The psychometric performance of the EQ-HWB in caregivers of persons living with dementia.

Abstract

Authors

Cate Bailey, Melbourne Health Economics, University of Melbourne, Australia.

Tessa Peasgood, School of Medicine and Population Health, University of Sheffield, UK

Bernhard Michalowsky, German Center for Neurodegenerative Diseases (DZNE) Rostock/Greifswald; Patient-Reported Outcomes & Health Economics Research, Greifswald, Germany.

Lidia Engel, School of Public Health and Preventive Medicine, Monash University, Australia

Funding – Funding was obtained from the EuroQol group, project RA1566, to undertake the analyses. Data collection was funded by Dementia Australia Research Foundation as part of a post-doc Fellowship for Lidia Engel,

Introduction

The EQ Health and Wellbeing (EQ-HWB) is a new generic instrument developed to use in evaluating interventions across health, public health, and social care, enabling the inclusion of impacts on caregivers. Evaluations of the instrument's performance in a range of informal caregiver populations is now needed.

Providing informal care to a person living with dementia can significantly impact the health and wellbeing of informal carers. However, it is unclear to what extent the EQ-HWB can capture such impacts. Using data collected through an Australian study of caregiver quality-of-life, this project aimed to investigate the psychometric properties of the EQ-HWB in caregivers of persons living with dementia.

Method

An online survey was developed which included demographic questions, informal care-related questions, and quality-of-life measures suitable for caregivers (C-DEMQOL and CarerQol) as well as the 25-item EQ-HWB. The psychometric properties of the EQ-HWB and EQ-HWB-S (9-item short-form) were assessed in terms of acceptability (missing data), distribution properties (ceiling and floor effects), known groups (Cohen's *d* and eta-squared) and convergent validity (Spearman correlations), and an examination of dimensionality using exploratory factor analysis (EFA).

Results

There were no floor or ceiling effects for either EQ-HWB instrument. Caregivers had high levels of exhaustion (34.7% selected 'often' or 'most/all of the time') and loneliness (25.7%). In known group validity analysis, the EQ-HWB instruments performed well against the C-DEMQOL and CarerQol in terms of effect size; however, the C-DEMQOL had higher effect sizes for dementia severity, and if there was more than one caregiver. At the item level, 13 of the 25 items could discriminate respondents by caregiving time and 23 of 25 items by caregivers' general self-reported health. Most hypothesised correlations in convergent validity analysis were found to be above 0.3. EQ-HWB items loaded onto 6 of the 9 EFA factors identified.

Conclusion

Our results suggest that the EQ-HWB instruments performed well in this population at the item, sum-score, and EQ-HWB-S index-score levels. Some EQ-HWB items may need further investigation as the instrument moves from the experimental stage into a final form.

The psychometric performance of the EQ-HWB in caregivers of persons living with dementia.

Introduction

Caregivers who provide informal care are critical to the quality-of-life of people living with dementia, but these caregivers can experience significant health and quality-of-life impacts, including social isolation, poor physical health, financial distress, and high levels of anxiety and depression [1]. The costs of value of time associated with providing informal care are higher for the dementias, particularly Alzheimer's disease, than other comparative diseases [2]. Whilst informal care costs are sometimes considered in economic evaluation, the caregivers' health and wellbeing impacts tend to be ignored [5], and it is rare to include caregiver quality-of-life-years (QALYs) that were collected using caregiver-sensitive preference-based questionnaires. This can lead to less-than-optimal decisions on the allocation of scarce resources for interventions that target people living with dementia [6]. Such impacts are often described as 'spillover effects'[3], and are an important consideration in economic evaluations adopting societal or welfare perspectives, as is now recommended by NICE in the UK [4].

To measure caregiver outcomes, suitable quality-of-life measures are required. There are several caregiver-specific quality-of-life measures; however, most of these are non-preference based and cannot be used in economic evaluations (ie C-DEMQOL[7]). The EQ-5D [8] and SF-6D [9] are often used instead, but whilst these measures have been validated extensively in the health sector, they may not capture important aspects of the quality-of-life of caregivers.

To address this, the 'Extending the QALY Project' aimed to develop a broad generic measure of quality-of-life for use in economic evaluation that would be applicable across health, social care and public health sectors and capture the impact of interventions upon caregivers [10]. Developed through an international collaboration [11], the EQ Health and Wellbeing measure (EQ-HWB) is now at an experimental stage, and not yet released for use except in research [12]. The measure is 25-items long, with a shorter version (EQ-HWB-S) of 9-items, and includes the concepts of loneliness, cognition, exhaustion, control as well as standard HRQoL domains such as mental health, mobility and self-care [13]. A pilot UK value-set has been produced for the short form [14], valued on the scale in which 1 indicates full health and wellbeing and zero is anchored as equivalent to dead (but can go below 0 into negative values) necessary to calculate Quality Adjusted Life Years (QALYs).

Recent work validating the questionnaire in caregivers has focussed on populations of parents of children with health conditions [15] and in caregivers of children where families have experienced adverse life events [16]. Both these studies found that the EQ-HWB-S seemed to be valid and reliable in these populations. Another study compared both EQ-HWB instruments to the EQ-5D in the general Australian population, and found that the EQ-HWB instruments had greater sensitivity for participants with mental health symptoms compared to the EQ-5D-5L [17]. In informal caregivers in the United States, the psychometric properties of the two EQ-HWB instruments were compared to comparable instruments, with authors finding support for the construct validity of the EQ-HWB and EQ-HWB-S [18]. The most recent study compares the EQ-HWB to the EQ-5D-5L for patient, caregiver and general public samples in China; the authors found that the EQ-HWB had good acceptability and construct validity, and that the EQ-HWB had better known-groups validity when the caregiver sample was included [19]. Across the above studies, the EQ-HWB performed well in regards to item response distribution [15–17, 19], floor/ceiling effects [18, 19], convergent analysis using Spearman correlations [16–19], exploratory factor analysis [17, 19], known group validity analysis using *t*-tests and one-way ANOVAs [15, 16, 18, 19], regression analysis [15, 17], and responsiveness to change [16].

The EQ-HWB instruments are potentially useful for measuring the quality-of-life of caregivers of people living with dementia, for inclusion in economic evaluations. There are no studies, to date, on validating the EQ-HWB instruments in caregivers of people living with dementia, and it is therefore important to validate the scale in this population. The aim of this study is to examine the psychometric performance of the EQ-HWB and EQ-HWB-S in caregivers of persons living with dementia.

2.0 Methods

2.1 Study design

Data were derived from 202 respondents in the COCOON project: incorporating Carer Outcomes in COst-effectiveness analyses Of dementia iNterventions”. This survey was not designed specifically to test the EQ-HWB, but was suitable for this purpose. An online survey was used to collect information in the COCOON study that included the full EQ-HWB (25 items) alongside a range of demographic questions and other carer-related quality of life instruments. Although there is no specific method to calculate the sample size for psychometric validation studies, 200 cases is seen as potentially suitable for testing the reliability and validity of Patient-Reported Outcome Measures [20]. Samples of over 100 participants are perceived as ‘very good’ in COSMIN guidelines for reliability analyses, comparisons to other measurement instruments, and construct validity [21].

Carers of people living with dementia were recruited with help from Step Up for Dementia, Dementia Australia, and carer organizations (e.g., Carers Victoria) (approximately 50% of sample). Additionally, carers were recruited through an online research portal PureProfile. Participants were invited to participate, informed about the study, and asked to provide written informed consent. A $10 gift voucher was provided to all respondents who completed the online survey in acknowledgement of their time; PureProfile panel members were reimbursed according to their specific panel agreements. Ethics was obtained from Monash University Human Research Ethics Committee #35073.

2.2 Materials

Survey data included a range of demographic characteristics of the caregiver and person living with dementia (caregiver: age, gender, education level, employment status, marital status, household income, state/country; cared-for (person living with dementia): gender, living situation, relationship to caregiver, type of dementia), informal care-related questions (length of caregiving/support in months, time spent caregiving per week, activities supported, income support for caring, as derived from the Resource Utilisation in Dementia (RED) Questionnaire [24]). Sum-scores were calculated for the EQ-HWB (25-items) and the EQ-HWB-S (9-items) by summing the items. Response options were scored from 1 to 5 for the 5 response levels. For the EQ-HWB, sum-scores could range from 25 (representing no problems on any item) and the maximum score is 125 (representing the most severe problems on all items). Sum-score range for the EQ-HWB-S was 9 to 45. We applied preference-weights from the recently published pilot UK value-set to produce index-scores for the EQ-HWB-S [14]. The three positive EQ-HWB items (items 19-21) were reverse scored prior to analysis.

Two other quality-of-life measures suitable for carers were included, the C-DEMQOL [7], and the CarerQol-7D [22] (scored using the Australian tariff [23]), as well as 35 candidate items for inclusion in the development of a dementia-specific quality-of-life measure for informal carers. The C-DEMQOL was developed specifically to measure the quality-of-life of family caregivers of people living with dementia and to define the concept of quality-of-life in this population. It is a 30-item questionnaire with 6 items for each of the five domains (meeting personal needs, carer wellbeing, carer-patient relationship, confidence in the future, feeling supported). The CareQol-7D was developed to be suitable as a care-related quality-of-life measure for informal caregivers, specifically for use in economic evaluations. The instrument has seven items plus a visual analogue scale measuring happiness. The recall periods for the three quality-of-life instruments are 7 days for the EQ-HWB, 4-weeks for the C-DEMQOL and "at the moment" for the CarerQol. The COCOON study data were conducted from May to November in 2023.

2.3 Statistical analysis

All analyses were performed in STATA version 15. Prior to running the analyses, data were cleaned by removing inconsistent and fraudulent respondents. These included respondents who provided inconstant responses to their location, respondents who completed the survey outside of Australia, as well as bots that were detected. Response distributions were calculated using frequency and percentages across responses. Responses were not forced (participants were not required to complete the question to proceed), so acceptability could be measured by missing data. Floor and ceiling effects were measured by determining percentages of participants scoring the lowest or highest scores across all items, and deemed significant if over 15%, (i.e. a floor effect in the EQ-HWB-S would mean more than 15% of respondents having a score of 111111111) [25].

We assessed known groups validity of the EQ-HWB sum-scores and EQ-HWB-S preference-weighted scores across variables where we hypothesised there were likely to be differences in health and wellbeing for carers, through a head-to-head comparison with the CarerQol and the C-DEMQOL. We hypothesised that caregivers were more likely to have lower quality-of-life (higher EQ-HWB sum-scores/lower EQ-WHB-S preference weighted scores) if 1) the caregiver had a chronic health condition (caregivers were asked about 13 specific chronic health conditions; this variable was then coded to Yes/No chronic health condition), 2) caregivers reported higher general health (5-item general health scale recoded to 2 groups (poor health (poor/fair), good health (good, very good, excellent), or 3 groups: poor health (poor/fair), moderate health (good), or high general health other (very good, excellent)) [26], 3) carers undertook caring for longer hours (caregiver time was coded to two equal groups (5 hours or less/ 5 or more hours per day) or 3 equal groups (less than 2.8, 2.8-11, 11-18 hours per day; with hours capped at 18 hours per day) and/or or 4) more severe dementia (categorised as mild or early, moderate or middle-stage, and severe or late stage dementia) [27, 28]. We further hypothesised that sole caregivers would have lower quality-of-life than where there was more than one caregiver, on the basis of reduced work due to shared care. We used *t*-tests when comparing two groups with Cohen's *d* scores for effect size, and one-way ANOVAs when comparing 3 groups with eta-squared for effect size. We compared known groups for each of the 25 EQ-HWB items by two variables: caregiver time (5 hours or less compared to more than 5 hours per day, with the time split chosen pragmatically as a mid-point) and caregiver self-reported health (good, very good or excellent compared to fair or poor) using *t*-tests, to understand how each item performed against these two variables.

Generalized linear regression models (GLMs) were conducted to investigate whether the differences observed by caregiver time (up to five hours/more than 5 hours; model 1) and general health (good to excellent health/fair to poor health; model 2) held when we controlled for caregiver gender and age. This analysis was included to determine whether the known groups results were being impacted by effects such as gender and age. To address the skewed distribution of utility values we used a gamma distribution and log link selected (as per Lee et al., 2024) for the GLM models.

To investigate convergent validity, we compared raw scores from the EQ-HWB (25 items) to the five C-DEMQOL dimensions and the 8 CarerQol items (including the VAS), using Spearman correlations for ordinal data. We made *a priori* hypotheses (presented in bold in the tables) regarding which items we expected to be at least moderately correlated (above .3) reaching consensus on expected correlations between three researchers (CB, TP and LE). Correlation strength was defined as per Cohen 1992 [29], where a correlation of 0.1-0.29 is considered weak, 0.3-0.49 moderate, and =>0.5 strong.

An exploratory factor analysis was conducted to investigate the structure of the EQ-HWB and the extent to which the instrument's items overlapped with other caregiver instruments (C-DEMQOL and CarerQol). We assessed the suitability of the data for factor analysis using Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity. We initially performed the analysis without rotation using the principal axis factor method of extracting factors in order to investigate the data. To determine the number of factors, we investigated the Kaiser criterion (with Eigen values >1) and a visual scree plot. We then used the Promax method for oblique rotation, allowing for correlated factors, to aid in interpreting the extracted factors. Factor loadings below .3 were removed to aid the identification of items that were able to represent a factor.

Results

3.1 Sample description

Analyses included 202 respondents who were caregivers of people living with dementia, of which 127 (63.2%) were women, the mean age of the sample was 60.4 (SD 13.7) years. One hundred participants were recruited from dementia and carers organisations and 102 participants from PureProfile. Baseline characteristics are presented in Table 1. For the caregivers, almost half of the sample had a university degree, nearly half were employed, 71% were married, 49% lived in a two-person household, and 93% spoke English as a first language. Of the people living with dementia, 54%lived with their caregiver, 20% lived in a care home and 18% lived alone (8% "other"), and 80% had another chronic health condition.

Table 1

Baseline characteristics of caregivers and persons living with dementia

	Full sample # (%)	Present (missing)
Caregivers		
Gender, n(%)		201 (1)
Female	127 (63.18)	
Male	74 (36.82)	
Age, mean (SD, range)	60.35 (13.70, 18-90)	199 (3)
Education, n(%)		201 (1)
Year 10 or less	15 (7.46)	
Year 11/12	26 (12.94)	
Cert III/IV or Diploma	60 (29.85)	
University degree	100 (49.75)	
Employment, n(%)		202 (0)
Employed	98 (48.51)	
Unemployed	7 (3.47)	
Student	3 (1.49)	
Retired	79 (39.11)	
Housewife/husband	7 (3.47)	
Other	8 (3.96)	
Marital status, n(%)		200 (2)
Single	32 (16.00)	
Married/De Facto	142 (71.00)	
Separated/Divorced	19 (9.50)	
Widowed	7 (3.50)	
Household size, n(%)		202 (0)
1 person	32 (15.84)	
2 people	98 (48.51)	
3 people	28 (13.86)	
4 people	28 (13.86)	
5+ people	15 (7.44)	
English is first language, n(%)		200 (2)
Yes	185 (92.50)	

No	15 (7.50)	
Time spent in caregiving, mean (SD, range))	8.14 (6.69, 0-18)	202 (0)
Other chronic health condition, n(%)		203 (0)
Yes	169 (83.66)	
No	33 (16.34)	
Household with children, n(%)		199 (3)
Yes	35 (17.59)	
No	164 (82.41)	
General health, n(%)		
Excellent	13 (6.47)	
Very good	57 (28.36)	
Good	72 (35.82)	
Fair	49 (24.38)	
Poor	10 (4.98)	
Person living with dementia		
Gender, n(%)		200 (2)
Female	90 (45.00)	
Male	109 (54.50)	
Other	1 (0.50)	
Age, mean (SD, range)	78.57 (10.61, 43-100))	198 (4)
Lives with carer, n(%)		200 (2)
Yes	110 (55.00)	
No	90 (45.00)	
Relationship to person living with dementia, n(%)		199 (3)
My partner	64 (32.16)	
My father or mother	82 (41.21)	
My son or daughter	2 (1.01)	
My grandparent	10 (5.03)	
Another family member	20 (10.05)	
My neighbour	1 (0.50)	
My friend	11 (5.53)	
Other (please specify):	9 (4.52)	
Diagnosis, n(%)		200 (2)
Alzheimer's Disease	91 (45.50)	
Vascular dementia	28 (14.00)	
Lewy Body Disease	10 (5.00)	
Frontotemporal dementia	13 (6.50)	
Don't know	39 (19.50)	
Other (please specify)	19 (9.50)	
Severity, n(%)		200 (2)
Mild or early-dementia	67 (33.50)	
Moderate or middle-stage dementia	101 (50.50)	
Severe or late-stage dementia	32 (16.00)	
Other chronic health condition, n(%)		202 (2)
Yes	163 (80.69)	
No	39 (19.31)	
Living situation, n(%)		200 (2)
Living alone	37 (18.50)	
Living with family	107 (53.50)	
Living in a care home	40 (20.00)	
Other (please specify)	16 (8.00)	

3.2 Response distribution and feasibility

Figure 1 shows the response distribution of the 25 EQ-HWB items for caregivers, the full results with means and standard deviations in Table 2. There were low levels for the "physical" items for seeing, hearing, getting around, daily activities, and self-care as well as item 13, "feel unsafe". There were high levels of exhaustion (35% selected 'often' or 'most/ all of the time') and sleep problems (27% selected 'often' or 'most/all of the time'). For the three positively worded questions (items 19-21), most people selected the three highest options (66%, 75% and 82% respectively) suggesting that most people felt accepted by others, felt good about themselves and were able to do the things they wanted to do at least some of the time. Pain and discomfort levels had a good spread over the response options; with pain frequency having the highest amount in the most severe category (6.4%) but also the most missing data (11.4%).

There were no missing data in items 1 to 20, suggesting good feasibility. There were 23 missing responses to item 22; in discussion between the authors, we wondered if this was due to the layout and that the item may have been overlooked by participants who did not need to answer questions to proceed with the survey (See figure S1). There were 2 missing data points 23 and 24, and one for item 25.

3.3 Known group validity

3.3.1 Known group validity for EQ-HWB-S index scores, EQ-HWB sum-scores, C-DEMQOL sum-score and CarerQol Australian preference-weighted scores

For the EQ-HWB and the EQ-HWB-S index scores, caregiver chronic health condition (2 groups), carer time (2 groups) and caregiver general health (2 groups) were all statistically significantly different in the hypothesised directions. Where we compared across three groups, time spent in caregiving, dementia severity, number of caregivers and caregivers' general health, were all statistically significantly different in hypothesised directions. All test results are shown in Table 3, and a summary of the effect sizes is shown in Table 4. When comparing effect sizes between the four instruments, the EQ-HWB sum-score out-performed the other instruments in the variables that related to health, but not to those items related to caregiving, where the C-DEMQOL had higher effect sizes. The carer time and general health questions were consistent whether analysed across 2 or 3 groups.

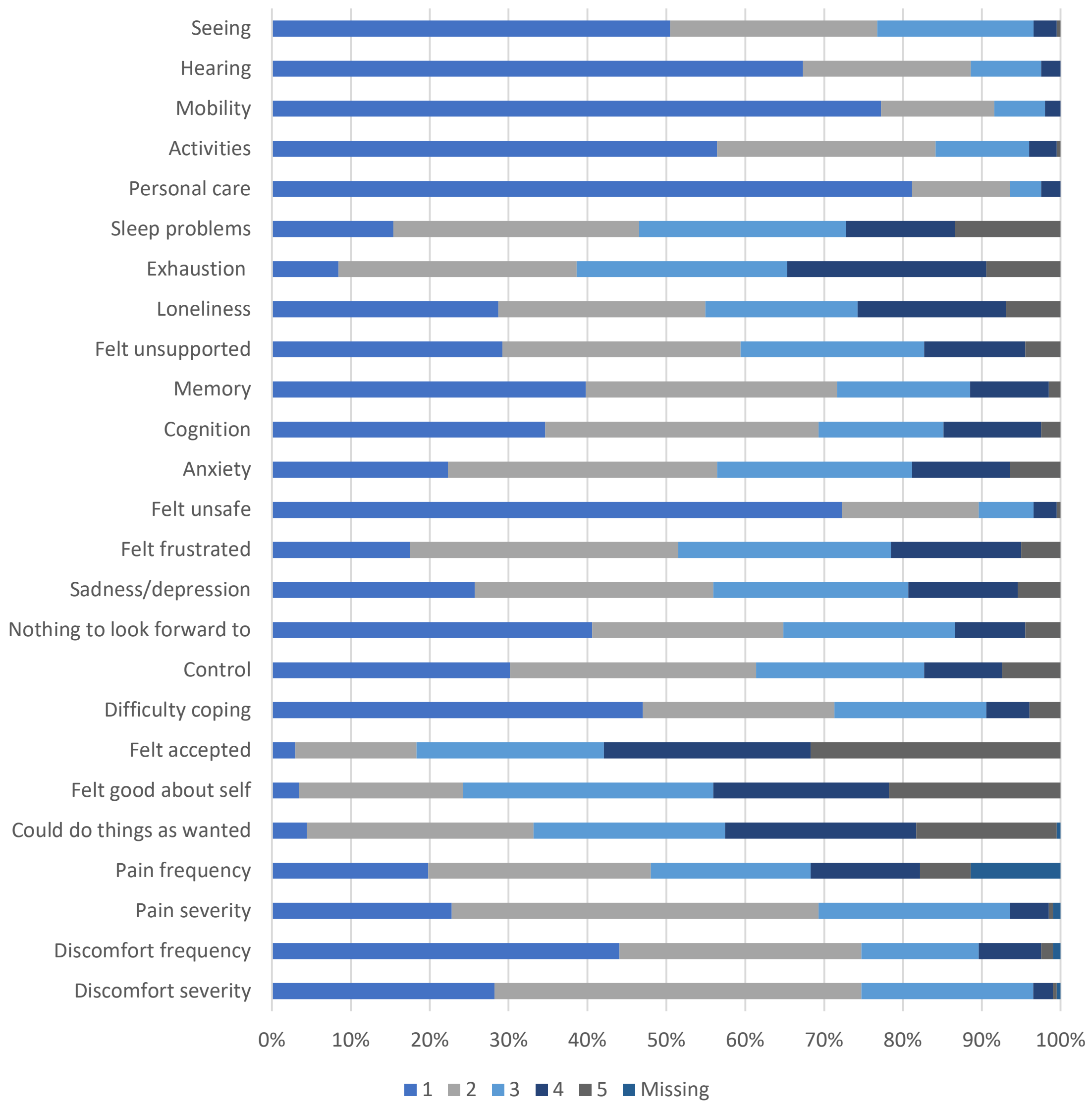


Figure 1

Distribution of EQ-HWB item responses (legend numbers match to the responses displayed in Table 2). OR Percentage of responses by EQ-HWB items over 5 levels where higher scores indicate lower quality-of-life.

Table 2 - Number and percentage of EQ-HWB item responses, with means and standard deviations (SD)

#	EQ-HWB item name	# (%)	# (%)	# (%)	# (%)	# (%)	# (%)	Mean	SD
		No difficulty	Slight difficulty	Some difficulty	A lot of difficulty	Unable	Missing		
1	Seeing	102 (50.5)	53 (26.2)	40 (19.8)	6 (3.0)	1 (0.5)	0 (0)	1.77	0.90
2	Hearing	136 (67.3)	43 (21.3)	18 (8.9)	5 (2.5)	0 (0)	0 (0)	1.47	0.76
3	Mobility	156 (77.2)	29 (14.1)	13 (6.4)	4 (2.0)	0 (0)	0 (0)	1.33	0.69
4	Activities	114 (56.4)	56 (27.7)	24 (11.9)	7 (3.5)	1 (0.5)	0 (0)	1.62	0.83
5	Personal care	164 (81.2)	25 (12.4)	8 (4.0)	5 (2.5)	0 (0)	0 (0)	1.28	0.66
		None of the time	Only occasionally	Sometimes	Often	Most/all of the time	Missing		
6	Sleep problems	31 (15.4)	63 (31.2)	53 (26.2)	28 (13.9)	27 (13.4)	0 (0)	2.79	1.25
7	Exhaustion	17 (8.4)	61 (30.2)	54 (26.7)	51 (25.3)	19 (9.4)	0 (0)	2.97	1.13
8	Loneliness	58 (28.7)	53 (30.2)	39 (19.3)	38 (18.8)	14 (6.9)	0 (0)	2.49	1.27
9	Felt unsupported	59 (29.2)	61 (30.2)	47 (23.3)	26 (12.9)	9 (4.5)	0 (0)	2.33	1.16
10	Memory	80 (39.6)	64 (31.7)	34 (16.8)	20 (9.9)	3 (1.5)	0 (0)	2.02	1.05
11	Cognition	70 (34.7)	70 (34.7)	32 (15.8)	25 (12.4)	5 (2.5)	0 (0)	2.13	1.10
12	Anxiety	45 (22.3)	69 (34.2)	50 (24.8)	25 (12.4)	13 (6.4)	0 (0)	2.47	1.16
13	Felt unsafe	146 (72.3)	35 (17.3)	14 (6.9)	6 (3.0)	1 (0.5)	0 (0)	1.40	0.75
14	Felt frustrated	35 (17.3)	68 (33.7)	54 (26.7)	33 (16.3)	10 (5.0)	0 (0)	2.58	1.11
15	Sadness/depression	52 (25.7)	61 (30.2)	50 (24.8)	28 (13.9)	11 (5.5)	0 (0)	2.43	1.17
16	Nothing to look forward to	82 (40.6)	49 (24.3)	44 (21.8)	18 (8.9)	9 (4.5)	0 (0)	2.12	1.17
17	Control	61 (30.2)	63 (31.2)	43 (21.3)	20 (9.9)	15 (7.4)	0 (0)	2.33	1.22
18	Difficulty coping	95 (47.0)	49 (24.3)	39 (19.3)	11 (5.5)	8 (4.0)	0 (0)	1.95	1.11
		Most/all of the time	Often	Sometimes	Only occasionally	None of the time			
19	Felt accepted	64 (31.7)	53 (26.2)	48 (23.8)	31 (15.4)	6 (3.0)	0 (0)	2.32	1.16
20	Felt good about self	44 (21.8)	45 (22.3)	64 (31.7)	42 (20.8)	7 (3.5)	0 (0)	2.61	1.14
21	Could do things as wanted	36 (17.8)	49 (24.3)	49 (24.3)	58 (28.7)	9 (4.5)	1 (.05)	2.78	1.18
		None of the time	Only occasionally	Sometimes	Often	Most/all of the time	Missing		
22	Pain frequency	40 (19.8)	57 (28.2)	41 (20.3)	28 (13.9)	13 (6.4)	23 (11.4)	2.54	1.21
		No physical pain	Mild	Moderate	Severe	Very severe	Missing		
23	Pain severity	46 (22.8)	94 (46.5)	49 (24.3)	10 (5.0)	1 (0.5)	2 (1.0)	2.13	0.84
		None of the time	Only occasionally	Sometimes	Often	Most/all of the time	Missing		
24	Discomfort frequency	89 (44.1)	62 (30.7)	30 (14.9)	16 (7.9)	3 (1.5)	2 (1.0)	1.91	1.02
		No physical discomfort	Mild	Moderate	Severe	Very severe	Missing		
25	Discomfort severity	57 (28.2)	94 (46.5)	44 (21.8)	5 (2.5)	1 (0.5)	1 (0.5)	2.00	0.81

Table 3

Known group validity analysis with mean scores for caregiver EQ-HWB-S index-scores, C-DEMQOL sum-score and CarerQol index scores (AU) for two and three groups

t-tests EQ-HWB-S index-score	*n*	Mean (SD)	*t* (*df*)	*p*-value	Mean difference	Cohen's *d*	Non-parametric *z* (*p*)
Caregiver chronic health condition (2 groups)			**5.41 (197)**	**<.001**	**0.034**	**0.77**	5.82 (<.001)
No	38	.872 (.116)					
Yes	161	.687 (.202)					
Caregiver health (2 groups)			**7.68 (196)**	**<.001**	**0.213**	**1.10**	6.50 (<.001)
Good to Excellent	140	.786 (.159)					
Fair to Poor	58	.572 (.217)					
Carer-time per day (2 groups)			**2.80 (197)**	**0.003**	**0.079**	**0.40**	3.62 (<.001)
5 hours or less	94	.764 (.207)					
More than 5 hours	105	.685 (.191)					
Caregiver chronic health condition (2 groups)			**5.34 (168)**	**<.001**	**15.47**	**0.82**	5.14 (<.001)
No	32	109.47 (11.71)					
Yes	138	94.00 (15.36)					
Caregiver health (2 groups)			**6.56 (168)**	**<.001**	**15.71**	**1.01**	5.52 (<.001)
Good to Excellent	120	101.53 (13.26)					
Fair to Poor	50	85.82 (16.38)					
Carer-time per day (2 groups)	86		**5.34 (168)**	**0.002**	**7.00**	**0.45**	3.11 (<.001)
5 hours or less	84	100.45 (15.81)					
More than 5 hours	105	93.45 (15.32)					
C-Dem-Qol							
Caregiver chronic health condition (2 groups)			**3.81 (197)**	**<.001**	**14.05**	**0.54**	3.60 (<.001)
No	39	105.87 (21.93)					
Yes	162	91.82 (20.39)					
Caregiver health (2 groups)			**5.69 (198)**	**<.001**	**17.54**	**0.81**	5.23 (<.001)
Good to Excellent	141	99.83 (19.46)					
Fair to Poor	59	82.29 (20.81)					
Carer-time per day (2 groups)			**4.17 (199)**	**<.001**	**12.1**	**0.59**	4.11 (<.001)
5 hours or less	95	100.93 (17.47)					
More than 5 hours	106	88.83 (22.97)					
CarerQol-7D Index-AU							
Caregiver chronic health condition (2 groups)			**4.12 (197)**	**<.001**	**12.46**	**0.59**	4.12 (<.001)
No	39	78.01 (12.03)					
Yes	158	65.55 (17.88)					
Caregiver health (2 groups)			**7.32 (194)**	**<.001**	**18**	**1.05**	5.66 (<.001)
Good to Excellent	139	73.27 (13.01)					
Fair to Poor	57	55.27 (20.72)					
Carer-time per day (2 groups)			**2.66 (195)**	**0.009**	**6.56**	**0.38**	2.87 (<.001)
5 hours or less	92	71.51 (17.05)					
More than 5 hours	105	64.95 (17.53)					

One-way ANOVAs EQ-HWB-S index score	n	M(SD)	F(*df*)	p-value	Scheffe mean difference (*p*)*	Eta Squared	χ^2 (*p*)
Carer-time per day (3 groups)**			**4.12 (2,196)**	**0.018**		0.040	14.20 (<.001)
0-2.8 hours	67	0.765 (0.221)			*Ref cat*		
2.8-11 hours	62	0.735 (0.184)			-0.031 (1.000)		
11 to 18 hours	70	0.670 (0.188)			-0.095 (0.016)		
Dementia Severity (3 groups)			**3.83 (2,196)**	**0.023**		0.038	9.12 (<.001)
Mild or early-dementia	67	0.748 (0.225)			*Ref cat*		
Moderate or middle-stage dementia	98	0.737 (0.175)			-0.011 (1.000)		
Severe or late-stage dementia	32	0.636 (0.207)			-0.112 (.028)		
Shared caregiving (3 groups			**4.06 (2, 195)**	**0.019**		0.040	10.34 (<.001)
Only caregiver	101	0.684 (.203)			*Ref cat*		
One other caregiver	60	0.751 (.192)			0.067 (.126)		

Two or more other caregivers	37	0.781 (.201)			0.097 (.043)		
Caregiver Health (3 groups)			**36.79 (2,195)**	**<.001**		0.274	55.71 (<.001)
excellent/very good	68	0.837 (.135)			*Ref cat*		
good	72	0.737 (.165)			-0.099 (.004)		
fair/poor	58	0.572 (.217)			-0.264 (<.001)		
EQ-HWB sum-score							
Carer-time per day (3 groups)**			**4.84 (2,167)**	**0.009**		0.055	11.44 (<.001)
0-2.8 hours	58	101.69 (16.45)			*Ref cat*		
2.8-11 hours	55	96.20 (14.70)			-5.49 (.188)		
11 to 18 hours	57	92.74 (15.44)			-8.95 (.007)		
Dementia Severity (3 groups)			**2.41 (2,167)**	**0.093**		0.028	5.87 (.053)
Mild or early-dementia	58	99.48 (17.43)			*Ref cat*		
Moderate or middle-stage dementia	87	96.84 (14.80)			-2.64 (0.614)		
Severe or late-stage dementia	25	91.20 (15.04)			-8.28 (.093)		
Shared caregiving (3 groups			**4.60 (2,168)**	**0.011**		0.053	10.37 (<.001)
Only caregiver	86	93.51 (15.10)			*Ref cat*		
One other caregiver	47	99.38 (17.10)			5.87 (0.119)		
Two or more other caregivers	36	102.08 (14.74)			8.57 (.019)		
Caregiver Health (3 groups)			**31.35 (2,197)**	**<.001**		0.273	44.15 (<.001)
excellent/very good	58	106.67 (12.25)			*Ref cat*		
good	62	96.73 (12.42)			-9.95 (<.001)		
fair/poor	50	85.82 (16.38)			-20.85 (<.001)		
C-Dem-Qol							
Carer-time per day (3 groups)**			**9.67 (2,198)**	**<.001**		0.089	18.50 (<.001)
0-2.8 hours	67	101.84 (17.43)			*Ref cat*		
2.8-11 hours	64	95.66 (19.56			-6.18 (.257)		
11 to 18 hours	70	86.54 (23.81)			-15.30 (<.001)		
Dementia Severity (3 groups)			**12.16 (2,197)**	**<.001**		0.110	21.82 (<.001)
Mild or early-dementia	67	103.06 (20.93)			*Ref cat*		
Moderate or middle-stage dementia	101	92.93 (20.74)			-10.13 (0.008)		
Severe or late-stage dementia	32	82.16 (17.35)			-20.90 (<.001)		
Shared caregiving (3 groups			**11.88 (2,197)**	**<.001**		0.108	23.30 (<.001)
Only caregiver	100	88.25 (22.23)			*Ref cat*		
One other caregiver	61	97.48 (18.70)			9.23 (0.022)		
Two or more other caregivers	39	106.26 (17.50)			18.01 (<.001)		
Caregiver Health (3 groups)			**20.61 (2,197)**	**<.001**		0.173	33.36 (<.001)
excellent/very good	69	104.49 (19.76)			*Ref cat*		
good	72	95.36 (18.21)			-9.13 (0.023)		
fair/poor	59	82.29 (20.81)			-22.20 (<.001)		
CarerQol-7D Index-AU							
Carer-time per day (3 groups)**			**4.25 (2,194)**	**0.016**		0.042	9.45 (.008)
0-2.8 hours	64	71.83 (18.06)			*Ref cat*		
2.8-11 hours	63	69.28 (14.94)			-2.55 (0.708)		
11 to 18 hours	70	63.38 (18.49)			-8.45 (0.02)		
Dementia Severity (3 groups)			**1.27 (2,192)**	**0.284**		0.013	3.16 (.206)
Mild or early-dementia	64	69.84 (17.81)			*Ref cat*		
Moderate or middle-stage dementia	99	68.34 (17.37)			-1.51 (0.867)		
Severe or late-stage dementia	32	63.82 (17.99)			-6.02 (0.29)		
Shared caregiving (3 groups			**4.45 (2, 193)**	**0.013**		0.044	7.52 (.023)
Only caregiver	99	64.90 (18.51)			*Ref cat*		
One other caregiver	59	68.90 (16.22)			4.00 (0.374)		
Two or more other caregivers	38	74.62 (15.61)			9.23 (0.014)		
Caregiver Health (3 groups)			**29.77 (2,193)**	**<.001**		0.236	39.41 (<.001)
excellent/very good	68	76.23 (13.38)			*Ref cat*		
good	71	70.43 (12.07)			-5.80 (0.09)		
fair/poor	57	55.26 (20.72)			-20.96 (<.001)		

Table 4

Known group validity analysis effect sizes

Cohen's *d*	**EQ-HWB-S index-score**	**EQ-HWB sum-score**	**C-Dem-Qol**	**CarerQol-7D Index-AU**
Caregiver, other chronic condition	0.77	0.82	0.54	0.59
Carer-time (RUD) per week	0.40	0.45	0.59	0.38
Caregiver Health (2 groups)	1.10	1.01	0.81	1.05
Eta-squared	**EQ-HWB-S index-score**	**EQ-HWB sum-score**	**C-Dem-Qol**	**CarerQol-7D Index-AU**
Carer-time	0.040	0.055	0.089	0.042
Dementia Severity	0.038	0.028	0.110	0.013
Shared caregiving	0.040	0.053	0.108	0.044
Caregiver Health (3 groups)	0.274	0.273	0.173	0.236

3.3.2 Known group validity by EQ-HWB items

We compared two known groups (caregiver time and caregiver general health) across all 25 items, as shown in Table 5 to understand whether each item was contributing to the instrument in terms of measurement and valuation (For example, an item might be given very low weight in some value sets so the instrument overall at utility score level might not discriminate between known groups). We have highlighted non-significant differences to show which items are not significantly different between groups. There were no significant differences between groups on the physical items (Items 1-5), memory, cognition, felt unsafe, and the four pain/discomfort items (Items 22-25) for time spent in caregiving. Only seeing and hearing failed to show significant differences between groups on caregiver general health.

3.3.3 Known group validity controlling for demographic covariates

For each model we ran without (univariate) and with (multivariate) the covariates for caregivers age and gender where the dependent variable was the EQ-HWB-S index score. In the multivariate model 1 (time spent in caregiving – 2 groups), neither gender nor age were significant in the model, and caregiver time remained significant (coefficient = -0.398 (SE = 0.147, *p* = .007). In model 2, caregiver gender was significant in the model (coefficient = -.0377 (SE = 0.148, *p* = .011), and caregiver general health remained significant (coefficient = -1.007 (SE = 0.133, *p* <.001). The table of results is shown in Table 6.

Table 5

Known group validity by EQ-HWB item for caregiving and general health with non-significant differences shaded and positive items in italics

		Caregiving				General health			
#	Item name	**5 hours or less mean**	**<5 hours mean**	**Mean difference**	***p***	**Good mean**	**Poor mean**	**Mean difference**	***p***
1	Seeing	1.779	1.757	-0.022	.568	1.725	1.864	-0.139	.161
2	Hearing	1.453	1.477	0.024	.412	1.430	1.525	-0.096	.207
3	Mobility	1.389	1.280	-0.109	.870	1.239	1.559	-0.320	.001
4	Activities	1.600	1.642	0.042	.362	1.430	2.086	-0.657	<.001
5	Personal care	1.305	1.252	-0.053	.716	1.148	1.576	-0.428	<.001
6	Sleep problems	2.589	2.963	0.373	.017	2.549	3.373	-0.824	<.001
7	Exhaustion	2.779	3.140	0.361	.011	2.697	3.627	-0.930	<.001
8	Loneliness	2.074	2.860	0.786	<.001	2.197	1.106	-0.955	<.001
9	Felt unsupported	1.989	2.636	0.646	<.001	2.099	2.864	-0.766	<.001
10	Memory	1.957	2.065	0.108	.234	1.922	2.237	-0.315	.027
11	Cognition	2.053	2.206	0.153	.163	1.972	2.525	-0.554	.001
12	Anxiety	2.253	2.654	0.402	.007	2.261	2.949	-0.689	<.001
13	Felt unsafe	1.316	1.481	0.165	.059	1.333	1.576	-0.243	.018
14	Felt frustrated	2.290	2.822	0.532	<.001	2.447	2.897	-0.450	.005
15	Sadness/depression	2.147	2.682	0.535	.001	2.218	2.915	-0.697	<.001
16	Nothing to look forward to	1.832	2.383	0.552	<.001	1.859	2.746	-0.887	<.001
17	Control	1.989	2.636	0.646	<.001	2.049	2.983	-0.934	<.001
18	Difficulty coping	1.747	2.131	0.383	.007	1.739	2.458	-0.718	<.001
19	Felt accepted	*3.842*	*3.542*	*-0.300*	*.033*	*3.887*	*3.220*	*0.667*	*<.001*
20	Felt good about self	*3.621*	*3.168*	*-0.453*	*.002*	*3.620*	*2.847*	*0.772*	*<.001*
21	Could do things as wanted	*3.596*	*2.897*	*-0.699*	*<.001*	*3.430*	*2.741*	*0.688*	*<.001*
22	Pain frequency	2.448	2.620	0.171	.172	2.168	3.389	-1.221	<.001
23	Pain severity	2.085	2.170	0.085	.239	1.893	2.695	-0.802	<.001
24	Discomfort frequency	1.800	2.010	0.210	.074	1.695	2.431	-0.736	<.001
25	Discomfort severity	1.947	2.047	0.100	.191	1.823	2.424	-0.601	<.001

3.4 Convergent validity

Of the 45 items we identified *a priori* as likely to be at least moderately correlated between EQ-HWB items and CarerQol items (bolded in Table 7a), 6 hypothesised items failed to meet this criterion. The seeing, hearing and mobility items failed to correlate over .3 with the CarerQol ‘own physical health’ item. The EQ-HWB loneliness and felt unsupported items failed to correlate over .3 with the CarerQol support item. EQH-HWB item felt unsafe did not correlate over .3 with CarerQol items of relational problems and VAS happiness. Of the 15 items expected to correlate between the EQ-HWB and C-DEMQOL domains (Table 7b), all 15 correlated at least moderately.

Table 6

Multivariable generalized linear regression modelling results on EQ-HWB-S utility values

Univariate - Caregiving (2 groups)		Coefficient	SE of Coef.	*z*	*p*-value
Caregiving	less than 5 hours (ref)	-			
	more than 5 hours	-0.398	0.147	-2.70	.007
Constant		1.175	0.120	9.77	<.001
Multivariate					
Caregiving (ref)	less than 5 hours (ref)	-			
	more than 5 hours	-0.379	0.148	-2.57	.010
Caregiver gender (ref)	male (ref)	-			
	female	-0.264	0.156	-1.7	.089
Caregiver age		-0.001	0.005	-0.18	.856
Constant		1.388	0.361	3.85	<.001
Univariate - Health (2 groups)		Coefficient	SE of Coef.	*z*	*p*-value
Health	fair/poor	-			
	good to excellent	-1.007	0.133	-7.57	<.001
Constant		1.298	0.105	12.37	<.001
Multivariate					
Health	fair/poor	-			
	good to excellent	-1.049	0.135	-7.77	<.001
Caregiver gender (ref)	male (ref)	-			
	female	-0.377	0.148	-2.55	.011
Caregiver age		0.004	0.006	0.64	.523
Constant		1.349	0.373	3.62	<.001

Table 7a - Spearman correlations between EQ-HWB and CarerQol items with hypothesised correlations over .3 (in bold).

#	EQ-HWB items	Fulfillment	Relational problems	Own mental health	Combining care tasks	Financial issues	Support	Own physical health	Happiness	CarerQol Index
1	**Seeing**	-0.198**	0.111	0.142**	0.181*	0.232**	0.027	**0.229***	-0.160*	-0.210**
2	**Hearing**	-0.006	0.085	0.131*	0.021	0.125*	0.078	**0.169****	-0.158*	-0.190**
3	**Mobility**	0.016	-0.020	0.138*	0.183**	0.120**	0.141*	**0.288*****	-0.064	-0.239***
4	**Activities**	-0.077	0.132	0.233**	**0.350*****	0.236***	0.047	0.461***	-0.200**	-0.371***
5	**Personal care**	-0.033	0.116	0.076	0.192**	0.228**	0.052	0.270***	-0.093	-0.228**
6	**Sleep problems**	-0.182*	0.218**	0.348***	**0.329*****	0.289***	-0.037	**0.472*****	-0.405***	-0.410***
7	**Exhaustion**	-0.280***	0.295***	0.508***	0.510***	0.384***	-0.018	**0.544*****	**-0.561*****	-0.497***
8	**Loneliness**	-0.208**	**0.351*****	0.502***	0.448***	0.440***	**-0.136**	0.470***	**-0.596*****	-0.508***
9	**Felt unsupported**	-0.319***	0.316***	0.494***	0.453***	0.439***	**-0.175***	0.518***	**-0.563*****	-0.461***
10	**Memory**	-0.098	0.196***	**0.450*****	0.307***	0.260***	0.121*	0.402***	-0.333***	-0.467***
11	**Cognition**	-0.127*	0.266**	**0.592*****	0.428***	0.325***	0.020	0.493***	-0.452***	-0.575***
12	**Anxiety**	-0.214**	0.290***	**0.639*****	0.439***	**0.443*****	-0.038	0.545***	**-0.557*****	-0.581***
13	**Felt unsafe**	-0.098	**0.150***	0.296***	0.328***	0.428***	0.009	0.369***	**-0.178***	-0.390***
14	**Felt frustrated**	-0.432***	**0.418*****	0.555***	0.501***	0.420***	0.021	0.457***	**-0.600*****	-0.512***
15	**Sadness/depression**	-0.247***	0.340***	**0.675*****	0.470***	0.450***	-0.079	0.490***	**-0.646*****	-0.585***
16	**Nothing to look forward to**	-0.286***	0.320***	**0.550*****	0.427***	0.429***	-0.015	0.481***	**-0.586*****	-0.507***
17	**Control**	-0.290***	0.262***	**0.551*****	**0.550*****	0.507***	-0.110	0.506***	**-0.660*****	-0.522***
18	**Difficulty coping**	-0.332***	0.312***	**0.639*****	**0.522*****	0.503***	-0.019	0.531***	**-0.586*****	-0.558***
19	**Felt accepted**	0.242***	-0.222**	**-0.393*****	-0.324***	-0.355***	0.215***	-0.343**	**0.382*****	0.304***
20	**Felt good about self**	**0.338*****	-0.299***	**-0.547*****	-0.431***	-0.344***	0.089	-0.431***	**0.624*****	0.433***
21	**Could do things as wanted**	0.270***	-0.289***	**-0.417*****	**-0.497*****	**-0.397*****	0.127*	-0.375***	**0.510*****	0.370***
22	**Pain frequency**	-0.124	0.178*	0.314***	0.262***	0.267***	0.117	**0.404*****	-0.299***	-0.384***
23	**Pain severity**	-0.140*	0.147*	0.275***	0.237***	0.192**	0.069	**0.425*****	-0.200**	-0.346***
24	**Discomfort frequency**	-0.092	0.098	0.356***	0.155*	0.215**	-0.036	**0.491*****	-0.297***	-0.357***
25	**Discomfort severity**	-0.115	0.107	0.348***	0.190**	0.182*	0.008	**0.514*****	-0.242***	-0.368***
	EQHWB sum-score	0.322***	-0.362***	-0.649***	-0.612***	-0.536***	0.0154	-0.692***	0.656***	0.642***

p<.05=*, *p*<.01=**, *p*<.001 =***; white = Rho < 0.3, light yellow Rho 0.3-< 0.5, dark yellow 0.5 < 0.7, gold Rho >= 0.7, green = items were hypothesized at least moderately but were not. EQ-HWB response scales: difficulty scale (1 = no difficulty, 2 = slight difficulty, 3 = some difficulty, 4 = a lot of difficulty, 5, unable), frequency scale (1 = none of the time, 2 = only occasionally, 3 = sometimes, 4 = often, 5 = most or all of the time), severity scale (1 = no pain, 2 = mild pain, 3 = moderate pain, 4 = severe pain, 5 = very severe pain). ^EQ-HWB-S items.

Table 7b - Spearman correlations between EQ-HWB and C-DEMQOL domains with hypothesised correlations over .3 (in bold).

		C-DEMQOL domain					
#	**EQ-HWB items**	Meeting personal needs	Carer Wellbeing	Carer-patient relationship	Confidence in the future	Feeling supported	C-DEMQOL Total
1	**Seeing**	-0.190**	-0.156*	-0.244***	-0.142*	-0.262***	-0.232**
2	**Hearing**	-0.017	-0.060	-0.093	-0.023	0.016	-0.029
3	**Mobility**	-0.049	-0.034	0.004	-0.040	-0.033	-0.050
4	**Activities**	-0.282***	-0.272***	-0.119	-0.235***	**-0.202****	-0.297***
5	**Personal care**	-0.190**	-0.141*	-0.015	-0.175*	-0.120	-0.175*
6	**Sleep problems**	-0.352***	-0.475***	-0.301***	-0.435***	-0.303***	-0.455***
7	**Exhaustion**	-0.603***	-0.676***	-0.418***	-0.627***	-0.448**	-0.676***
8	**Loneliness**	**-0.587*****	-0.668***	**-0.432*****	-0.564***	-0.441*	-0.659***
9	**Felt unsupported**	**-0.610*****	-0.614***	-0.476***	-0.526***	-0.631***	-0.706***
10	**Memory**	-0.306***	-0.346***	-0.284***	-0.324***	-0.108	-0.338***
11	**Cognition**	-0.439***	-0.517***	-0.322***	-0.448***	-0.229**	-0.481***
12	**Anxiety**	-0.559***	**-0.664*****	-0.418***	-0.591***	-0.328***	-0.632***
13	**Felt unsafe**	-0.291***	-0.265***	-0.221**	-0.287***	-0.192**	-0.298***
14	**Felt frustrated**	-0.605***	**-0.684*****	**-0.620*****	-0.608***	-0.477***	-0.728***
15	**Sadness/depression**	-0.570***	**-0.694*****	-0.483***	-0.617***	-0.341***	-0.654***
16	**Nothing to look forward to**	-0.542***	**-0.582*****	-0.437***	**-0.500*****	-0.336***	-0.591***
17	**Control**	-0.704***	**-0.657*****	**-0.432*****	**-0.609*****	-0.444***	-0.704***
18	**Difficulty coping**	-0.563***	**-0.635*****	**-0.484*****	**-0.586*****	-0.383***	-0.648***
19	**Felt accepted**	0.390***	0.408***	0.398***	0.366***	0.442***	0.484***
20	**Felt good about self**	0.452***	0.535***	0.487***	0.436***	0.388***	0.550***
21	**Could do things as wanted**	0.646***	0.571***	0.395***	0.489***	0.414***	0.618***
22	**Pain frequency**	-0.207**	-0.286***	-0.200**	-0.260***	-0.188*	-0.284**
23	**Pain severity**	-0.194**	-0.240***	-0.208**	-0.180*	-0.192**	-0.242***
24	**Discomfort frequency**	-0.243***	-0.288***	-0.206**	-0.228**	-0.216**	-0.277***
25	**Discomfort severity**	-0.197**	-0.248***	-0.212**	-0.206**	-0.218***	-0.251***
	EQHWB sum-score	0.674***	0.723***	0.510***	0.653***	0.482***	-0.351***

p<.05=*, *p*<.01=**, *p*<.001 =***; white = Rho < 0.3, light yellow Rho 0.3-< 0.5, dark yellow 0.5 < 0.7, gold Rho >= 0.7, green = items were hypothesized at least moderately but were not. EQ-HWB response scales: difficulty scale (1 = no difficulty, 2 = slight difficulty, 3 = some difficulty, 4 = a lot of difficulty, 5, unable), frequency scale (1 = none of the time, 2 = only occasionally, 3 = sometimes, 4 = often, 5 = most or all of the time), severity scale (1 = no pain, 2 = mild pain, 3 = moderate pain, 4 = severe pain, 5 = very severe pain). ^EQ-HWB-S items.

3.5 Exploratory factor analysis (EFA)

An exploratory factor analysis was conducted that included all items from each of the instruments (EQ-HWB, C-DEMQOL, CarerQol). Results from the Kaiser-Mayer-Olkin statistic (.905) and Bartlett's test of sphericity ($\chi^2 = 8234.5$, degrees of freedom = 1953, $p<.001$) confirm that the data were suitable for EFA. Unrotated factor loadings suggested that there were 13 items with eigenvalues over 1; however, the final four factors had items loading higher on other factors, reducing the number extracted for the final solution to 9 (Supplementary Tables not shown). Factors (with variance in brackets) include the caregiver's: 1) Caregiving responsibilities and needs (18.2%), 2) Mental health (17.2%), 3) Cognition (11.4%), 4) Relationship (10.8), 5) Personal support (10.2), 6) Confidence in the future (10.0), 7) Professional support (9.5%), 8) Pain (6.5%), and 9) Functioning (physical ability and safety) (5.7%). Total cumulative variance for the 9 factors was 99.4%. Most EQ-HWB items loaded onto factors 1 (2 items), 2 (6 items), 3 (3 items), 5 (2 items), 8 (4 items) and 9 (4 items). Factors 3, 8 and 9 only included EQ-HWB items. Factors 1,4,5,6 and 7 strongly related to the C-DEMQOL dimensions. Of the four EQ-HWB items that didn't load on any factor, seeing, hearing and sleep are not included in the short form, but exhaustion is.

Discussion

We assessed the psychometric performance of the EQ-HWB and the EQ-HWB-S for measuring quality-of-life of caregivers of people living with dementia. Overall, the instruments performed well in this sample, suggesting appropriateness in this setting. We found significant differences between groups on all known group validity tests. At the item level, items could differentiate between groups for the variables time spent caregiving (2 groups) and caregiver general health (2 groups), differences between groups were retained at the multivariate level for the two variables that were checked, and most hypothesised correlations were moderately correlated. The EQ-HWB instruments were superior on caregiver health items, but not on the caregiver-specific variables (caregiver hours per day, dementia severity or whether there was more than one caregiver) where the C-DEMQOL outperformed the EQ-HWB instruments.

Most EQ-HWB items were well distributed, with the physical items more likely to be negatively skewed. For items included in the EQ-HWB-S, the distribution was similar to previous research in other caregiver populations: parents of children with health conditions [15] and caregivers of children either year or less in families who had experienced adversity

[11]. In all these caregiver populations, we see higher levels of exhaustion than in the Australian population sample [17]. In this sample of caregivers of people living with dementia, we see high levels of sleep problems, feeling frustrated, pain frequency and loneliness. The lack of floor and ceiling effects at the instrument level could be expected, given the large number of items in both EQ-HWB instruments.

Known group analysis was conducted at the EQ-HWB item, sum-score, and EQ-HWB-S index-score levels. When comparing items by general health (2 groups) EQ-HWB item scores were significantly different between groups on all items except Seeing and Hearing. When investigating differences between groups for time spent caregiving (2 groups), most of the emotional and mental health EQ-HWB items were different between groups, but the physical items (Seeing, Hearing, Mobility, Activities, Personal care, Memory, Cognition, Feeling unsafe, and all four of the pain/discomfort items) were not significantly different. This outcome suggests that the EQ-HWB may capture carer impacts mainly via the psycho-social items rather than the physical items. For the EQ-HWB sum-scores and EQ-HWB-S index-scores, general health and caregiver time were significantly different between groups at the 2-group and 3-group levels, suggesting that the instruments were able to differentiate between groups as expected. Being able to determine differences between groups on time spent in caregiving is particularly important for this instrument for it to work well in caregiver populations. Results were comparable between the EQ-HWB sum-scores and the EQ-HWB-S index scores. Given that the EQ-HWB-S is significantly shorter and thus more suitable for inclusion in trials, this result suggests that the use of the EQ-HWB-S could be recommended.

When compared to the two caregiver specific instruments, the C-DEMQOL and the CarerQol, both EQ-HWB instruments performed well. Results from the EQ-HWB sum score *t*-tests suggested that the instrument out-performed the other caregiver instruments on the health items. When comparing 3 groups with ANOVAs, the C-DEMQOL performed better for dementia severity, as could be expected given that it is a dementia specific instrument, as well as time spent caregiving (3 groups) and whether caregiving was shared. We found that our results held when controlling for relevant covariates. Even though the EQ-HWB is a generic instrument that aims to measure quality-of-life in a range of health and social care situations and populations, including for caregivers, it out-performed the CarerQol, which was designed specifically to measure quality-of-life in caregivers. Overall, these results suggest that the EQ-HWB instruments performed well in this population.

In convergent validity analysis, all the correlations we hypothesised to be at least moderate, were so between the EQ-HWB items and C-DEMQOL domains. Of the 45 hypothesised correlations between the EQ-HWB and CarerQol items, 6 items did not meet this criterion. Interestingly, the EQ-HWB items Loneliness and Felt Unsupported did not correlate to at least .3 with the CarerQol Support item. We hypothesised that Felt unsafe would correlate with CareQol item Relationship problems plus the happiness VAS – it is possible that those answering the Felt unsafe question were referring to being physically unsafe rather than experiencing abuse (examples for this question are "*fear of falling, physical harm, abuse")*. Further, feeling unsafe may not be expected to correlate with the happiness VAS. We also expected the EQ-HWB Seeing and Hearing items to correlate at least moderately with the CarerQol Own physical health item, but they did not. The exploratory factor analysis results may also be useful when considering whether to use the long or short form of the instruments [17]. We note that the four EQ-HWB items that didn't load on any factors (seeing, hearing, sleep and exhaustion), also did not load when we ran a 13-factor model. EQ-HWB items were included in the factors of Caregiving responsibilities and needs, Mental health, Cognition, Personal support, Pain, and Functioning (physical ability and safety), but not in Relationships, Confidence in future or Professional support, and more research would be helpful in determining how important these factors might be, to determine which scale is the most useful.

The Seeing and Hearing items (items 1 and 2) appear to be the weakest items in the EQ-HWB. They both had considerably negatively skewed distributions, where most participants had no or slight difficulty, they failed to discriminate between caregiver time (2 groups) and general health (2 groups), did not load onto a factor in the EFA, and did not perform as expected in the convergent validity analyses. These results suggest that these items may not work well in the context of caregiver-related quality-of-life. We note that such items may be important for older people and may work less well for younger caregivers (such as parents); however, our sample in the current study had a mean age of 63 years. Further qualitative work on these items may be needed to determine their relevance in a range of samples.

An important consideration for determining the relative merits of quality-of-life instruments is that they adequately capture relevant domains whilst also balancing length and ease of use [17, 30]. When comparing the EQ-HWB to the EQ-HWB-S, we note that the EQ-HWB-S has almost as high effect sizes as the EQ-HWB, despite the reduced number of items, on the known group analyses. We also note that items in the EQ-HWB-S performed well in terms of item distribution and convergent validity analysis. On the whole, the EQ-HWB-S may be a more suitable tool for determining quality-of-life for economic analysis given the lower number of items leading to lower participant burden. It will also be easier to create value-set for the smaller instrument.

Limitations

Key strengths of this study were the comparisons to two validated scales for caregivers, the use of preference-weights from a recently published pilot value-set, the specific population of caregivers of people living with dementia, and that the sample size was adequate for psychometric analysis. A further strength was that the order of the EQ-HWB, the C-DEMQOL and the CarerQol were randomised in the survey pack, to eliminate participant fatigue relative to each scale (noting also that these scales were presented in a random order after the respondents had already completed 35 candidate items for a separate project).

Limitations to this study include that there were differences in the recall periods of each scale (EQ-HWB was 7 days, C-DEMQOL was 4 weeks, and the CarerQol was "at the moment"). Differences in the recall periods may help explain the lower correlations we found, but we could reasonably expect that if the variables used in the known-groups differences are stable then we can justifiably use as a comparison in known-group validity testing. We also need to better understand the representativeness of the caregivers of people living with dementia, given that this was an online survey. We also note that the value-set used is UK specific and may not capture the nuances of the preferences of an Australian population. The missing data points for item 22 suggest that it is important to determine a clear layout for the items so that respondents do not skip a question by accident. Future research could concentrate on other caregiver populations and, specifically, on the ability of the EQ-HWB to measure change over time, where there is currently only minimal information.

Conclusion

We found that the EQ-HWB instruments showed validity through known group and convergent validity analysis and compared well to the C-DEMQOL and the CarerQol instruments. As the first study to investigate the validity of the EQ-HWB in caregivers of people living with dementia, this study helps build the evidence for use of the EQ-HWB in caregiver populations. The analyses included here demonstrate the potential value of the EQ-

HWB instruments for capturing health and wellbeing dimensions that are not well represented in other multi-attribute utility instruments. Including caregiver quality-of-life in economic evaluations of interventions for people living with dementia is essential to accurately account for spillover effects.

References

1. Brodaty H, Donkin M (2009) Family caregivers of people with dementia. Dialogues Clin Neurosci 11:217–228. https://doi.org/10.31887/DCNS.2009.11.2/hbrodaty

2. Oliva-Moreno J (2017) The Valuation of Informal Care in Cost-of-Illness Studies: A Systematic Review. PharmacoEconomics 35:331–345. https://doi.org/DOI 10.1007/s40273-016-0468-y

3. Brouwer WBF (2019) The Inclusion of Spillover Effects in Economic Evaluations: Not an Optional Extra. PharmacoEconomics 37:451–456. https://doi.org/10.1007/s40273-018-0730-6

4. NICE; National Institute for Health and Care Excellence (2013) Guide to the methods of technology appraisal 2013

5. Dawoud D, Lamb A, Moore A, Bregman C, Rupniewska E, Paling T, Wolfram V, Lovett RES, Dent R (2022) Capturing what matters: updating NICE methods guidance on measuring and valuing health. Qual Life Res 31:2167–2173. https://doi.org/10.1007/s11136-022-03101-6

6. Wu Y, Al-Janabi H, Mallett A, Quinlan C, Scheffer IE, Howell KB, Christodoulou J, Leventer RJ, Lockhart PJ, Stark Z, Boughtwood T, Goranitis I (2020) Parental health spillover effects of paediatric rare genetic conditions. Qual Life Res 29:2445–2454. https://doi.org/10.1007/s11136-020-02497-3

7. Brown A, Page TE, Daley S, Farina N, Basset T, Livingston G, Budgett J, Gallaher L, Feeney Y, Murray J, Bowling A, Knapp M, Banerjee S (2019) Measuring the quality of life of family carers of people with dementia: development and validation of C-DEMQOL. Qual Life Res 28:2299–2310. https://doi.org/10.1007/s11136-019-02186-w

8. EuroQol Research Foundation (2020) EQ-5D. https://euroqol.org/eq-5d-instruments/. Accessed 14 Oct 2020

9. Brazier J, Roberts J, Deverill M (2002) The estimation of a preference-based measure of health from the SF-36. J Health Econ

10. EuroQol Research Group (2022) Extending the QALY Project

11. Brazier, J., Peasgood T, Mukuria C, Marten O, Kreimeier S, Luo N, Mulhern B, Pickard AS, Augustovski F, Greiner W, Engel L, Belizan M, Yang Z, Monteiro A, Kuharic M, Gibbons L, Ludwig K, Carlton J, Connell J, Rand S, Devlin N, Jones K, Tsuchiya A, Lovett R, Naidoo B, Rowen D, Rejon-Parrilla JC (2022) The EQ-HWB: Overview of the Development of a Measure of Health and Wellbeing and Key Results. Value Health 25:482–491. https://doi.org/10.1016/j.jval.2022.01.009

12. Peasgood, T., Mukuria, C., Brazier, J., Marten. O., Kreimeier, S., Nan Luo, Brendan Mulhern, Wolfgang Greiner, A. Simon Pickard, Federico Augustovski, Lidia Engel, Luz Gibbons, Zhihao Yang, Andrea L. Monteiro, Maja Kuharic, Maria Belizan, Jakob Bjørner (2022) Developing a New Generic Health and Wellbeing Measure: Psychometric Survey Results for the EQ-HWB. Value Health 25:525–533. https://doi.org/10.1016/j.jval.2021.11.1361

13. Peasgood T, Mukuria C, Carlton J, Connell J, Devlin N, Jones K, Lovett R, Naidoo B, Rand S, Rejon-Parrilla JC, Rowen D, Tsuchiya A, Brazier J (2021) What is the best approach to adopt for

identifying the domains for a new measure of health, social care and carer-related quality of life to measure quality-adjusted life years? Application to the development of the EQ-HWB? Eur J Health Econ 22:1067–1081. https://doi.org/10.1007/s10198-021-01306-z

14. Mukuria, C., Peasgood, T., McDool, E., Norman, R., Rowen, D., Brazier, J. (2023) Valuing the EQ Health and Wellbeing Short (EQ-HWB-S) using time trade-off and a discrete choice experiment: A feasibility study. Value Health. https://doi.org/10.1016/j.jval.2023.02.008.

15. Bailey C, Dalziel K, Jones R, Hiscock H, Devlin NJ, Peasgood T, In collaboration with the Quality of Life in Kids: Key evidence to strengthen decisions in Australia (QUOKKA) Project team (2024) The Validity of the EuroQol Health and Wellbeing Short Version (EQ-HWB-S) Instrument in Parents of Children With and Without Health Conditions. PharmacoEconomics. https://doi.org/10.1007/s40273-024-01351-5

16. Bailey C, Dalziel K, Constable L, Devlin NJ, Hiscock H, Skouteris H, Peasgood T (2024) The performance of the EQ-HWB-S as a measure of quality-of-life of caregivers in families that have experienced adverse events. Eur J Health Econ. https://doi.org/10.1007/s10198-024-01688-w

17. Lee P, Engel L, Lubetkin E, Gao L (2024) Exploring the Comparability Between EQ-5D and the EQ Health and Wellbeing in the General Australian Population. Value Health S1098301524000329. https://doi.org/10.1016/j.jval.2024.01.004

18. Kuharic M, Mulhern B, Sharp LK, Turpin RS, Pickard AS (2024) Comparison of EQ Health and Well-Being Long and Short With Other Preference-Based Measures Among United States Informal Caregivers. Value Health S1098301524001177. https://doi.org/10.1016/j.jval.2024.03.003

19. Long C, Mao Z, Yang Z (2024) A Head-to-Head Comparison of EQ-HWB and EQ-5D-5L in Patients, Carers, and General Public in China. Value Health 27:848–856. https://doi.org/10.1016/j.jval.2024.02.012

20. Frost MH, Reeve BB, Liepa AM, Stauffer JW, Hays RD, Mayo/FDA Patient-Reported Outcomes Consensus Meeting Group (2007) What is sufficient evidence for the reliability and validity of patient-reported outcome measures? Value Health 10:S94–S105

21. Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, de Vet HC, Terwee CB (2019) COSMIN Study Design checklist for Patient-reported outcome measurement instruments

22. Brouwer WBF, Van Exel NJA, Van Gorp B, Redekop WK (2006) The CarerQol instrument: A new instrument to measure care-related quality of life of informal caregivers for use in economic evaluations. Qual Life Res 15:1005–1021. https://doi.org/10.1007/s11136-005-5994-6

23. Hoefman RJ, van Exel J, Brouwer WBF (2017) Measuring Care-Related Quality of Life of Caregivers for Use in Economic Evaluations: CarerQol Tariffs for Australia, Germany, Sweden, UK, and US. PharmacoEconomics 35:469–478. https://doi.org/10.1007/s40273-016-0477-x

24. Wimo A, Wetterholm AL, Mastey V, Winblad B. (1998) Evaluation of the resource utilization and caregiver time in Anti-dementia drug trials - a quantitative battery. In: The Health Economics of dementia. Wiley’s, London

25. Jones R, Mulhern, B., Devlin, N., Hiscock, H., Chen, G., O’Loughlin, R., Xiong X., Bahrampour, M., McGregor, K., Yip, S., Dalziel, K. (2023) Australian Paediatric Multi-Instrument Comparison (P-MIC) Study: Technical Methods Paper [Online].

26. Sambasivam R, Liu J, Vaingankar JA, Ong HL, Tan M, Fauziana R, Picco L, Chong SA, Subramaniam M (2019) The hidden patient: chronic physical morbidity, psychological distress, and quality of life in caregivers of older adults. Psychogeriatrics 19:65–72. https://doi.org/10.1111/psyg.12365

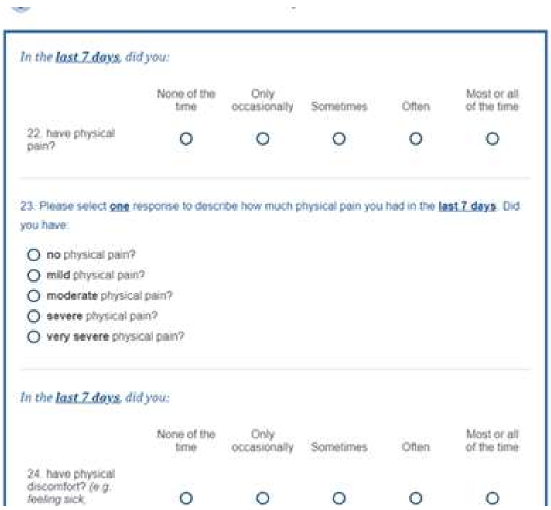
27. Oba H, Matsuoka T, Kato Y, Narumoto J (2018) Factors associated with quality of life of dementia caregivers: Direct and indirect effects. J Adv Nurs 74:2126–2134. https://doi.org/10.1111/jan.13721

28. Thomas P, Lalloué F, Preux P, Hazif-Thomas C, Pariel S, Inscale R, Belmin J, Clément J (2006) Dementia patients caregivers quality of life: the PIXEL study. Int J Geriatr Psychiatry 21:50–56. https://doi.org/10.1002/gps.1422

29. Cohen, J. (1992) Quantitative methods in psychology: A power primer. Psychol Bull 112:1155–1159

30. Devlin, N., Pickard, S., Busschbach, J. (2022) The development of the EQ-5D-5L and its value sets. In: Value Sets for EQ-5D-5L: A Compendium, Comparative Review & User Guide. Springer, Cham, Switzerland

Appendices



*In the **last 7 days**, did you:*

	None of the time	Only occasionally	Sometimes	Often	Most or all of the time
22. have physical pain?	○	○	○	○	○

23. Please select **one** response to describe how much physical pain you had in the **last 7 days**. Did you have:

○ **no** physical pain?
○ **mild** physical pain?
○ **moderate** physical pain?
○ **severe** physical pain?
○ **very severe** physical pain?

*In the **last 7 days**, did you:*

	None of the time	Only occasionally	Sometimes	Often	Most or all of the time
24. have physical discomfort? *(e.g. feeling sick,*	○	○	○	○	○

Figure S1

Image of item placement of EQHWB-22