



# Comparing the German Translation of the ICECAP-A Capability Wellbeing Measure to the Original English Version: Psychometric Properties across Healthy Samples and Seven Health Condition Groups

Myles-Jay Linton<sup>1,2</sup> · Paul Mark Mitchell<sup>1</sup>  · Hareth Al-Janabi<sup>3</sup> · Michael Schlander<sup>4,5,6</sup> · Jeff Richardson<sup>7</sup> · Angelo Iezzi<sup>7</sup> · Jasper Ubels<sup>4,6</sup> · Joanna Coast<sup>1,2</sup>

Received: 17 July 2018 / Accepted: 26 November 2018 / Published online: 7 December 2018  
© The Author(s) 2018

## Abstract

As the cross-cultural use of outcome measures grows, it is important to determine whether these instruments are: appropriate for use in other settings, translated accurately, and perform in a similar manner to their original tools. This research aimed to compare the validity of the German translation of the ICECAP-A to the original English version of the instrument, across healthy adults and seven health condition groups (arthritis, asthma, cancer, depression, diabetes, hearing loss and heart disease). Data were analysed from a cross-cultural study, which recruited participants through online panels in 2012. Data were analysed on capability wellbeing (ICECAP-A), health-related quality of life (EQ-5D-5L and SF-6D), satisfaction with life (SWLS), and a series of other condition-specific outcome measures. The ICECAP-A was assessed for internal consistency, convergent validity and construct validity. 2501 individuals were included in the analysis. The ICECAP-A demonstrated good internal consistency within Germany and the UK population, and across all seven health condition sub-groups ( $\alpha = .74-.86$ ). In both countries, ICECAP-A scores were significantly correlated with SWLS, SF-6D and EQ-5D-5L scores for healthy participants and health condition groups ( $r = .35-.77$ ). Finally, experiencing one of the seven health conditions (compared to being healthy) was significantly associated with lower levels of capability wellbeing in the German and UK samples (construct validity). The German translation of the ICECAP-A yielded valid and reliable data, in both healthy respondents and the seven health condition groups. Further work could be undertaken to develop a German specific value-set for the ICECAP-A.

**Keywords** ICECAP-A · German translation · Cultural validation

---

✉ Paul Mark Mitchell  
paul.mitchell@bristol.ac.uk

## Background

There is increasing pressure for economic evaluations to acknowledge the impact of health and social care interventions on the wider aspects of wellbeing, above and beyond health-related quality of life (Brazier and Tsuchiya 2015). Several western governments have also sought to incorporate comparatively ‘subjective’ forms of wellbeing alongside more ‘traditional’ economic indicators in their assessments of societal progress (Hicks et al. 2013; Stiglitz et al. 2010). As the measurement of wellbeing continues to grow (Linton et al. 2016), it becomes increasingly important to assess how well these tools are able to perform across countries, populations and settings.

The ICECAP tools represent one approach to the measurement of wellbeing for use in economic evaluations. These tools were influenced by the work of Nobel Prize winning economist, Amartya Sen (Sen 1993), and conceptualise wellbeing in terms of how capable a person is to achieve a personally valuable life (Coast et al. 2015). The primary rationale for focussing on capability in the health field is a growing concern that measures may be missing important aspects of quality of life; currently measures used in health economic analysis tend to focus primarily on health (Coast et al. 2008b). The ICECAP-O was the first generic measure of capability wellbeing for use across patient populations used in health economics, and was specifically designed for use in older-adults (i.e. 65 years old and over) (Coast et al. 2008a; Grewal et al. 2006). Following this, ICECAP-A was developed for use in the general adult population (i.e. 18 years old and over) (Al-Janabi et al. 2012). ICECAP measures are now being recommended for use in economic evaluations assessing interventions in social care and long-term conditions, both in the UK where the ICECAP measures were originally developed (National Institute for Health and Care Excellence 2014) but also internationally (Versteegh et al. 2016)

Given the impact of economic evaluations on policy and practice, it remains vital to assess the quality of the instruments involved in this process (Mokkink et al. 2010). Psychometric evidence provides a source of insight into the validity and reliability of the measures in use, but these investigations are infrequently conducted across different cultural settings. Nonetheless, experiences of health and wellbeing may differ between countries and there is a need for internationally used instruments to be comparable and applicable across cultures (Taggart et al. 2013).

A growing literature on the psychometric performance of the ICECAP tools exists. In the UK, studies have demonstrated that the ICECAP-A is responsive to deteriorations in clinical symptoms among women with irritative lower urinary tract symptoms (Goranitis et al. 2016) and to changes in health-related quality of life following knee pain (Keeley et al. 2015). Additionally, support has been found for the convergent validity of the attributes within the ICECAP-A with the attributes of the EQ-5D, in the general population (Al-Janabi et al. 2013). Convergent validity describes the extent to which scores produced by a new instrument are correlated with scores on comparable (or conceptually identical) existing instruments (DeVon et al. 2007). Cross-culturally, the ICECAP-O has demonstrated good convergent validity with existing measures in a sample of post-hospitalised older adults in the Netherlands (Makai et al. 2013) and among nursing home residents with dementia in Germany (Makai et al. 2014). Similar studies have demonstrated the reliability and validity of the ICECAP-O in Spain (Sarabia-Cobo et al. 2017) and Sweden (Hörder et al. 2016) and for the ICECAP-A

in China (Tang et al. 2018). However, to date, the validity of a German translation of the ICECAP-A has not been investigated. In addition, no translated ICECAP-A or ICECAP-O measure has been directly compared to the original version using similar population groups across different countries.

The challenge of translating instruments into different languages is critical to the use of outcomes measures across countries. One of the central challenges here concerns the extent to which questionnaire items will be interpreted in the same way across cultures, and in alternative languages (Reeve et al. 2013). Specific guidance on the translation of outcome measures highlights how poor translation may result in new measures that misinterpret the concepts underpinning the original tools or violate the regular speech patterns of the target language (Wild et al. 2005). In summary, the translation of outcome measures requires careful consideration, in addition to an exploration of psychometric properties.

The aim of this work was to quantitatively assess the psychometric comparability of the newly translated German version of the ICECAP-A to the original English language version. To further investigate the validity of the ICECAP-A across contexts, this research investigates the psychometric performance of the tool in healthy and health condition populations. Although ICECAP measures have been translated into numerous languages, this is the first psychometric study to simultaneously compare a translated ICECAP instrument in its country of use (Germany) with the original English language version of the instrument in the UK. This study will contribute to the growing literature of articles evaluating the quality and applicability of capability wellbeing measures.

## Methods

### Dataset

This study used data collected as part of the Multi Instrument Comparison (MIC) dataset, a large study of health and wellbeing measures collected across different population groups and countries in 2012 (Richardson et al. 2015b). The survey was conducted by a global panel company CINT Pty Ltd., using online panels to recruit relevant individuals. Participants were classified into health condition groups if they self-reported one of seven primary health conditions (asthma, arthritis, cancer, depression, diabetes, hearing loss or heart disease). ‘Healthy’ participants were defined as those reporting 70 or higher on a 0–100 visual analogue scale measuring overall health, and without any other illnesses lasting longer than three months in the past year. This study utilised data collected in Germany and the UK. Quotas were employed to obtain a representative sample in terms of age, sex and education in the healthy population, while target quotas of 150 individuals per health condition group per country were employed to reach similar numbers of health condition groups within and across countries. The survey was approved by the Monash University Human Research Ethics Committee (reference number: CF11/3192–2,011,001,748). In addition to questions related to health and wellbeing, standard sociodemographic data were collected (age, sex, gender, educational attainment).

## Measures

### Capability Wellbeing

The ICECAP-A comprises five questionnaire items with four response levels each, and is designed to capture people's capability to live a life that they value (Al-Janabi et al. 2012). These 'capabilities' are *stability* (feeling settled and secure), *attachment* (able to achieve love, friendship and support), *autonomy* (able to be independent), *achievement* (able to achieve and progress in life) and *enjoyment* (able to experience enjoyment and pleasure) (Al-Janabi et al. 2012). The methodology used to convert raw scores into capability values is described elsewhere (Flynn et al. 2015), however scores range from 0 which represents 'no capability' to 1 which represents 'full capability'. In the absence of a scoring tariff for both countries, UK scores were applied to the German and UK sample.

In accordance with available guidance on the topic, preparation for the translation involved establishing a multi-lingual study team and undergoing a process of familiarisation with the concepts of the ICECAP-A (Wild et al. 2005). Next, a first language German speaker was identified by the Australian MIC study team to forward translate the instrument. Plausibility and accuracy of this translation was undertaken in collaboration with two members of the German MIC data collaborators (including MS) prior to data collection. Subsequent to data collection, a back-translation was undertaken by one German native with prolonged experience living in English speaking countries and who did not have any knowledge of the ICECAP-A. The back-translated English version was then compared with the German translation by an American native with prolonged experience living in Germany, who also did not know the ICECAP-A, to check for any literal and semantic mistakes. Semi-structured interviews were conducted by MS and JU, who are both familiar with the ICECAP-A, with the aim to understand if there were conceptual differences between the original ICECAP-A and the back translated version. It was concluded that the semantic translation was conducted well, with little if any potential for improvement on the translated version (further information available on request). Both the original, and the translated version of the ICECAP-A are available through the ICECAP project website (<https://www.birmingham.ac.uk/ICECAP>).

### Health Status

The EQ-5D is the most common measure used in health economics to generate quality-adjusted life years (QALYs) (Wisløff et al. 2014). The tool consists of five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression), with the most recent version (EQ-5D-5L) consisting of five response levels per dimension (no problems, slight problems, moderate problems, severe problems and extreme problems). English population values have been developed for the EQ-5D-5L (Devlin et al. 2017). Values are used to estimate the relative preferences for different health states as measured by the concepts captured in the measure. Values on the EQ-5D-5L are anchored on a 0–1 dead to full health scale, with values below zero possible (minimum value for England EQ-5D-5L value set of  $-0.285$ ). The EQ-5D-5L was translated into German and validated in a previous study (Hinz et al. 2014). The English population value-set for the EQ-5D-5L was applied to the UK and German sample in the absence of a German value-set.

The Short Form 36 (SF-36) is a widely used instrument for measuring generic health status worldwide. It consists of 36 items that are compiled into eight sub-scales: vitality, physical functioning, bodily pain, general health perceptions, physical role functioning, emotional role functioning, social role functioning and mental health (Ware Jr and Sherbourne 1992). For each of the eight sub-scales, scores are summed to a 0 (worst health state) to 100 (best health state). The SF-36 was also previously translated into a German language version and validated within an existing study (Bullinger 1995). SF-6D scores were derived from responses to the SF-36 using econometric modelling methods (Brazier et al. 2002) in line with previous studies using the Multi-Instrument Comparison (MIC) dataset (Richardson et al. 2015a).

### Satisfaction with Life

The Satisfaction with Life Scale (SWLS) was developed as a measure of ‘life satisfaction’, a key component of subjective well-being (Diener et al. 1985). The scale consists of five items and is scored on a five-point Likert scale (ranging from strongly disagree to strongly agree). An overall scale score is calculated by summing the item scores; higher scores indicate greater levels of life satisfaction. SWLS was chosen after previous research in English speaking samples suggest subjective wellbeing measures have comparable correlation with the ICECAP-A and the physical and mental health subscales of the SF-36 (Richardson et al. 2016).

### Condition Severity Measures

The severity of disease burden experienced by participants in each of the health condition sub-samples was measured using the following condition specific tools: Arthritis Impact Measurement Scales (Guillemin et al. 1997), Asthma Quality of Life Questionnaire (Marks et al. 1992), European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Cancer-30 (Aaronson et al. 1993), Depression Anxiety and Stress Scale (Lovibond and Lovibond 1995), Diabetes-39 Questionnaire (Boyer and Earp 1997), Abbreviated Profile of Hearing Aid Benefit (Cox and Alexander 1995) and MacNew Heart Disease HRQL questionnaire (Höfer et al. 2004). Scores on these instruments were transformed into total scale scores ranging from ‘0’ (low) to ‘1’ (severe) health condition severity. A full outline of the method used to develop these scores has been detailed previously (Mitchell et al. 2015).

### Statistical Analysis

All analyses were conducted using STATA 14. Descriptive statistics are used to report the socio-demographic characteristics of the sample, split by country context and health condition sub-sample. The psychometric analyses were split by country to test whether the German translation of the ICECAP-A performs similarly to the original UK version of the tool. Analyses were further split by health condition group, to test whether the tool performed as well as expected across populations with a range of health conditions.

The internal consistency of ICECAP-A was tested using Cronbach’s alpha, and refers to the extent to which the items of an instrument are conceptually inter-related (Tavakol and Dennick 2011). Although ICECAP-A is expected to assess five distinct components

(stability, attachment, autonomy, achievement and enjoyment), it would be useful to examine the extent to which the tool measures some underlying and unifying concept of ‘capabilities’. Alpha scores range from ‘0 – 1’ and low scores indicate poor internal consistency, while scores of ‘.7’ or higher indicate satisfactory internal consistency (Bland and Altman 1997).

To assess convergent validity, Pearson’s correlation coefficients between scores for the ICECAP-A and related measures of health-related quality of life and wellbeing were calculated (EQ-5D-5L, SF-6D and SWLS). Correlations between the ICECAP-A and other measures are considered strong if over ‘0.5’, moderate if between ‘0.3’ and ‘0.5’ and weak if less than ‘0.3’ (Cohen 1988).

Construct validity describes a broad range of measurement concerns; however, overall it concerns investigations into the relationships between scores on a measure with other variables that it theoretically should be associated with (Westen and Rosenthal 2003). A central component of construct validation is ongoing hypothesis testing concerning other questionnaire instruments or between sub-groupings of respondents (De Vet et al. 2011). In line with guidance, we investigated the ability of the ICECAP-A to differentiate between groups hypothesised to differ in their levels of the construct capability-wellbeing (the healthy sub-sample, compared to respondents reporting one of the seven primary healthy conditions).

Construct validity was assessed using a two-stage process. Firstly, ordinary least squares (OLS) regression methods were used to test whether type of health condition (reporting the presence of one of seven primary health conditions compared to being healthy) had a significant association with capability scores, controlling for age, gender and educational level. Although previous research has employed exploratory factor analysis to validate ICECAP measures compared to other health measures (Davis et al. 2013; Keeley et al. 2016; Engel et al. 2017), our research question required a method to compare the performance of two versions of ICECAP simultaneously. Previous research has demonstrated that OLS models perform well in health-related quality of life outcome studies (Al-Janabi et al. 2017), and studies with capability wellbeing measures (Mitchell et al. 2013; Mitchell et al. 2017b; Franklin et al. 2018). Secondly, the sample was split into the seven health condition groups, and follow-up OLS analyses were conducted to examine the associations between health condition severity and capability wellbeing for each of the health conditions, controlling for demographic differences (age, gender and education). This analysis was undertaken to explore the associations between health condition within each of the health condition groups, rather than to provide a directly comparable indicator of severity between conditions.

## Results

### Sample Characteristics

Missing data were found for a very small proportion of the respondents (0.08%, 2/2503), therefore data were only analysed for complete cases ( $N=2501$ ). The descriptive characteristics of this sample are presented in Table 1. The German sample ( $N=1212$ ) and UK sample ( $N=1289$ ) are further split into healthy and health condition group sub-samples. The youngest respondents were found in the depression sub-sample and the oldest respondents were found in the heart disease sub-sample across the German and

**Table 1** Sample descriptive characteristics – German sample and UK sample

Sample <i>N</i> = sample size	Sub-sample <i>n</i> = sub-sample	Age Mean (SD)	Gender % female	Education % University	ICECAP-A Mean (SD)
Germany <i>N</i> = 1212	Healthy <i>n</i> = 203	44.90 (15.20)	51.72	14.29	.883 (.114)
	Arthritis <i>n</i> = 159	52.40 (9.54)	53.16	18.35	.766 (.175)
	Asthma <i>n</i> = 147	42.95 (15.80)	60.00	41.33	.812 (.171)
	Cancer <i>n</i> = 115	53.38 (13.24)	56.93	38.69	.792 (.185)
	Depression <i>n</i> = 160	42.16 (12.62)	63.39	34.81	.595 (.224)
	Diabetes <i>n</i> = 140	52.86 (12.78)	32.30	29.81	.766 (.211)
	Hearing problems <i>n</i> = 136	51.31 (11.86)	47.62	36.51	.830 (.169)
	Heart disease <i>n</i> = 152	55.38 (11.50)	31.74	25.75	.789 (.197)
	UK <i>N</i> = 1289	Healthy <i>n</i> = 234	45.92 (16.05)	49.15	30.34
Arthritis <i>n</i> = 158		58.27 (11.25)	52.83	23.27	.799 (.153)
Asthma <i>n</i> = 150		41.31 (15.80)	48.98	25.85	.846 (.134)
Cancer <i>n</i> = 137		59.91 (11.42)	45.22	38.26	.785 (.184)
Depression <i>n</i> = 158		38.94 (13.17)	60.63	16.25	.613 (.205)
Diabetes <i>n</i> = 161		57.25 (12.29)	36.43	32.14	.817 (.181)
Hearing problems <i>n</i> = 124		55.27 (17.43)	47.58	37.10	.831 (.169)
Heart disease <i>n</i> = 167		59.17 (14.02)	34.21	32.24	.795 (.185)

*SD* Standard deviation

UK samples. Similarly, the depression sub-sample contained the most female respondents while the heart disease sub-sample contained the most males in both the German and UK samples. The cancer sub-sample had the highest level of university attendance in the German sample, while people with Asthma had the highest educational level in the UK data. Finally, across Germany and the UK, ICECAP-A scores were highest amongst healthy respondents and lowest amongst respondents with depression.

The characteristics of respondents who scored at different levels of the five ICECAP-A attributes are described in Table 2. In the German sample, there were comparatively higher scores for achievement, while in the UK sample, over 50% of the sample indicated that they experienced the highest level of autonomy. Levels of life

**Table 2** Percentage of respondents at each ICECAP-A response level, and their mean SWLS and EQ-5D-5 L scores

Attributes and levels	Respondents at each level of the ICECAP-A - %		SWLS scores - Mean (SD)		EQ-5D scores - Mean (SD)	
	Germany	UK	Germany	UK	Germany	UK
<b>Stability</b>						
Level 4	29%	19%	.707 (.010)	0.737 (.011)	.909 (.007)	.904 (.008)
Level 3	52%	48%	.561 (.008)	0.583 (.008)	.811 (.007)	.832 (.007)
Level 2	17%	25%	.299 (.013)	0.344 (.011)	.614 (.016)	.658 (.014)
Level 1	2%	7%	.248 (.036)	0.139 (.016)	.502 (.039)	.528 (.031)
<b>Attachment</b>						
Level 4	36%	43%	.680 (.009)	0.637 (.009)	.864 (.008)	.837 (.009)
Level 3	42%	37%	.554 (.009)	0.511 (.010)	.812 (.008)	.779 (.010)
Level 2	20%	17%	.349 (.013)	0.288 (.014)	.676 (.015)	.669 (.017)
Level 1	2%	2%	.222 (.036)	0.167 (.040)	.583 (.039)	.512 (.050)
<b>Autonomy</b>						
Level 4	37%	53%	.639 (.010)	0.572 (.010)	.882 (.007)	.883 (.004)
Level 3	46%	34%	.554 (.009)	0.496 (.012)	.801 (.008)	.755 (.010)
Level 2	15%	11%	.357 (.016)	0.385 (.020)	.627 (.018)	.460 (.023)
Level 1	2%	2%	.302 (.050)	0.342 (.058)	.454 (.066)	.245 (.074)
<b>Achievement</b>						
Level 4	27%	21%	.696 (.011)	0.702 (.012)	.919 (.006)	.915 (.006)
Level 3	49%	47%	.571 (.009)	0.577 (.008)	.828 (.007)	.844 (.006)
Level 2	22%	26%	.375 (.013)	0.349 (.012)	.626 (.014)	.638 (.014)
Level 1	3%	6%	.213 (.029)	0.209 (.023)	.499 (.045)	.441 (.033)
<b>Enjoyment</b>						
Level 4	26%	33%	.692 (.011)	0.689 (.009)	.903 (.007)	.895 (.006)
Level 3	49%	39%	.586 (.009)	0.540 (.009)	.838 (.006)	.807 (.008)
Level 2	23%	26%	.354 (.012)	0.302 (.011)	.629 (.013)	.615 (.015)
Level 1	2%	2%	.169 (.030)	0.098 (.036)	.464 (.045)	.450 (.064)

Attributes refer to each of the domains within the ICECAP-A

The four levels within the attributes range from having full capability (4) to having no capability (1)

satisfaction (SWLS) and health-related quality of life (EQ-5D-5 L) were also examined for participants scoring at different levels of the ICECAP-A attributes. In Germany and the UK, the highest mean SWLS scores were found for respondents with the highest levels of 'Stability'. Similarly, in both samples, the highest mean EQ-5D scores were found in respondents with the highest levels of 'Achievement'.

### Internal Consistency Reliability

Cronbach's alpha was calculated separately for the German and UK samples and is presented in Table 3. The German translation of the ICECAP-A demonstrated good



**Table 3** Internal consistency (Cronbach's alpha -  $\alpha$ ) of the ICECAP-A across sub-samples

	Germany	UK
Overall sample	.83	.85
Healthy	.78	.80
Arthritis	.74	.78
Asthma	.77	.83
Cancer	.86	.83
Depression	.78	.79
Diabetes	.83	.86
Hearing loss	.74	.84
Heart disease	.83	.85

internal consistency ( $\alpha = .83$ ) and was comparable with the internal consistency of the scale in the UK data ( $\alpha = .85$ ). When internal consistency was assessed separately for each health condition group the scale was found to be similarly reliable (Germany:  $\alpha = .74$ –.86, UK:  $\alpha = .78$ –.86).

### Convergent Validity

Correlation coefficients highlighting the relationships between the ICECAP-A and three related tools are presented in Table 4. All of the correlation coefficients were

**Table 4** Correlation coefficients between the ICECAP-A and three related measures across sub-samples

	Samples <sup>a</sup>	EQ-5D-5 L	SWLS	SF-6D
Germany	Overall sample	.62	.66	.64
	Healthy	.35	.61	.47
	Arthritis	.53	.58	.64
	Asthma	.61	.52	.56
	Cancer	.57	.70	.57
	Depression	.48	.54	.56
	Diabetes	.55	.62	.57
	Hearing loss	.58	.60	.57
	Heart disease	.66	.66	.70
UK	Overall sample	.61	.68	.65
	Healthy	.36	.63	.40
	Arthritis	.64	.57	.66
	Asthma	.43	.65	.54
	Cancer	.68	.67	.70
	Depression	.57	.64	.59
	Diabetes	.60	.67	.60
	Hearing loss	.58	.77	.61
	Heart disease	.68	.66	.73

(All coefficients were statistically significant ( $p < .001$ ))

<sup>a</sup> Samples sizes are stated in Table 1

statistically significant ( $p < .001$ ) and positive in direction. In the German sample, all of the correlation coefficients were moderate to large in size ( $r \geq .3$ ). These results demonstrate that scores on the German translation of the ICECAP-A converge with scores on measures of health-related quality of life and wellbeing, and that this convergence is generalizable across sub-populations. These results were comparable to the results yielded with the English language version of the ICECAP-A in the UK sample. One notable difference in results across study populations was that capabilities

**Table 5** Correlation coefficients between the ICECAP-A and dimension level scores on the EQ-5D-5 L, SWLS and SF-36, in Germany and the UK

Country context	Dimensions <sup>a</sup>	ICECAP-A attributes				
		Stability	Attachment	Autonomy	Achievement	Enjoyment
Germany	EQ-5D Mobility	0.28	0.19	0.33	0.40	0.31
	EQ-5D Self-care	0.29	0.25	0.33	0.33	0.31
	EQ-5D Usual activities	0.42	0.30	0.43	0.53	0.45
	EQ-5D Pain	0.31	0.19	0.27	0.37	0.35
	EQ-5D Anxiety/Depression	0.61	0.47	0.42	0.53	0.55
	SWLS <sup>b</sup>	0.57	0.53	0.39	0.52	0.54
	SF Physical functioning	0.33	0.24	0.36	0.45	0.36
	SF Role-Physical	0.39	0.27	0.34	0.46	0.38
	SF Bodily pain	0.37	0.22	0.30	0.41	0.39
	SF General health	0.44	0.32	0.36	0.50	0.45
	SF Vitality	0.57	0.41	0.45	0.60	0.53
	SF Social functioning	0.54	0.48	0.40	0.53	0.51
	SF Role-emotional	0.51	0.37	0.41	0.48	0.46
	SF Mental health	0.64	0.49	0.44	0.56	0.58
UK	EQ-5D Mobility	0.27	0.16	0.54	0.39	0.32
	EQ-5D Self-care	0.25	0.13	0.53	0.36	0.29
	EQ-5D Usual activities	0.33	0.20	0.60	0.47	0.39
	EQ-5D Pain	0.27	0.15	0.45	0.37	0.29
	EQ-5D Anxiety/Dep	0.65	0.48	0.39	0.57	0.60
	SWLS <sup>b</sup>	0.67	0.52	0.25	0.57	0.62
	SF Physical functioning	0.27	0.14	0.55	0.41	0.33
	SF Role-Physical	0.27	0.20	0.38	0.34	0.33
	SF Bodily pain	0.32	0.17	0.47	0.41	0.33
	SF General health	0.44	0.31	0.49	0.52	0.49
	SF Vitality	0.57	0.40	0.43	0.58	0.60
	SF Social functioning	0.54	0.38	0.53	0.54	0.56
	SF Role-emotional	0.46	0.33	0.38	0.42	0.45
	SF Mental health	0.69	0.52	0.39	0.60	0.67

(All coefficients were statistically significant ( $p < .001$ ))

<sup>a</sup> Higher levels of health related quality of life (EQ-5D-5L and SF-6D), and higher levels of life satisfaction (SWLS) <sup>a</sup> <sup>b</sup> correlations between the ICECAP-A and were not conducted with the SWLS at item level, as the five questions in this scale measure a single dimension (life satisfaction).

(ICECAP-A) were more related to life satisfaction (SWLS) than health-related quality of life (EQ-5D-5 L and SF-6D) among healthy respondents in the German and UK study samples.

Correlation coefficients between the ICECAP-A attributes and the dimensions within these three outcome measures for the overall samples in Germany and the UK are presented in Table 5. In both country contexts, SWLS scores had good correlation with most ICECAP-A attributes in both countries, except for the autonomy attribute. The anxiety/depression dimension on EQ-5D-5L had good correlation with all ICECAP-A attributes, with only the autonomy attribute on ICECAP-A having higher correlation with usual activities on EQ-5D-5L across both countries. For SF-36, the dimensions of vitality, social functioning and mental health also had good correlations across all ICECAP-A attributes in both countries.

**Table 6** OLS regressions examining the associations between each of the health conditions and capability scores

	OLS regression in German sample			OLS regression in UK sample		
	Coef. (SE)	Lower 95% CI	Upper 95% CI	Coef. (SE)	Lower 95% CI	Upper 95% CI
<b>Socio-Demographics</b>						
Age	.001* (.000)	.000	.002	.001** (.000)	.001	.002
Gender (Female)	-.003 (.010)	-.022	.015	.001 (.011)	-.020	.021
Education (University)	.054*** (.011)	.033	.075	.027* (.011)	.005	.048
<b>Health conditions</b>						
Arthritis	-.095*** (.017)	-.129	-.061	-.123*** (.019)	-.161	-.085
Asthma	-.041* (.017)	-.075	-.007	-.063** (.019)	-.100	-.025
Cancer	-.119*** (.019)	-.156	-.081	-.105*** (.020)	-.145	-.065
Depression	-.267*** (.017)	-.301	-.234	-.274*** (.019)	-.311	-.237
Diabetes	-.083*** (.018)	-.118	-.048	-.125*** (.019)	-.163	-.088
Hearing loss	-.065*** (.018)	-.100	-.029	-.060** (.021)	-.100	-.019
Heart disease	-.107*** (.018)	-.142	-.073	-.104*** (.019)	-.141	-.067
_cons	0.840			.811		
R <sup>2</sup>	.21			.18		
N	1212			1289		

*Dependent variable* ICECAP-A scores, *Coef.* Regression coefficient, *SE* Standard error, *CI* 95% Confidence interval, \* $p < .05$ , \*\* $p < .01$  and \*\*\* $p < .001$

## Construct Validity

The effects of experiencing one of the seven health conditions (compared to being in the healthy sub-sample) on capability wellbeing are presented in Table 6. The coefficients from the OLS regression indicate that in the German sample, the presence of any of the seven of the health conditions had a significantly negative impact on capability scores, controlling for socio-demographic characteristics. Depression ( $\beta = -.267, p < .001$ ), followed by cancer ( $\beta = -.119, p < .001$ ) and heart disease ( $\beta = -.107, p < .001$ ) had the strongest effects on capabilities, while the smallest effect was for the presence of asthma ( $\beta = -.041, p < .05$ ). The seven health conditions were also associated with significantly lower levels of capability wellbeing in the UK sample, in which depression also had the strongest detrimental effect ( $\beta = -.274, p < .001$ ). The  $R^2$  statistics indicate that the

**Table 7** Summary table of regression coefficients indicating the impact of condition-specific severity on capabilities, controlling for socio-demographics (age, gender and education) across health condition groups in Germany and the UK

Country	Condition severity <i>n</i> = sub-sample	Coef. (SE)	Lower 95% CI	Upper 95% CI
Germany	Arthritis <i>n</i> = 159	-0.556*** (0.075)	-0.704	-0.408
	Asthma <i>n</i> = 147	-0.305*** (0.048)	-0.400	-0.210
	Cancer <i>n</i> = 115	-0.409*** (0.062)	-0.532	-0.286
	Depression <i>n</i> = 160	-0.499*** (0.067)	-0.631	-0.366
	Diabetes <i>n</i> = 140	-0.396*** (0.062)	-0.519	-0.272
	Hearing problems <i>n</i> = 136	-0.170* (0.067)	-0.303	-0.036
	Heart disease <i>n</i> = 152	-0.673*** (0.053)	-0.778	-0.567
United Kingdom	Arthritis <i>n</i> = 158	-0.593*** (0.086)	-0.759	-0.481
	Asthma <i>n</i> = 150	-0.620*** (0.070)	-0.764	-0.423
	Cancer <i>n</i> = 137	-0.502*** (0.052)	-0.606	-0.399
	Depression <i>n</i> = 158	-0.628*** (0.063)	-0.752	-0.503
	Diabetes <i>n</i> = 161	-0.429*** (0.069)	-0.566	-0.292
	Hearing problems <i>n</i> = 124	-0.313*** (0.076)	-0.464	-0.163
	Heart disease <i>n</i> = 167	-0.747*** (0.053)	-0.852	-0.641

Dependent variable ICECAP-A scores, *Coef.* Regression coefficient, *SE* Standard error, *CI* 95% Confidence interval, \* $p < .05$ , \*\* $p < .01$  and \*\*\* $p < .001$

demographics and health conditions accounted for approximately a fifth of the variance in ICECAP-A scores (Germany = 21% and UK = 18%).

Finally, the sample was split by health condition group to determine the extent to which *condition-specific severity* had a negative association with ICECAP-A scores *within the seven health condition groups*. The OLS regression coefficients for health condition severity in each of the separate health condition groups are presented in Table 7 (Appendix 1 and 2: Full description of the individual regression analyses). In each of the German health condition groups, health condition severity had a significant negative impact on capability wellbeing, most notably for heart disease ( $\beta = -.673$ ,  $p < .001$ ), arthritis ( $\beta = -.556$ ,  $p < .001$ ) and depression ( $\beta = -.499$ ,  $p < .001$ ). A similar pattern of results was observed for the UK health condition groups.

## Discussion

### Summary of Main Findings

This study set out to determine whether the German translation of the ICECAP-A is psychometrically valid, and capable of yielding results comparable to the previously validated original UK version of the tool. The German ICECAP-A demonstrated good internal consistency and good convergent validity when compared to existing measures of health-related quality of life and satisfaction with life. Further, scores on the newly translated version of the tool were significantly lower amongst respondents reporting one of the seven studied health condition groups (compared to being healthy). Finally, the results yielded by the German translation of the ICECAP-A used in the German sample were largely comparable to the results yielded by the original English language version of the tool used in the UK.

### Interpretation of Findings

The findings from this study contribute to the growing literature concerning the psychometric validity of ICECAP measures across cultural contexts. The ICECAP-A also demonstrated convergent validity with the SWLS and SF-6D, suggesting that it is able to capture elements of a person's health-related quality of life, but also wider aspects of subjective wellbeing. This work also extends existing cross-cultural (Australia, Canada, UK and US) insight on the detrimental impact of depression on capabilities to a German context (Mitchell et al. 2015). Meeting the challenges caused by mental health difficulties in Europe has become a widely recognised policy priority (World Health Organization 2008).

### Strengths and Limitations

The key strength of this study is its utilisation of data collected from participants reporting a variety of health conditions, from multiple country contexts. It represents the first study to validate a translated ICECAP measure across countries compared to the original version. Further, evidence is presented for multiple psychometric properties, indicating a comprehensive description of both validity and reliability. There are, however also some limitations. In particular, the pragmatic nature of the study meant that formal back-translation of the measure was not undertaken until after the data collection was conducted. A further caveat is

that due to the use of cross-sectional data, further research is needed to establish the causality of the significant relationships identified. It should also be noted that there were some socio-demographic differences in the participants between the samples in the German and UK data. Reliance on a UK value set to demonstrate the validity of the German translation is also not ideal, but necessary given no value set for ICECAP-A currently exists outside the country of origin (Flynn et al. 2015). In addition, there is no “gold-standard” for the measurement of capability wellbeing, with the ICECAP-A representing one of the first attempts to directly measure self-reported capability (Mitchell et al. 2017a). Validity assessment in this study relies on common measures of health status currently used in health economics and a measure of subjective wellbeing. Finally, condition severity is measured using condition-specific measures in each health condition group, therefore the follow-up OLS results highlight the importance of condition severity within conditions, *rather* than differences in the importance of severity between conditions.

## Implications

This study demonstrates that the German translation of the ICECAP-A performs similarly to the original English version of the tool and is able to collect reliable and valid data when applied in a German setting. This study therefore provides initial evidence that the ICECAP-A capability measure may be appropriate to be used in countries outside where the measure was developed. Specifically, this paper provides evidence for German decision makers to recommend the use of the ICECAP-A in economic evaluations, as has been previously done for the ICECAP-O in the Netherlands for evaluations of long-term conditions (Versteegh et al. 2016).

## Future Research

Although this study presents evidence of the validity of the German translated ICECAP-A, there are a number of questions that remain unaddressed in the use of translated capability measures. Two main areas that need to be addressed concern the appropriateness of the ICECAP-A descriptive and valuation system outside of the UK. This question is not only aimed at countries where English is not the first language, as both the specific capabilities that are important and how they are valued could vary in both English and non-English speaking settings.

In terms of further validation of the German translated ICECAP-A, ongoing research is needed to establish whether the German translation of the ICECAP-A performs well when tested for additional psychometric properties. For example, a longitudinal study design would enable researchers to explore whether scores are stable across reasonable durations of time (test re-test reliability) and whether scores are sensitive to changes in life circumstances and health status (responsiveness).

## Conclusion

This is the first paper to provide psychometric evidence that simultaneously compares a translated ICECAP instrument in its country of use (Germany) with the original English language version of the instrument in the UK. Although the current study is based on cross-

sectional data, the results indicate that the German translation of the ICECAP-A yields valid and reliable data and is comparable to the original English language version of the tool. This study adds to growing body of literature on the psychometric properties of the ICECAP-A, whilst also extending our understanding of the relationship between condition severity and capability wellbeing.

**Acknowledgements** We would like to thank Dr. Harald Hofmann from the University of Queensland, Australia, for providing the ICECAP-A German translation. Thanks also to Erica Niebauer and Jana Mader from the Division of Health Economics at the DKFZ, in providing support for the back translation report. We would also like to thank those in attendance at the European Health Economics Association (EuHEA) Conference, in Hamburg, Germany, July 2016, for comments on an earlier version of this research.

**Funding** Data collection was funded by the Australian National Health and Medical Research Council [grant number 1006334].

This research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West). The views expressed in this article are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

**Data Availability** The data used in this study were part of a larger Multi Instrument Comparison survey undertaken by the Centre for Health Economics at Monash University. It is available without charge and universally accessible online via the AqoL website ([www.aqol.com.au](http://www.aqol.com.au)). From the homepage there are instructions for accessing the data.

## Compliance with Ethical Standards

**Ethical Approval and Consent to Participate** Ethics approval was obtained from Monash University Human Research Ethics Committee (MUHREC Approval CF 11/1758:2011 00074). At the start of the survey, a Participant Information and Consent form was provided. Proceeding with the survey was deemed as consent.

**Conflict of Interest** HA and JC developed the ICECAP-A measure. JR and AI created the MIC database. ML, PM, MS and JU have no competing interests to declare.

## Appendix 1

OLS regression analyses investigating the impact of disease severity on capabilities (using the German translation of the ICECAP-A in Germany) in six patient populations

**Table 8** OLS regression analysis investigating the association between Arthritis Impact Measurement scores on capability wellbeing (using the German translation of the ICECAP-A in Germany) in the arthritis disease-group

	Coef.	Std. Err.	t	<i>P</i> >t	[95% Conf.	Interval]
Age	0.002	0.001	1.420	0.157	−0.001	0.004
Gender	0.012	0.021	0.570	0.567	−0.030	0.054
Education (University)	0.028	0.025	1.120	0.262	−0.021	0.077
Arthritis severity	−0.556	0.075	7.430	0.000	−0.704	−0.408
_cons	0.322	0.083	3.870	0.000	0.157	0.486

R-squared = 0.2909

**Table 9** OLS regression analysis investigating the association between Asthma Quality of Life Questionnaire scores on capability wellbeing (using the German translation of the ICECAP-A in Germany) in the asthma disease-group

	Coef.	Std. Err.	t	P	[95% Conf. Interval]
Age	0.000	0.001	0.08	0.940	-0.001 0.001
Gender	-0.015	0.020	-0.76	0.450	-0.054 0.024
Education (University)	-0.007	0.023	-0.31	0.759	-0.054 0.039
Asthma severity	-0.305	0.048	6.35	0.000	-0.400 -0.210
_cons	0.686	0.059	11.72	0.000	0.571 0.802

R-squared = 0.2506

**Table 10** OLS regression analysis investigating the association between European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Cancer-30 scores on capability wellbeing (using the German translation of the ICECAP-A in Germany) in the cancer disease-group

	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]
Age	0.000	0.001	-0.080	0.933	-0.002 0.002
Gender	0.016	0.030	0.520	0.603	-0.045 0.076
Education (University)	0.071	0.030	2.400	0.018	0.012 0.129
Cancer severity	-0.409	0.062	6.590	0.000	-0.532 -0.286
_cons	0.501	0.088	5.700	0.000	0.327 0.675

R-squared = 0.3467

**Table 11** OLS regression analysis investigating the association between Depression Anxiety and Stress Scale scores on capability wellbeing (using the German translation of the ICECAP-A in Germany) in the depression disease-group

	Coef.	Std. Err.	t	P>t	[95% Conf. Interval]
Age	0.002	0.001	1.680	0.095	0.000 0.004
Gender	0.017	0.029	0.570	0.571	-0.041 0.074
Education (University)	0.028	0.038	0.720	0.470	-0.048 0.103
Depression severity	-0.499	0.067	7.440	0.000	-0.631 -0.366
_cons	0.259	0.083	3.120	0.002	0.095 0.423

R-squared = 0.2838



**Table 12** OLS regression analysis investigating the association between Diabetes-39 Questionnaire scores on capability wellbeing (using the German translation of the ICECAP-A in Germany) in the diabetes disease-group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.000	0.001	0.030	0.973	-0.002 0.002
Gender	-0.014	0.028	-0.490	0.627	-0.069 0.042
Education (University)	0.029	0.030	0.970	0.336	-0.030 0.087
Diabetes severity	-0.396	0.062	6.360	0.000	-0.519 -0.272
_cons	0.562	0.082	6.840	0.000	0.400 0.725

R-squared = 0.2615

**Table 13** OLS regression analysis investigating the association between Abbreviated Profile of Hearing Aid Benefit scores on capability wellbeing (using the German translation of the ICECAP-A in Germany) in the hearing loss disease-group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.000	0.001	-0.480	0.633	-0.003 0.002
Gender	-0.011	0.025	-0.440	0.659	-0.059 0.038
Education (University)	0.014	0.024	0.570	0.566	-0.034 0.062
Hearing loss severity	-0.170	0.067	-2.520	0.013	-0.303 -0.036
_cons	0.925	0.070	13.200	0.000	0.786 1.063

R-squared = 0.0593

**Table 14** OLS regression analysis investigating the association between MacNew Heart Disease HRQL questionnaire scores on capability wellbeing (using the German translation of the ICECAP-A in Germany) in the heart disease group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.001	0.001	0.810	0.421	-0.001 0.003
Gender	0.043	0.022	1.920	0.057	-0.001 0.087
Education (University)	0.028	0.023	1.190	0.236	-0.018 0.073
Heart disease severity	-0.673	0.053	12.590	0.000	-0.778 -0.567
_cons	0.234	0.069	3.390	0.001	0.098 0.370

R-squared = 0.5558

## Appendix 2

OLS regression analyses investigating the impact of disease severity on capabilities (using the ICECAP-A in the UK) in six patient populations

**Table 15** OLS regression analysis investigating the association between Arthritis Impact Measurement scores on capability wellbeing (using the ICECAP-A in the UK) in the arthritis disease-group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.000	0.001	0.420	0.674	-0.001 0.002
Gender	0.057	0.025	2.240	0.026	0.007 0.107
Education (University)	0.033	0.025	1.310	0.192	-0.017 0.082
Arthritis severity	-0.593	0.086	6.880	0.000	-0.764 -0.423
_cons	0.171	0.107	1.600	0.112	-0.040 0.382

R-squared = 0.2627

**Table 16** OLS regression analysis investigating the association between Asthma Quality of Life Questionnaire scores on capability wellbeing (using the ICECAP-A in the UK) in the asthma disease-group

	Coef.	Std. Err.	t	<i>P</i>	[95% Conf. Interval]
Age	0.000	0.001	0.420	0.674	-0.001 0.002
Gender	0.057	0.025	2.240	0.026	0.007 0.107
Education (University)	0.033	0.025	1.310	0.192	-0.017 0.082
Asthma severity	-0.593	0.086	6.880	0.000	-0.764 -0.423
_cons	0.171	0.107	1.600	0.112	-0.040 0.382
	0.000	0.001	0.420	0.674	-0.001 0.002

R-squared = 0.2627

**Table 17** OLS regression analysis investigating the association between European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire Cancer-30 scores on capability wellbeing (using the ICECAP-A in the UK) in the cancer disease-group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.002	0.001	1.990	0.049	0.000 0.004
Gender	0.003	0.025	0.130	0.899	-0.046 0.052
Education (University)	0.004	0.025	0.160	0.873	-0.045 0.053
Cancer severity	-0.502	0.052	9.610	0.000	-0.606 -0.399
_cons	0.316	0.090	3.520	0.001	0.138 0.494

R-squared = 0.4533

**Table 18** OLS regression analysis investigating the association between Depression Anxiety and Stress Scale scores on capability wellbeing (using the ICECAP-A in the UK) in the depression disease-group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	-0.001	0.001	-0.910	0.363	-0.003 0.001
Gender	0.020	0.030	0.660	0.509	-0.039 0.078
Education (University)	0.025	0.030	0.840	0.402	-0.034 0.083
Depression severity	-0.628	0.063	9.960	0.000	-0.752 -0.503
_cons	0.269	0.080	3.350	0.001	0.110 0.427

R-squared = 0.3993

**Table 19** OLS regression analysis investigating the association between Diabetes-39 Questionnaire scores on capability wellbeing (using the ICECAP-A in the UK) in the diabetes disease-group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.000	0.001	-0.360	0.716	-0.003 0.002
Gender	0.041	0.033	1.260	0.211	-0.024 0.107
Education (University)	0.004	0.033	0.130	0.901	-0.061 0.070
Diabetes severity	-0.429	0.069	6.180	0.000	-0.566 -0.292
_cons	0.445	0.101	4.410	0.000	0.246 0.644

R-squared = 0.2110

**Table 20** OLS regression analysis investigating the association between Abbreviated Profile of Hearing Aid Benefit scores on capability wellbeing (using the ICECAP-A in the UK) in the hearing loss disease-group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.002	0.001	1.760	0.081	0.000 0.003
Gender	-0.026	0.030	-0.880	0.382	-0.085 0.033
Education (University)	0.020	0.030	0.670	0.501	-0.039 0.080
Hearing loss severity	-0.313	0.076	-4.130	0.000	-0.464 -0.163
_cons	0.905	0.087	10.450	0.000	0.733 1.076

R-squared = 0.1996

**Table 21** OLS regression analysis investigating the association between MacNew Heart Disease HRQL questionnaire scores on capability wellbeing (using the ICECAP-A in the UK) in the heart disease group

	Coef.	Std. Err.	t	<i>P</i> > t	[95% Conf. Interval]
Age	0.000	0.001	-0.210	0.838	-0.002 0.001
Gender	0.000	0.023	0.000	0.997	-0.045 0.045
Education (University)	-0.005	0.024	-0.200	0.843	-0.053 0.043
Heart disease severity	-0.747	0.053	14.020	0.000	-0.852 -0.641
_cons	0.295	0.069	4.290	0.000	0.159 0.431

R-squared = 0.5507

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## References

- Aaronson, N. K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N. J., et al. (1993). The European organization for research and treatment of cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *JNCI: Journal of the National Cancer Institute*, *85*(5), 365–376.
- Al-Janabi, H., Flynn, T. N., & Coast, J. (2012). Development of a self-report measure of capability wellbeing for adults: The ICECAP-A. *Quality of Life Research*, *21*(1), 167–176.
- Al-Janabi, H., Peters, T. J., Brazier, J., Bryan, S., Flynn, T. N., Clemens, S., et al. (2013). An investigation of the construct validity of the ICECAP-A capability measure. *Quality of Life Research*, *22*(7), 1831–1840.
- Al-Janabi, H., Manca, A., & Coast, J. (2017). Predicting carer health effects for use in economic evaluation. *PLoS One*, *12*(9), e0184886.
- Bland, J. M., & Altman, D. G. (1997). Statistics notes: Cronbach's alpha. *Bmj*, *314*(7080), 572.
- Boyer, J. G., & Earp, J. A. L. (1997). The development of an instrument for assessing the quality of life of people with diabetes: Diabetes-39. *Medical Care*, *35*(5), 440–453.
- Brazier, J., & Tsuchiya, A. (2015). Improving cross-sector comparisons: Going beyond the health-related QALY. *Applied Health Economics and Health Policy*, *13*(6), 557–565.
- Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, *21*(2), 271–292.
- Bullinger, M. (1995). German translation and psychometric testing of the SF-36 health survey: Preliminary results from the IQOLA project. *Social Science & Medicine*, *41*(10), 1359–1366.
- Coast, J., Flynn, T. N., Natarajan, L., Sproston, K., Lewis, J., Louviere, J. J., et al. (2008a). Valuing the ICECAP capability index for older people. *Social Science & Medicine*, *67*(5), 874–882.
- Coast, J., Smith, R., & Lorgelly, P. (2008b). Should the capability approach be applied in health economics? *Health Economics*, *17*(6), 667–670.
- Coast, J., Kinghorn, P., & Mitchell, P. (2015). The development of capability measures in health economics: Opportunities, challenges and progress. *The Patient-Patient-Centered Outcomes Research*, *8*(2), 119–126.
- Cohen, J. (1988). Set correlation and contingency tables. *Applied Psychological Measurement*, *12*(4), 425–434.
- Cox, R. M., & Alexander, G. C. (1995). The abbreviated profile of hearing aid benefit. *Ear and Hearing*, *16*(2), 176–186.
- Davis, J. C., Liu-Ambrose, T., Richardson, C. G., & Bryan, S. (2013). A comparison of the ICECAP-O with EQ-5D in a falls prevention clinical setting: Are they complements or substitutes? *Quality of Life Research*, *22*(5), 969–977. <https://doi.org/10.1007/s11136-012-0225-4>.
- De Vet, H. C., Terwee, C. B., Mokkink, L. B., & Knol, D. L. (2011). *Measurement in medicine: a practical guide*. Cambridge: Cambridge University Press.
- Devlin, N., Shah, K., Feng, Y., Mulhern, B., & van Hout, B. (2017). Valuing health-related quality of life: An EQ-5D-5L value set for England. *Health Economics*. <https://doi.org/10.1002/hec.3564>.
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., et al. (2007). A psychometric toolbox for testing validity and reliability. *Journal of Nursing Scholarship*, *39*(2), 155–164.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75.
- Engel, L., Mortimer, D., Bryan, S., Lear, S. A., & Whitehurst, D. G. T. (2017). An investigation of the overlap between the ICECAP-A and five preference-based health-related quality of life instruments. *PharmacoEconomics*, *35*(7), 741–753. <https://doi.org/10.1007/s40273-017-0491-7>.
- Flynn, T. N., Huynh, E., Peters, T. J., Al-Janabi, H., Clemens, S., Moody, A., et al. (2015). Scoring the Icecap-a capability instrument. Estimation of a UK general population tariff. *Health Economics*, *24*(3), 258–269.

- Franklin, M., Payne, K., & Elliott, R. A. (2018). Quantifying the relationship between capability and health in older people: Can't map, Won't map. *Medical Decision Making*, 38(1), 79–94. <https://doi.org/10.1177/0272989X17732975>.
- Goranitis, I., Coast, J., Al-Janabi, H., Latthe, P., & Roberts, T. E. (2016). The validity and responsiveness of the ICECAP-A capability-well-being measure in women with irritative lower urinary tract symptoms. *Quality of Life Research*, 25(8), 2063–2075.
- Grewal, I., Lewis, J., Flynn, T., Brown, J., Bond, J., & Coast, J. (2006). Developing attributes for a generic quality of life measure for older people: Preferences or capabilities? *Social Science & Medicine*, 62(8), 1891–1901.
- Guillemin, F., Coste, J., Pouchot, J., Ghézail, M., Bregeon, C., & Sany, J. (1997). The AIMS2-SF. A short form of the arthritis impact measurement scales 2. *Arthritis & Rheumatology*, 40(7), 1267–1274.
- Hicks, S., Tinkler, L., & Allin, P. (2013). Measuring subjective well-being and its potential role in policy: Perspectives from the UK office for national statistics. *Social Indicators Research*, 114(1), 73–86.
- Hinz, A., Kohlmann, T., Stöbel-Richter, Y., Zenger, M., & Brähler, E. (2014). The quality of life questionnaire EQ-5D-5L: Psychometric properties and normative values for the general German population. *Quality of Life Research*, 23(2), 443–447.
- Höfer, S., Lim, L., Guyatt, G., & Oldridge, N. (2004). The MacNew heart disease health-related quality of life instrument: A summary. *Health and Quality of Life Outcomes*, 2(1), 3.
- Hörder, H., Gustafsson, S., Rydberg, T., Skoog, I., & Waern, M. (2016). A cross-cultural adaptation of the ICECAP-O: Test–retest reliability and item relevance in Swedish 70-year-olds. *Societies*, 6(4), 30.
- Keeley, T., Al-Janabi, H., Nicholls, E., Foster, N., Jowett, S., & Coast, J. (2015). A longitudinal assessment of the responsiveness of the ICECAP-A in a randomised controlled trial of a knee pain intervention. *Quality of Life Research*, 24(10), 2319–2331.
- Keeley, T., Coast, J., Nicholls, E., Foster, N. E., Jowett, S., & Al-Janabi, H. (2016). An analysis of the complementarity of ICECAP-A and EQ-5D-3 L in an adult population of patients with knee pain. *Health and Quality of Life Outcomes*, 14(1), 1–5. <https://doi.org/10.1186/s12955-016-0430-x>.
- Linton, M.-J., Dieppe, P., & Medina-Lara, A. (2016). Review of 99 self-report measures for assessing well-being in adults: Exploring dimensions of well-being and developments over time. *BMJ Open*, 6(7), e010641.
- Lovibond, S., & Lovibond, P. (1995). *Manual for the depression anxiety stress scales*. Sydney: Psychological Foundation of Australia. Inc..
- Makai, P., Koopmanschap, M. A., Brouwer, W. B., & Nieboer, A. A. (2013). A validation of the ICECAP-O in a population of post-hospitalized older people in the Netherlands. *Health and Quality of Life Outcomes*, 11(1), 57.
- Makai, P., Beckebans, F., van Exel, J., & Brouwer, W. B. (2014). Quality of life of nursing home residents with dementia: Validation of the German version of the ICECAP-O. *PLoS One*, 9(3), e92016.
- Marks, G. B., Dunn, S. M., & Woolcock, A. J. (1992). A scale for the measurement of quality of life in adults with asthma. *Journal of Clinical Epidemiology*, 45(5), 461–472.
- Mitchell, P. M., Roberts, T. E., Barton, P. M., Pollard, B. S., & Coast, J. (2013). Predicting the ICECAP-O capability index from the WOMAC osteoarthritis index: Is mapping onto capability from condition-specific health status questionnaires feasible? *Medical Decision Making*, 33(4), 547–557. <https://doi.org/10.1177/0272989x12475092>.
- Mitchell, P. M., Al-Janabi, H., Richardson, J., Iezzi, A., & Coast, J. (2015). The relative impacts of disease on health status and capability wellbeing: A multi-country study. *PLoS One*, 10(12), e0143590.
- Mitchell, P. M., Roberts, T. E., Barton, P. M., & Coast, J. (2017a). Applications of the capability approach in the health field: A literature review. *Social Indicators Research*, 133(1), 345–371. <https://doi.org/10.1007/s11205-016-1356-8>.
- Mitchell, P. M., Venkatapuram, S., Richardson, J., Iezzi, A., & Coast, J. (2017b). Are quality-adjusted life years a good proxy measure of individual capabilities? *Pharmacoeconomics*, 35(6), 637–646. <https://doi.org/10.1007/s40273-017-0495-3>.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., et al. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Quality of Life Research*, 19(4), 539–549.
- National Institute for Health and Care Excellence. (2014). *Developing NICE guidelines: the manual* (pp. 240). Manchester.
- Reeve, B. B., Wyrwich, K. W., Wu, A. W., Velikova, G., Terwee, C. B., Snyder, C. F., et al. (2013). ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Quality of Life Research*, 22(8), 1889–1905.

- Richardson, J., Iezzi, A., & Khan, M. A. (2015a). Why do multi-attribute utility instruments produce different utilities: The relative importance of the descriptive systems, scale and 'micro-utility' effects. *Quality of Life Research*, 24(8), 2045–2053.
- Richardson, J., Khan, M. A., Iezzi, A., & Maxwell, A. (2015b). 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. *Medical Decision Making*, 35(3), 276–291.
- Richardson, J., Iezzi, A., Khan, M. A., Chen, G., & Maxwell, A. (2016). Measuring the sensitivity and construct validity of 6 utility instruments in 7 disease areas. *Medical Decision Making*, 36(2), 147–159. <https://doi.org/10.1177/0272989X15613522>.
- Sarabia-Cobo, C. M., Parás-Bravo, P., Amo-Setién, F. J., Alconero-Camarero, A. R., Sáenz-Jalón, M., Torres-Manrique, B., et al. (2017). Validation of the Spanish version of the ICECAP-O for nursing home residents with dementia. *PLoS One*, 12(1), e0169354.
- Sen, A. (1993). *Capability and Well-Being*. In M. C. Nussbaum & A. Sen, *The Quality of Life* (pp. 30–53). Oxford: Oxford University Press.
- Stiglitz, J. E., Sen, A., & Fitoussi, J.-P. (2010). *Report by the commission on the measurement of economic performance and social progress*. Paris: Commission on the Measurement of Economic Performance and Social Progress.
- Taggart, F., Friede, T., Weich, S., Clarke, A., Johnson, M., & Stewart-Brown, S. (2013). Cross cultural evaluation of the Warwick-Edinburgh mental well-being scale (WEMWBS)-a mixed methods study. *Health and Quality of Life Outcomes*, 11(1), 27.
- Tang, C., Xiong, Y., Wu, H., & Xu, J. (2018). Adaptation and assessments of the Chinese version of the ICECAP-A measurement. *Health and Quality of Life Outcomes*, 16(1), 45. <https://doi.org/10.1186/s12955-018-0865-3>.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53.
- Versteegh, M., Knies, S., & Brouwer, W. (2016). From good to better: New Dutch guidelines for economic evaluations in healthcare. *Pharmacoeconomics* 34(11), 1071–1074.
- Ware Jr, J. E., & Sherbourne, C. D. (1992). The MOS 36-item short-form health survey (SF-36): I. Conceptual framework and item selection. *Medical Care*, 473–483.
- Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, 84(3), 608.
- Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., et al. (2005). Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: Report of the ISPOR task force for translation and cultural adaptation. *Value in Health*, 8(2), 94–104.
- Wisløff, T., Hagen, G., Hamidi, V., Movik, E., Klemp, M., & Olsen, J. A. (2014). Estimating QALY gains in applied studies: A review of cost-utility analyses published in 2010. *Pharmacoeconomics*, 32(4), 367–375.
- World Health Organization. (2008). *Integrating mental health into primary care: A global perspective*. Singapore: WHO Press.

## Affiliations

Myles-Jay Linton<sup>1,2</sup> · Paul Mark Mitchell<sup>1</sup> · Hareth Al-Janabi<sup>3</sup> · Michael Schlander<sup>4,5,6</sup> · Jeff Richardson<sup>7</sup> · Angelo Iezzi<sup>7</sup> · Jasper Ubels<sup>4,6</sup> · Joanna Coast<sup>1,2</sup>

Myles-Jay Linton  
mj.linton@bristol.ac.uk

Hareth Al-Janabi  
h.aljanabi@bham.ac.uk

Michael Schlander  
m.schlander@dkfz-heidelberg.de

Jeff Richardson  
jeffrey.richardson@monash.edu

Angelo Iezzi  
angelo.iezzi@monash.edu

Jasper Ubels  
j.ubels@dkfz-heidelberg.de

Joanna Coast  
jo.coast@bristol.ac.uk

- <sup>1</sup> Health Economics at Bristol (HEB), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
- <sup>2</sup> The National Institute for Health Research Collaboration for Leadership in Applied Health Research and Care West (NIHR CLAHRC West) at University Hospitals Bristol NHS Foundation Trust, Bristol, UK
- <sup>3</sup> Health Economics Unit, Institute of Applied Health Research, University of Birmingham, Birmingham, UK
- <sup>4</sup> Division of Health Economics, German Cancer Research Center (DKFZ), Heidelberg, Germany
- <sup>5</sup> Institute for Innovation & Valuation in Health Care (InnoValHC), Wiesbaden, Germany
- <sup>6</sup> Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany
- <sup>7</sup> Centre for Health Economics, Monash University, Melbourne, Australia