1998;79(4):395–401.
90. Dijkers MPJ. Correlates of Life Satisfaction Among Persons With Spinal Cord Injury. Arch Phys Med Rehabil. 1999;80(8):867–76.
91. Adriaansen JJE, Ruijs LEM, Van Koppenhagen CF, Van Asbeck FWA, Snoek GJ, Van Kuppevelt D, et al. Secondary health conditions and quality of life in persons living with spinal cord injury for at least ten years. J Rehabil Med. 2016;48(10):853–60.
92. Gurcay E, Bal A, Eksioglu E, Cakci A. Quality of life in patients with spinal cord injury. Int J Rehabil Res. 2010;33(4):356–8.
93. Hackert MQN, van Exel J, Brouwer WBF. Valid outcome measures in care for older people: comparing the ASCOT and the ICECAP-O. Value Health. 2017;20(7):936–44.
94. Coast J, Peters TJ, Natarajan L, Sproston K, Flynn T. An assessment of the construct validity of the descriptive system for the ICECAP capability measure for older people. Qual Life Res. 2008;17(7):967–76.
95. Davis JC, Liu-Ambrose T, Richardson CG, Bryan S. A comparison of the ICECAP-O with EQ-5D in a falls prevention clinical setting: are they complements or substitutes? Qual Life Res. 2013;22(5):969–77.

96. Geyh S, Ballert C, Sinnott A, Charlifue S, Catz A, D'Andrea Greve JM, et al. Quality of life after spinal cord injury: a comparison across six countries. Spinal Cord. 2013;51(4):322–6.
97. McColl MA, Arnold R, Charlifue S, Glass C, Savic G, Frankel H. Aging, spinal cord injury, and quality of life: structural relationships. Arch Phys Med Rehabil. 2003;84(8):1137–44.
98. Westgren N, Levi R. Quality of life and traumatic spinal cord injury. Arch Phys Med Rehabil. 1998;79(11):1433–9.
99. Vogel LD, Krajci KA, Anderson C. Adults with pediatric-onset spinal cord injuries: part 3: impact of medical complications. J Spinal Cord Med. 2002;25(4):297–305.
100. Haran MJ, Lee BB, King MT, Marial O, Stockler MR. Health status rated with the medical outcomes study 36-item short-form health survey after spinal cord injury. Arch Phys Med Rehabil. 2005;86(12):2290–5.
101. Craven C, Hitzig SL, Mittmann N. Impact of impairment and secondary health conditions on health preference among Canadians with chronic spinal cord injury. J Spinal Cord Med. 2012;35(5):361–70.
102. Jörgensen S, Hedgren L, Sundelin A, Lexell J. Global and domain-specific life satisfaction among older adults with long-term spinal cord injury. J Spinal Cord Med. 2019. https://doi.org/10.1080/10790268.2019.1610618.
103. Leduc BE, Lepage Y. Health-related quality of life after spinal cord injury. Disabil Rehabil. 2002;24(4):196–202.
104. Holicky R, Charlifue S. Ageing and spinal cord injury: the impact of spousal support. Disabil Rehabil. 1999;21(5–6):250–7.
105. Putzke JD, Elliott TR, Richards JS. Marital status and adjustment 1 year post-spinal-cord-injury. J Clin Psychol Med Settings. 2001;8(2):101–7.
106. McColl MA, Stirling P, Walker J, Corey P, Wilkins R. Expectations of independence and life satisfaction among ageing spinal cord injured adults. Disabil Rehabil. 1999;21(5–6):231–40.
107. Clayton KS, Chubon RA. Factors associated with the quality of life of long-term spinal cord injured persons. Arch Phys Med Rehabil. 1994;75(6):633–8.
108. Fekete C, Siegrist J, Reinhardt JD, Brinkhof MWG. Is Financial Hardship Associated with Reduced Health in Disability? The Case of Spinal Cord Injury in Switzerland. PLoS One. 2014;9(2):e90130.
109. Iorio-Morin C, Noonan VK, White B, Noreau L, Leblond J, Dumont FS, et al. Quality of life and health utility scores among Canadians living with traumatic spinal cord injury—a national cross-sectional study. Spine. 2018;43(14):999–1006.
110. Elfström ML, Rydén A, Kreuter M, Taft C, Sullivan M. Relations between coping strategies and health-related quality of life in patients with spinal cord lesion. J Rehabil Med. 2005;37(1):9–16.
111. Krause JS. Aging and life adjustment after spinal cord injury. Spinal Cord. 1998;36(5):320–8.
112. Post MWM, Reinhardt JD. Participation and life satisfaction in aged people with spinal cord injury: does age at onset make a difference? Topics Spinal Cord Inj Rehabil. 2015;21(3):233–40.
113. Stensman R. Adjustment to traumatic spinal cord injury. A longitudinal study of self-reported quality of life. Paraplegia. 1994;32(6):416–22.
114. Zeileis A. Econometric computing with HC and HAC covariance matrix estimators. J Stat Softw. 2004;11(10):1–17.
115. Streiner DL, Norman GR, Cairney J. Health measurement scales: a practical guide to their development and use. 5th ed. New York: Oxford University Press; 2015.
116. Richardson J, Khan MA, Iezzi A, Maxwell A. Comparing and explaining differences in the magnitude, content, and sensitivity of utilities predicted by the EQ-5D, SF-6D, HUI 3, 15D, QWB, and AQoL-8D multiattribute utility instruments. Med Decis Making. 2015;35(3):276–91.
117. Richardson J, Iezzi A, Khan MA. Why do multi-attribute utility instruments produce different utilities: the relative importance of the descriptive systems, scale and 'micro-utility' effects. Qual Life Res. 2015;24(8):2045–53.
118. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019.
119. Shrout PE. Measurement reliability and agreement. Stat Methods Med Res. 1998;7(3):301–17.
120. Kinghorn P. Using deliberative methods to establish a sufficient state of capability well-being for use in decision-making in the contexts of public health and social care. Soc Sci Med. 2019;240:112546.
121. Mitchell PM, Al-Janabi H, Richardson J, Iezzi A, Coast J. The relative impacts of disease on health status and capability wellbeing: a multi-country study. PLoS One. 2015;10(12):e0143590.