



This is a repository copy of *The validity of the EuroQol health and wellbeing short version (EQ-HWB-S) instrument in parents of children with and without health conditions.*

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/216736/>

Version: Published Version

Article:

Bailey, C. orcid.org/0000-0001-5030-430X, Dalziel, K. orcid.org/0000-0003-4972-8871, Jones, R. orcid.org/0000-0002-5332-5226 et al. (3 more authors) (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*, 42 (S1). pp. 163-179. ISSN 1170-7690

<https://doi.org/10.1007/s40273-024-01351-5>

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



eprints@whiterose.ac.uk
<https://eprints.whiterose.ac.uk/>



The Validity of the EuroQol Health and Wellbeing Short Version (EQ-HWB-S) Instrument in Parents of Children With and Without Health Conditions

Cate Bailey¹ · Kim Dalziel¹ · Renee Jones¹ · Harriet Hiscock^{2,3,4} · Nancy J. Devlin¹ · Tessa Peasgood¹ · In collaboration with the Quality of Life in Kids: Key evidence to strengthen decisions in Australia (QUOKKA) Project team

Accepted: 2 January 2024 / Published online: 18 January 2024
© The Author(s) 2024

Abstract

Background The EuroQol Health and Wellbeing Short Version (EQ-HWB-S) instrument has been developed to measure the health and wellbeing of care-recipients and their caregivers for use in economic evaluation. The EQ-HWB-S has nine items, and pilot UK preference weights have now been developed.

Objective We aimed to investigate the validity of the instrument in parents of children with and without health conditions.

Methods EQ-HWB-S data were sourced from an Australian paediatric multi-instrument comparison study. We analysed the baseline characteristics and response distribution of the EQ-HWB-S items. Assessment of known-group validity was conducted for EQ-HWB-S items, level sum-scores and preference-weighted scores, including partial effects. Known-group analyses included three child health variables and where caregivers reported coronavirus disease 2019 (COVID-19) had impacted their wellbeing. We included analyses across gender, controlled for child and parent demographic variables, and compared scores across child health conditions.

Results Item responses were distributed as expected, with higher skew for mobility and activities. Parents experienced high levels of exhaustion. We detected significant differences between groups for level sum-scores and preference-weighted scores, as hypothesised; all tests were significant ($p < 0.001$), with moderate effect sizes (effect sizes were slightly higher for female than male parents). The regression analysis identified significantly different EQ-HWB-S scores for child health samples compared with the general population after controlling for demographic variables. Differences were observed between child health conditions.

Conclusion The EQ-HWB-S is a useful instrument to measure parent quality of life for economic evaluation in this population. Data were limited to one time point; further research should investigate the instrument's sensitivity to change and test–retest reliability in this population.

✉ Cate Bailey
cate.bailey@unimelb.edu.au

¹ Melbourne Health Economics, Melbourne School of Population and Global Health, The University of Melbourne, Level 4, 207 Bouverie St, Carlton, VIC 3053, Australia

² Health Services and Economics, Murdoch Children's Research Institute, Melbourne, VIC, Australia

³ Health Services Research Unit, Royal Children's Hospital, Melbourne, VIC, Australia

⁴ Department of Paediatrics, University of Melbourne, Melbourne, VIC, Australia

Key Points for Decision Makers

The EQ-HWB-S is a new nine-item instrument designed for measuring health, social care and carer-related quality of life for use in economic evaluation.

We assess the validity of the EQ-HWB-S in a population of parents of children with and without health conditions, using pilot UK preference weights.

Parents of children with health conditions had poorer wellbeing scores on all items than the general population sample. On all tests, the EQ-HWB-S was shown to be a valid instrument in this population.

1 Introduction

Appropriate and relevant measurement of the health and wellbeing of patients is a priority in the evaluation of health interventions, and plays a key role in economic evaluation. A body of research is now building around the effects on health and wellbeing for informal caregivers and the importance of including these effects in decision-making for funding new health technologies [1]. Examples of the effects on health and wellbeing from caregiving may include declining health when a partner is unwell; caring for a person living with dementia, which can cause fatigue and anxiety; or caring for a child with a disability or health condition, affecting parents' ability to work, sleep and care for themselves and their other children [2]. These effects are often described as 'spillover effects', and there is a strong case for accounting for this impact in economic evaluation where adopting a societal or 'welfare' perspective is now considered recommended practice [1].

Including carer direct health effects has been recommended by National Institute for Health and Care Excellence (NICE) since 2013 [3], and there is sufficient evidence that caregiver effects are measurable [4]. Despite this, a recent paper found that family/caregiver spillover effects were rarely included in technology appraisals and highly specialised technologies (HST) guidance [5]. It is increasingly recognised, as overlooking spillover effects leads to a risk of interventions being undervalued [6].

To include spillover effects for caregivers, appropriate instruments for measuring caregiver wellbeing are required [7]. To this end, a new instrument has been developed to measure health, social care and carer-related health and wellbeing, the EuroQol Health and Wellbeing (EQ-HWB) instrument [7, 8]. This instrument is preference weighted and thus suitable for economic analysis and has been developed to be used in caregiver (and other) populations. As such, the instrument may be a useful measure for quantifying spillover effects for economic evaluation. It is therefore essential that this instrument is evaluated in a range of caregiver populations. The domains and items included in the EQ-HWB instruments were generated predominantly through a large literature review of qualitative work that explored how patients, social care service users and carers report that their conditions, circumstances and related interventions impact their wellbeing (we are unclear whether caregivers of children with health conditions were included in development work) [9–11].

There are two EQ-HWB Instruments, a 25-item instrument (EQ-HWB) and a 9-item instrument [EQ-HWB Short Version (EQ-HWB-S)], which is intended for economic evaluation. The nine-item EQ-HWB-S was chosen for inclusion in the current study, as it was a shorter scale, and

we expected that preference weights would soon be available [12]. Currently the EQ-HWB-S is only available for use in research, as it is still considered experimental [13]. The EQ-HWB-S could be considered a 'cross-over' instrument that can capture and measure elements of health and wellbeing that are relevant to users of both healthcare and/or social care [9]. Thus, the instrument goes beyond the more purely health aspects measured by other health-related quality of life instruments (such as the EQ-5D or the SF-6D [14]).

The generation of the items in the two EQ-HWB instruments has been documented [9, 11, 15], and face validity studies have been conducted across six countries: Argentina, Australia, China, Germany, the UK, and the USA [16]. Whilst it was challenging to ensure validity during development for every group, the conceptual framework [11] and face validity work [16] included carers. Thus, the need for evidence for validity across different groups of carers in different countries remains. Evidence is now beginning to emerge on the validity of the EQ-HWB-S for caregivers. In a study which investigated overlap and differences between the EQ-5D and the EQ-HWB-S [17], both instruments were able to distinguish between participants with and without physical and mental health illnesses, but only the EQ-HWB-S items could distinguish between caregivers and non-caregivers or between groups with high or low caregiver burden.

The evidence is still very limited on the use of the EQ-HWB-S in caregiver populations, and there is no current research on validating the EQ-HWB-S in caregivers of children with health conditions. Having a disabled child is associated with higher risk of parental mental health problems and poorer quality of life compared with parents of healthy children [18]. Caregiving for children living with autism spectrum disorder has also been found to have a negative impact on quality of life [19] as has cerebral palsy [20]. Some of the EQ-HWB-S items may be particularly pertinent to caregivers of children with health conditions, as these caregivers have been found to have higher levels of exhaustion [21, 22], loneliness [23], mental health problems (anxiety and depression), cognition issues and lower sense of control [24]. Research into how well the EQ-HWB-S measures health and wellbeing in parents of children with health conditions is therefore an important next step towards validating this instrument for caregivers/parents of children living with health conditions.

The aim of the current study is to investigate the validity of the EQ-HWB-S in parents of children with and without health conditions, and to assess whether the instrument may be suitable to measure the health and wellbeing of caregivers in this population. We specifically aimed to assess: response distributions by health condition samples compared with the general population, the ability of the instrument to

distinguish between known groups for child health conditions (parents of children in the child health condition samples, with special healthcare needs, where parents reported that the child had a chronic health condition) and a comparison of EQ-HWB-S item scores across a range of child health conditions. We also included whether coronavirus disease 2019 (COVID-19) had negatively affected parents' wellbeing in the known group analysis as a robustness check, to confirm whether the EQ-HWB-S could detect differences in parent's own wellbeing in addition to the impact of caring. There were few parent-specific variables available in this large dataset which had focused on children's health-related quality of life; however, this variable was included, as the data were collected during extended COVID-19 lockdowns in Melbourne.

2 Methods

2.1 Research Design

Data for the study were obtained from the Quality Of Life in Kids: Key Evidence to Strengthen Decisions in Australia (QUOKKA) Research Program's Paediatric Multi-Instrument Comparison (P-MIC) study [25]. The P-MIC study was conducted in 2021–2023 to provide new evidence on the performance of paediatric generic and disease-specific health-related quality of life (HRQoL) measures by comparing their psychometric properties, with the view to informing their use in economic evaluation and in routine hospital care. During the development stage of the P-MIC study, the QUOKKA Consumer Advisory Group provided feedback on the importance of also collecting the quality of life measurement of caregivers. The EQ-HWB-S was subsequently included in the P-MIC survey to inform measurement of caregiver wellbeing and thus enhance the richness of the data. P-MIC data from Data Cut 2, dated 10 August 2022 were used. Data Cut 2 includes approximately 94% of the total planned P-MIC participants.

Data were collected through an online panel, Pure-Profile, and through a sample collected at a paediatric tertiary hospital in Melbourne, Australia (see Technical Methods Paper for further details on the study methodology [26]). The online panel contained eight targeted samples of children with specific health conditions. If none of the children of the participant had these conditions, they were assigned to the 'general population' group. The hospital sample contained healthcare conditions ranging from acute to chronic, and from a wide range of hospital departments, including intensive care, short stay and outpatient clinics. The study was approved by The Royal Children's Hospital (RCH) Human Research Ethics Committee

(HREC/71872/RCHM-2021) and was prospectively registered with the Australia New Zealand Clinical Trials Registry (ANZCTR—ACTRN12621000657820).

2.2 Participants

The full sample included 6787 caregivers (only one caregiver per child). We restricted the sample to parents to enable a focus on the psychometric performance of a relatively homogenous group of caregivers, thus excluding survey respondents who were grandparents ($n=104$), unrelated carers ($n=80$), siblings ($n=43$) or other relatives ($n=43$). We also excluded cases where children were recruited through the hospital but may not have had health conditions ($n=238$). In the hospital sample, if the child presenting to hospital was out of the age range for the condition, parents were asked to complete the questionnaire for another child within the age range, with an emphasis on rarer and more severe health conditions if there was more than one child to choose from; however, we removed these cases, as we were unclear whether they should be in a health condition or general population sample. In total, we excluded 530 survey responses (some participants fell into more than one category), giving a final sample of 6257 survey responses from parents of children between the ages of 2 and 18 years. Of these, 4432 (70.83%) were parents of children with a health condition and 1825 (29.17%) were from the general population. The conditions included in the online panel were recurrent abdominal pain (377, 6.0%), attention deficit/hyperactivity disorder (ADHD; 505, 8.1%), anxiety or depression (452, 7.2%), autism spectrum disorder (ASD; 476, 7.6%), asthma (465, 7.4%), dental problems (472, 7.5%), eating disorders (168, 2.7%), epilepsy (272, 4.4%) and sleep issues (442, 7.1%). The hospital sample contained 807 participants, where the study child may have had a range of medical conditions (condition types are presented in Table S1). We ran analyses excluding the hospital sample and results were found to be robust; therefore, we retained this sample in the analysis. Further information on study recruitment can be found in the published protocol and technical methods paper [25, 26].

2.3 Study Variables

Demographic characteristics for the study child included whether a language other than English was spoken at home, child gender (female, male, other), whether the parent identified that the child was of Aboriginal or Torres Strait Islander descent, whether the study child had a disability, whether the child had special healthcare needs and whether the child had a chronic health condition. Demographic characteristics for parents included gender (female, male, other), household weekly income by four income brackets, education level

in four groups, government healthcare card (government healthcare cards are means tested and issued to those meeting relevant criteria such as receiving payments or supplements, earning low income or meeting age requirements), number of adults in the home and whether the parent was negatively impacted by COVID-19.

The nine EQ-HWB-S dimensions are reported against a five-item response scale [12] (scored 1 to 5). Participants are asked to consider their responses over the last 7 days. The nine items include (1) getting around inside and outside (mobility), (2) ability to do day-to-day activities (activities), (3) feeling exhausted (exhaustion), (4) feeling lonely (loneliness), (5) having trouble concentrating or thinking clearly (cognition), (6) feeling anxious (anxiety), (7) feeling sad or depressed (sad/depressed), (8) feeling as though one has no control over day-to-day life (control) and (9) how much pain they had experienced over the last 7 days (pain). A pilot value set has been developed to score the instrument on the basis of the preferences of the UK public [12]. As an Australian value set is not currently available, this study uses two methods to score the EQ-HWB-S: (i) the UK preference weights with possible values ranging from -0.384 for the worse state described by the instrument to 1 for the best state and (ii) a simple level sum-score which aggregates the score for each of the items on the basis of the reported response level (1 for the best response to 5 for the worst response) giving scores between 9 and 45 (level sum-scores from here on are described as 'sum-scores'). Item questions are presented in Table S2.

2.3.1 Study Variables for Known Group Validity Tests

Differences between known groups were hypothesised a priori by the study team. For the child health conditions, we expected lower EQ-HWB-S preference-weighted scores/higher sum-scores (lower wellbeing) for parents of children in the child health condition samples, with special healthcare needs and where parents reported that the child had a chronic health condition compared with their counterparts. These hypotheses are in line with the QUOKKA statistical analysis plan [26]. For the parent-specific condition, we expected that, where parents reported that COVID-19 had negatively affected them, their wellbeing would be lower compared with their counterparts not reporting COVID-19 impacts.

2.3.1.1 Child Health Condition Sample The QUOKKA technical methods paper [26] defines the following samples: (1) recruited via hospital, (2) general population recruited via online panel and (3) health-condition-specific groups recruited via online panels. Parents of children from the health condition samples (combined Samples 1 and 3) were compared to parents of children from the general population Sample 2.

2.3.1.2 Child Special Healthcare Needs Parents of children with a special healthcare need as per the previously validated Child Special Healthcare Needs Screener (CSHNS) [27] were compared with parents of children without a special healthcare need. Questions included in survey for the SCHNS were: (1) "Does the Study Child currently need or use medicine prescribed by a doctor (other than vitamins)?" and (2) "Does the Study Child need or use more medical care, mental health or educational services than is usual for most children of the same age?" Each of these two questions was followed by: "Is this because of ANY medical, behavioural, or other health conditions? Is this a condition that has lasted or is expected to last for at least 12 months?" [27]. Children who score "yes" to one or more of the first two questions and both follow-up questions are classified as having a special healthcare need.

2.3.1.3 Parent-Reported Child Chronic Health Condition Parents of children with a parent-reported chronic health condition or disability, defined as a health condition expected to last or has lasted at least 6 months ("Does the Study Child have any medical conditions or disabilities that have lasted or are likely to last for six months or more?") were compared with parents of children without a chronic health condition.

2.3.1.4 Impact of COVID-19 on Parent Parents were asked the following question in the survey: "Is your (the parent's) quality-of-life currently affected by the COVID-19 pandemic?". Parents who responded "Yes, in a bad way" were compared with parents who responded "Yes, in a good way", "No, not at all", or "I'm not sure". This was assessed as a binary variable.

2.4 Statistical Analysis

All analyses were performed in Stata version 15 (Statacorp, Texas, US). Participants were required to complete each question in the online survey before they could proceed, thus excluding the possibility of missing data.

2.4.1 Baseline Characteristics and Response Distribution

Demographic variables were compared between child health condition samples and the general population sample using chi-square tests for categorical data and *t*-tests for continuous, normally distributed data. Response distributions were conducted for the full sample, the health conditions sample and the general population sample, by calculating the number and percentage of cases of each response by item. We expected that items 1 (mobility), 2 (activities) and 9 (pain) might have a more skewed distribution than items 3–8 (exhaustion, loneliness, cognition, anxiety, sad/depressed

and control), as we would expect parents of children with health conditions to be less likely to experience issues with mobility, activities and pain than with exhaustion, loneliness, cognition, anxiety, sadness/depression and control. We expected that the health conditions sample would have lower preference-weighted scores and higher sum-scores (worse wellbeing) than the general population sample.

2.4.2 Known Group Validity by Items

We conducted non-parametric (Mann–Whitney) and parametric (*t*-test) tests for known group validity for the health conditions versus general population samples for each EQ-HWB-S item. Non-parametric tests were conducted due to the non-normal distribution of the responses, and *t*-tests were included for interpretability. We expected that the health population samples would have higher mean response scores than the general population sample, and that these differences would be higher for items such as exhaustion [21, 22], loneliness [23], mental health problems (anxiety and depression), cognition issues and lower sense of control [24].

To investigate the difference in the probability of selecting each response option between the health condition and general population samples, we conducted nine ordinal multivariate Probit regression analyses to estimate average partial effects whilst controlling for caregiver age, gender (two groups), number of children in the house (1, 2, 3 or more), single parent in household (yes/no), weekly income (four groups), caregiver education level (four groups) and receiving a government healthcare card (yes/no) [28]. For this analysis, we aggregated the moderate, severe and extreme/unable categories, due to the low observations across the more severe range of responses (hence, we combined responses 3 to 5 to form the third group), as per Henry and Cullinan, 2021 [28].

2.4.3 Known Group Validity for Preference-Weighted Scores and Sum-Scores by a Range of Variables

Known groups for EQ-HWB-S sum-scores and preference-weighted scores were compared using independent *t*-tests, and Cohen's *d* scores to measure effect size [26] across four variables (as outlined above). As per Cohen (1992) [29], effect sizes of 0.2–0.49 were considered small, 0.5–0.79 moderate and ≥ 0.8 large.

2.4.4 Known Group Validity—Subgroup Analyses by Parent Gender

Because the sample was primarily female, we investigated whether the instrument was also valid for male parents [we removed the removed the 23 participants (0.34%) who stated

a gender other than female or male, as it was too small to analyse]. We calculated Cohen's *d* scores to investigate the effect size differences between female and male parents. We used gender rather than sex, as this is how the question was worded in the survey.

2.4.5 Known Group Validity—Controlling for Demographic Covariates

We used linear regression to investigate whether the differences observed between these groups held when we controlled for the parent age and gender, number of children in the household, sole parent status, weekly income, parent education level and receipt of a government healthcare card (as a marker of adversity), as presented in the baseline characteristics table.

2.4.6 Comparison Across Health Conditions

We compared EQ-HWB-S item scores across the different child health condition samples using means and standard deviations across each health condition and ordered them in a table descending by EQ-HWB-S sum-score. Separate regression analyses were conducted for preference-weighted scores and sum-scores with each health condition dummy-coded to investigate the influence of the health condition sample in the model compared with the general population. Beta values were ordered by strength, and the order was compared between preference-weighted scores and sum-scores.

3 Results

3.1 Baseline Characteristics

Baseline characteristics for the general population sample and for the child health condition groups are presented in Table 1. Parents in the general population sample were statistically significantly more likely to have a child speaking a language other than English at home, less likely to have a disability and slightly more likely to be female than the child health condition group. Parents in the general population group were more likely to be female, in a higher income bracket and more highly educated, and less likely to hold a government healthcare card, and less likely to be a single parent than parents in the child health condition group. Children were slightly older in the child health conditions sample, at 10.15 years [standard deviation (SD) = 4.27] compared with the general population sample at 9.69 years (SD = 4.88; $t(6255) = 3.69$, $p < 0.001$) because the age of eligibility was often older for the child health condition groups (e.g. eating disorders). Child age ranged from 2 to 18 years.

Table 1 Baseline characteristics by health condition and general population samples

	Health condition Number (%)	General population Number (%)	Total Number (%)	χ^2 (df)	p-value
Child characteristics					
Child speaks language other than English at home				46.30 (1)	<0.001
No	4098 (92.46)	1588 (87.01)	5686 (90.87)		
Yes	334 (7.53)	237 (12.99)	571 (9.13)		
Child gender				24.71 (2)	<0.001
Female	1993 (44.97)	870 (47.67)	2863 (45.76)		
Male	2359 (53.20)	950 (52.05)	3309 (52.88)		
Other	80 (1.80)	5 (0.27)	85 (1.36)		
Indigenous status*				3.76 (1)	0.052
No	4385 (98.94)	1815 (99.45)	6200 (99.09)		
Yes	47 (1.06)	10 (0.55)	57 (0.91)		
Child has a disability**				469.05 (1)	<0.001
No	3276 (73.92)	1781 (97.59)	5057 (80.82)		
Yes	940 (21.21)	44 (2.41)	984 (15.73)		
Child special healthcare needs				1200.00 (1)	<0.001
No SHCN	1969 (44.43)	1664 (91.18)	3633 (58.06)		
SHDN	2463 (55.57)	161 (8.82)	2624 (41.94)		
Child chronic health conditions				966.59 (1)	<0.001
No chronic condition	2038 (45.98)	1617 (88.60)	3655 (58.41)		
Chronic condition	2394 (54.02)	208 (11.40)	2602 (41.59)		
Parent characteristics					
Parent gender				35.14 (2)	<0.001
Female	3728 (84.12)	1433 (78.52)	5161 (82.48)		
Male	685 (15.44)	390 (21.37)	1075 (17.17)		
Other	19 (0.43)	2 (0.11)	21 (0.34)		
Weekly income				44.54 (4)	<0.001
Less than \$500	291 (6.56)	117 (6.42)	408 (6.52)		
\$500–999	1020 (23.01)	301 (16.49)	1321 (21.11)		
\$1000–1999	1596 (35.98)	672 (36.82)	2268 (36.25)		
\$2000+	1415 (31.93)	701 (38.41)	2116 (33.82)		
Missing	110 (2.48)	34 (1.86)	144 (2.30)		
Parent education level				27.99 (3)	<0.001
Bachelor's or above	1603 (36.17)	751 (41.15)	2354 (37.62)		
Cert III/IV	1692 (38.18)	588 (32.22)	2280 (36.44)		
Year 12 or Cert I/II	654 (14.76)	312 (17.10)	966 (15.44)		
Not completed Year 12	483 (10.90)	174 (9.53)	657 (10.50)		
Government healthcare card***				65.27 (1)	<0.001
No	2757 (62.21)	1327 (72.71)	4084 (65.27)		
Yes	1675 (37.79)	498 (27.29)	2173 (34.71)		
Number of adults in the home				31.23 (3)	<0.001
1	886 (20.00)	259 (14.19)	1145 (18.30)		
2	3026 (68.28)	1343 (73.59)	4369 (69.83)		
3	354 (7.99)	140 (7.67)	494 (7.90)		
4+	166 (3.75)	83 (4.55)	249 (3.98)		
Number of children in the home				206.01 (2)	<0.001
1	1137 (25.65)	737 (40.38)	1874 (29.95)		
2	1903 (42.94)	794 (43.51)	2697 (43.10)		
3+	1392 (31.41)	294 (16.11)	1686 (25.95)		
Impact of COVID-19				52.10 (1)	<0.001
Impacted in a bad way	1338 (35.37)	470 (25.75)	1808 (32.24)		
Other	2445 (64.63)	1355 (74.25)	3800 (67.76)		

Table 1 (continued)

SHCN special healthcare needs, *SHDN* special health disability needs

*Aboriginal or Torres Strait Islander

**Question asked was: “Does the study child have a disability?”

***Government healthcare card (means tested to provide relief for low-income individuals and families)

Parents were slightly younger in the child health conditions samples at 39.60 years ($SD = 7.74$) compared with the general population sample at 40.45 years ($SD = 8.75$; $t(6255) = 3.79$, $p < 0.001$). Parent ages ranged from 18 to 74 years.

3.2 Response Distribution

Response distribution for the nine EQ-HWB-S items is shown in Table 2 for the full, general population and health condition samples. As expected, there was a greater distribution of responses for items 3–8 (exhaustion, loneliness, cognition, anxiety, sad/depressed and control) than items 1 (mobility), 2 (activities) and 9 (pain) in all samples presented. In the full sample, the mobility item was highly skewed, with 85% of participants having no difficulty with mobility. The activities item (item 2) was also skewed, with 64% of participants having no difficulty with their day-to-day activities. For item 9 (pain), only 35.8% of participants had no physical pain. In items 3–8, many participants selected one of the highest two responses for these items: exhaustion (37.9%), loneliness (17.5%), cognition (22.4%), anxiety (26.1%), sad/depression (17.3%) and control (16.1%). Only 8.8% of participants selected that they were exhausted none of the time. The overall trend between the general population and health condition samples was towards wider response spread in the health conditions samples, as expected. The general population and health condition samples are presented graphically in Fig. 1.

3.3 Known Group Validity—Items

We conducted non-parametric (Mann–Whitney) tests and parametric tests (t -tests) to compare item scores between the health condition samples and the general population sample. Non-parametric and parametric tests were comparable; thus, we present the t -tests in Table 3, as they are more easily interpretable (non-parametric tests are included in Table S3). There were significant differences between health condition samples and the general population samples on all items, as expected. We had hypothesised that mean differences would be greater for the exhaustion, loneliness, cognition, mental health (anxiety and sad/depression) and control than for mobility, activities and pain. Mean differences for exhaustion, loneliness, cognition, mental health (anxiety and sad/depression) and

control were between 0.45 and 0.62, whilst mean differences for mobility, activities and pain were between 0.12 and 0.36, confirming our hypothesis.

The estimated average partial effects for the health condition versus general population samples are presented in the final columns of Table 3. Our key findings from this analysis were that there was a consistent trend across all dimensions for the health conditions sample to have worse item scores than the general population sample when controlling for key demographics of the caregiver. For instance, the results suggest that respondents in the health conditions group were 7.8 percentage points (ppts) less likely to report no difficulty with mobility problems, 3.0 ppts more likely to have slight difficulty with mobility and 4.8 ppts more likely to have some difficulty, have a lot of difficulty or indicate unable. Similar independent associations were found across all domains. The partial effects from the ordered probit models suggest that items with the greatest probability of being in the third category are cognition, exhaustion and anxiety.

3.4 Known Group Validity—Preference-Weighted Scores and Sum-Scores

In all t -tests where the outcome variable was the EQ-HWB-S preference-weighted score or sum-score, there were significant differences between known groups in the hypothesised directions with p -values of less than 0.001, as shown in Table 4 for the whole sample, and subgroups for women and men. Effect sizes (Cohen’s d) were moderate for the three child health variables [29] and small for parent quality of life (QoL) impacted by COVID-19. Women had slightly higher effect sizes (0.45–0.50) than men (0.40–0.50) for the child health conditions. There were stronger differences in effect sizes from COVID-19 on women (0.40) compared with men (0.32). A histogram of the distribution of the EQ-HWB preference-weighted score is shown in Fig. S1, and for the sum-score in Fig. S2.

3.5 Known Group Validity—Subgroup Analyses by Parent Gender

Effect sizes were similar between female and male parents, with slightly lower Cohen’s d scores for male parents on all variables (Table 4).

Table 2 Frequency of observations by EQ-HWB-S item for full, general population and child health condition samples ($N=6257$)

	Full sample $n=6257$	General population $n=1825$	Child health condition $n=4432$	Percentage difference (%)
	Number (%)	Number (%)	Number (%)	
Mobility				
No difficulty	5305 (84.8)	1638 (89.8)	3667 (82.7)	-7.0
Slight difficulty	463 (7.4)	97 (5.3)	366 (8.3)	2.9
Some difficulty	323 (5.2)	59 (3.2)	264 (5.9)	2.7
A lot of difficulty	137 (2.2)	22 (1.2)	115 (2.6)	1.4
Unable	29 (0.5)	9 (0.5)	20 (0.5)	0.0
Activities				
No difficulty	3997 (63.9)	1414 (77.5)	2583 (58.3)	-19.2
Slight difficulty	1213 (19.4)	267 (14.6)	946 (21.3)	6.7
Some difficulty	715 (11.4)	97 (5.3)	618 (13.9)	8.6
A lot of difficulty	291 (4.7)	40 (2.2)	251 (5.7)	3.5
Unable	41 (0.7)	7 (0.4)	34 (0.8)	0.4
Exhaustion				
None of the time	548 (8.8)	287 (15.7)	261 (5.9)	-9.8
Only occasionally	1576 (25.2)	591 (32.4)	985 (22.2)	-10.2
Sometimes	1762 (28.2)	533 (29.2)	1229 (27.7)	-1.5
Often	1559 (24.9)	311 (17.0)	1248 (28.2)	11.1
Most or all of the time	812 (13.0)	103 (5.6)	709 (16.0)	10.4
Loneliness				
None of the time	2321 (37.1)	875 (47.9)	1446 (32.6)	-15.3
Only occasionally	1322 (21.1)	387 (21.2)	935 (21.1)	-0.1
Sometimes	1520 (24.3)	377 (20.7)	1143 (25.8)	5.1
Often	808 (12.9)	145 (7.9)	663 (14.9)	7.0
Most or all of the time	286 (4.6)	41 (2.3)	245 (5.5)	3.3
Cognition				
None of the time	1384 (22.1)	620 (33.9)	764 (17.2)	-16.7
Only occasionally	1725 (27.6)	566 (31.0)	1159 (26.2)	-4.9
Sometimes	1745 (27.9)	423 (23.2)	1322 (29.8)	6.7
Often	1043 (16.7)	169 (9.3)	874 (19.7)	10.5
Most or all of the time	360 (5.8)	47 (2.6)	313 (7.1)	4.5
Anxiety				
None of the time	1478 (23.6)	639 (35.0)	839 (18.9)	-16.1
Only occasionally	1599 (25.6)	509 (27.9)	1090 (24.6)	-3.3
Sometimes	1547 (24.7)	387 (21.2)	1160 (26.2)	5.0
Often	1214 (19.4)	220 (12.1)	994 (22.4)	10.4
Most or all of the time	419 (6.7)	70 (3.8)	349 (7.9)	4.0
Sad/depression				
None of the time	2007 (32.1)	790 (43.3)	1217 (27.5)	-15.8
Only occasionally	1762 (28.2)	522 (28.6)	1240 (28.0)	-0.6
Sometimes	1406 (22.5)	311 (17.0)	1095 (24.7)	7.7
Often	778 (12.4)	159 (8.7)	619 (14.0)	5.3
Most or all of the time	304 (4.9)	43 (2.4)	261 (5.9)	3.5
Control				
None of the time	2568 (41.0)	984 (53.9)	1584 (35.7)	-18.2
Only occasionally	1450 (23.2)	396 (21.7)	1054 (23.8)	2.1
Sometimes	1229 (19.6)	278 (15.2)	951 (21.5)	6.2
Often	690 (11.0)	125 (6.9)	565 (12.8)	5.9
Most or all of the time	320 (5.1)	42 (2.3)	278 (6.3)	4.0

Table 2 (continued)

	Full sample <i>n</i> = 6257	General population <i>n</i> = 1825	Child health condition <i>n</i> = 4432	Percentage difference (%)
	Number (%)	Number (%)	Number (%)	
Pain				
No physical pain	2241 (35.8)	798 (43.7)	1443 (32.6)	-11.2
Mild pain	2474 (39.5)	735 (40.3)	1739 (39.2)	-1.0
Moderate pain	1140 (18.2)	221 (12.1)	919 (20.7)	8.6
Severe pain	317 (5.07)	55 (3.0)	262 (5.9)	2.9
Very severe pain	85 (1.4)	16 (0.9)	69 (1.6)	0.7

Cognition is interpreted here as impaired cognition; control is interpreted as loss of control

Fig. 1 a Distribution of EQ-HWB-S item responses for the general population sample (legend numbers match to the responses displayed in Table 2). **b** Distribution of EQ-HWB-S item responses for the health condition samples (legend numbers match to the responses displayed in Table 2).

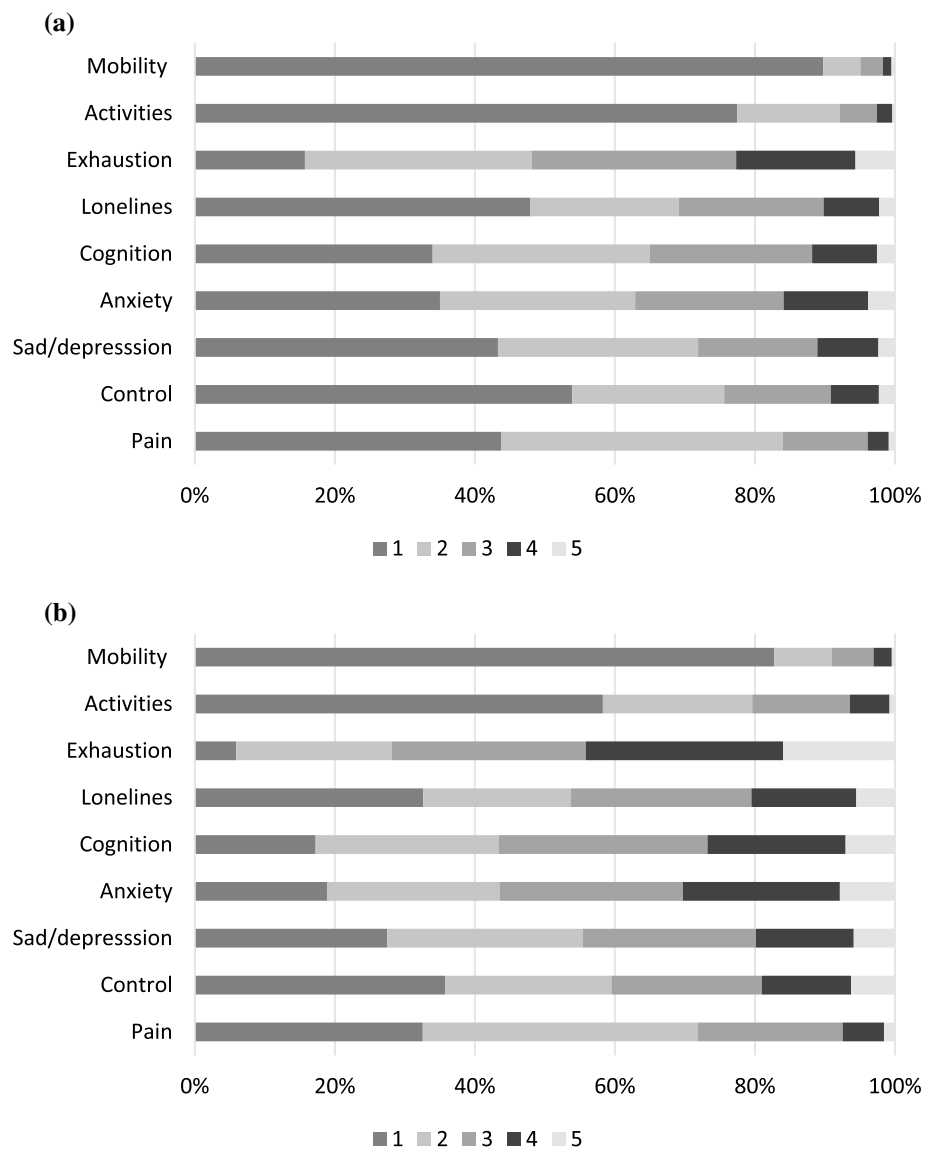


Table 3 Known group comparison analysis for EQ-HWB-S items by health condition samples ($n=4432$) and the general population sample ($n=1825$) and estimated partial effects across EQ-HWB-S items

EQ-HWB-S	t-tests				Partial effects		
	Health condition Mean (SD)	General population Mean (SD)	Mean difference	$t(p)$	No problems*	Slight problems*	Moderate to severe problems*
Mobility	1.30 (0.73)	1.17 (0.59)	0.12	6.43 (<0.001)	-0.078 (0.010)	0.030 (0.004)	0.048 (0.007)
Activities	1.70 (0.96)	1.33 (0.71)	0.36	14.42 (<0.001)	-0.179 (0.013)	0.059 (0.004)	0.120 (0.009)
Exhaustion	3.26 (1.15)	2.64 (1.11)	0.62	19.55 (<0.001)	-0.068 (0.005)	-0.091 (0.007)	0.159 (0.011)
Loneliness	2.40 (1.23)	1.95 (1.10)	0.44	13.33 (<0.001)	-0.120 (0.012)	-0.004 (0.001)	0.124 (0.012)
Cognition	2.73 (1.17)	2.15 (1.07)	0.58	18.23 (<0.001)	-0.134 (0.009)	-0.048 (0.004)	0.182 (0.012)
Anxiety	2.76 (1.22)	2.22 (1.16)	0.54	16.14 (<0.001)	-0.127 (0.009)	-0.039 (0.003)	0.166 (0.012)
Sad/depression	2.43 (1.19)	1.98 (1.08)	0.45	13.80 (<0.001)	-0.132 (0.011)	-0.011 (0.002)	0.143 (0.012)
Control	2.30 (1.25)	1.82 (1.07)	0.48	14.44 (<0.001)	-0.153 (0.012)	-0.006 (0.001)	0.148 (0.012)
Pain	2.05 (0.95)	1.77 (0.84)	0.28	10.78 (<0.001)	-0.113 (0.011)	0.017 (0.002)	0.096 (0.010)

Regarding the regression analysis, covariates included parent age and gender, number of children in the household, sole parent status, weekly income, parent education level, and receipt of a government healthcare card. Partial effects were obtained from multivariate, ordered Probit regression models (full results in Table S4)

SD standard deviation

*Problems: For mobility and activities=no difficulty, slight difficulty, some and a lot of difficulty and unable; for exhaustion, loneliness, cognition, anxiety, sad/depression and control=none of the time/only occasionally/sometimes and often and most or all of the time; for pain=no physical pain, mild physical pain, moderate and severe and very severe physical pain

3.6 Known Group Validity—Controlling for Demographic Covariates

Multiple linear regression was used to test if a range of demographic variables, found to be significantly different between groups in Table 1, were associated with EQ-HWB-S preference-weighted scores and sum-scores (the dependent variables). All independent variables were retained in the models. We used r^2 to test model fit. We tested for multicollinearity by conducting correlation analysis for all included independent variables, and none were higher than 0.80 (income and healthcare card were correlated at 0.55). The results for the EQ-HWB-S preference-weighted score and sum-score regression models are presented in Table 5. The variables of the child health condition sample (child health condition samples versus general population), parent age, parent gender (male, female), single parent household, weekly income (four groups) and government healthcare card (yes/no) were significantly associated with a better outcome for the EQ-HWB-S preference-weighted scores ($F(10, 6082) = 96.53, p < 0.001, r^2 = 0.14$). Holding a government healthcare card and having a child in the health condition sample (compared with general population) had the largest associations with the EQ-HWB-S preference-weighted score, thus confirming our hypotheses. Further, higher income was associated with higher parent EQ-HWB-S preference-weighted scores (higher wellbeing). The analysis using the sum-scores, as shown in the second section of the table,

showed similar results: ($F(10, 6082) = 111.13, p < 0.001, r^2 = 0.15$).

3.7 Comparison Across Health Conditions

EQ-HWB-S items, preference-weighted scores and sum-scores were compared across all eight child health condition samples, the hospital sample and the general population sample. Table 6 shows means and standard deviations across EQ-HWB-S items; the table is ordered by descending sum-score values. Overall, parents of children with ASD, eating disorders and sleep problems had higher EQ-HWB-S item and sum-scores means (lower wellbeing), whilst parents of children in the general population, dental problem, asthma and the hospital sample had lower EQ-HWB-S item scores and sum-scores (higher wellbeing). The items with the largest wellbeing impacts were exhaustion, cognition and anxiety. In Fig. 2, the pattern is similar between the health condition samples and the general population; however, the general population sample have overall lower EQ-HWB-S item scores (higher wellbeing).

The regression analyses with dummy coding for health conditions to investigate the influence of each health condition sample in the model compared with the general population are shown in Table S5. When ordered by the absolute beta values to compare preference-weighted scores with sum-scores, the conditions for preference-weighted scores were ordered: autism spectrum disorder

Table 4 Known group comparison analysis for total sample, women and men caregivers for EQ-HWB-S preference-weighted scores and sum-scores, over four known groups

Total sample	Yes		No		Mean difference	<i>t</i>	<i>p</i>	Cohen's <i>d</i> ⁵
	<i>n</i>	Mean (SD)	<i>n</i>	Mean (SD)				
Preference-weighted score								
Child health condition ¹	4432	0.702 (0.228)	1825	0.810 (0.187)	0.108	17.92	<0.001	0.45
Special healthcare needs ²	2624	0.670 (0.243)	3633	0.779 (0.194)	0.109	19.73	<0.001	0.50
Chronic health condition ³	2602	0.667 (0.245)	3655	0.781 (0.191)	0.114	20.66	<0.001	0.52
Parent QoL impacted by COVID ⁴	1808	0.670 (0.228)	3800	0.762 (0.214)	0.092	14.76	<0.001	0.39
Sum-score								
Child health condition ¹	4432	20.91 (7.08)	1825	17.05 (6.29)	-3.86	-20.24	<0.001	0.51
Special healthcare needs ²	2624	21.89 (7.31)	3633	18.27 (6.51)	-3.63	-20.65	<0.001	0.52
Chronic health condition ³	2602	21.94 (7.33)	3655	18.25 (6.47)	-3.69	-21.03	<0.001	0.53
Parent QoL impacted by COVID ⁴	1808	22.10 (6.94)	3800	18.75 (6.90)	-3.35	-16.95	<0.001	0.45
Women (<i>n</i> = 5161) ⁶								
Preference-weighted score								
Child health condition ¹	3728	0.693 (0.226)	1433	0.801 (0.185)	0.108	16.09	<0.001	0.45
Special healthcare needs ²	2254	0.664 (0.236)	2907	0.769 (0.196)	0.105	17.42	<0.001	0.49
Chronic health condition ³	2271	0.663 (0.192)	2890	0.771 (0.192)	0.108	17.97	<0.001	0.50
Parent QoL impacted by COVID ⁴	1512	0.659 (0.226)	3072	0.752 (0.212)	0.092	13.66	<0.001	0.40
Sum-score								
Child health condition ¹	3728	21.29 (6.99)	1433	17.52 (6.25)	-3.77	-17.88	<0.001	0.50
Special healthcare needs ²	2907	22.22 (7.10)	2907	18.72 (6.52)	-3.50	-18.41	<0.001	0.51
Chronic health condition ³	2271	22.19 (7.17)	2890	18.72 (6.47)	-3.47	-18.23	<0.001	0.51
Parent QoL impacted by COVID ⁴	1512	22.53 (6.82)	3072	19.23 (6.84)	-3.31	-15.41	<0.001	0.46
Men (<i>n</i> = 1010) ⁶								
Preference-weighted score								
Child health condition ¹	685	0.752 (0.231)	390	0.844 (0.190)	0.092	6.69	<0.001	0.41
Special healthcare needs ²	355	0.714 (0.272)	720	0.821 (0.182)	0.107	7.64	<0.001	0.47
Chronic health condition ³	315	0.703 (0.277)	759	0.820 (0.183)	0.118	8.16	<0.001	0.50
Parent QoL impacted by COVID ⁴	293	0.728 (0.225)	717	0.806 (0.220)	0.078	5.06	<0.001	0.32
Sum-score								
Child health condition ¹	685	18.82 (7.14)	390	15.34 (6.17)	-3.48	-8.07	<0.001	0.49
Special healthcare needs ²	355	19.74 (8.05)	720	16.48 (6.16)	-3.16	-7.36	<0.001	0.45
Chronic health condition ³	316	20.11 (8.08)	759	16.50 (6.21)	-3.62	-7.93	<0.001	0.48
Parent QoL impacted by COVID ⁴	293	19.84 (7.10)	717	16.71 (6.85)	-3.13	-6.52	<0.001	0.41

QoL quality of life, SD standard deviation

¹All child health condition samples versus the general population sample, as defined in Sect. 2.3

²Special healthcare needs of the child

³'Chronic health condition' is defined here as the child having a chronic health condition for more than 6 months

⁴Impact on parent's quality of life from COVID-19 pandemic, defined as: affected in bad way by COVID-19 versus affected in a good way, not affected or not sure

⁵Cohen's *d* effect sizes defined as 0.2–0.49 are considered small, 0.5–0.79 moderate, and ≥ 0.8 large [29]

⁶0.34% of participants defined their gender as 'other' and were removed from the analysis

(beta score = -0.177), eating disorder (-0.174), recurrent abdominal pain (-0.125), sleep disorder (-0.121), anxiety/depression (-0.109), ADHD (-0.101), epilepsy (-0.093), hospital sample (multiple conditions; -0.058), asthma (-0.033) and dental (-0.027), as shown

in Table S6. There were only slight differences in the relative weights of the conditions when using preference-weighted scores and sum-scores: autism spectrum disorder, eating disorders, recurrent abdominal pain and sleep disorder had the highest beta values indicating higher

Table 5 Regression analysis summary for parents' EQ-HWB-S sum-scores and EQ-HWB-S preference-weighted scores

	Model 1	
	Coefficient (SE)	<i>t</i> (<i>p</i>)
Index score		
General population (non-health-condition sample)	0.092 (0.006)	15.50 (<0.001)
Parent age	0.002 (0.0003)	4.81 (<0.001)
Parent gender (female)	-0.031 (0.007)	-4.23 (<0.001)
Number of children in household		
1	Base	
2	0.011 (0.006)	1.78 (0.09)
3+	0.012 (0.007)	1.63 (0.16)
Single-parent household	-0.021 (0.008)	-2.75 (0.007)
Weekly household income		
Less than \$500 per week	Base	
\$500-999 per week	0.028 (0.012)	2.33 (0.020)
\$1000-1999 per week	0.048 (0.012)	3.96 (<0.001)
\$2000 or more per week	0.076 (0.013)	5.92 (<0.001)
Education level (bachelor's degree or higher)	0.018 (0.006)	3.08 (0.002)
Healthcare card holder	-0.081 (0.007)	-11.64 (<0.001)
Constant	0.663 (0.025)	26.91 (<0.001)
Sum-score		
General population (non-health-condition sample)	-3.31 (0.19)	-17.62 (<0.001)
Parent age	-0.09 (0.01)	-8.12 (<0.001)
Parent gender (female)	1.52 (0.23)	6.67 (<0.001)
Number of children in household		
1	Base	
2	-0.26 (0.20)	-1.29 (0.20)
3+	-0.24 (0.23)	-1.03 (0.31)
Single-parent household	0.83 (0.24)	3.43 (0.001)
Weekly household income		
Less than \$500 per week	Base	
\$500-999 per week	-0.64 (0.37)	-1.74 (0.08)
\$1000-1999 per week	-1.19 (0.38)	-3.13 (0.002)
\$2000 or more per week	-2.29 (0.40)	-5.70 (<0.001)
Education level (bachelor's degree or higher)	-0.48 (0.18)	-2.63 (0.009)
Healthcare card holder	2.18 (0.22)	9.95 (<0.001)
Constant	22.30 (0.78)	28.71 (<0.001)

SE standard error

(1) Government healthcare cards are used here as a measure of adversity. These cards are means tested and issued to those meeting relevant criteria for low income. (2) Higher preference-weighted scores and lower sum-scores mean higher wellbeing

impact on caregiver wellbeing when using both preference-weighted scores and sum-scores and the lowest in both were dental and asthma. In general, there was a high level of consistency in conditions showing the largest impact regardless of whether sum-scores or preference weighted scores were used. Overall, autism spectrum disorder, eating disorders, recurrent abdominal pain and sleep disorder had the highest beta values, indicating high caregiving impact on health and wellbeing.

4 Discussion

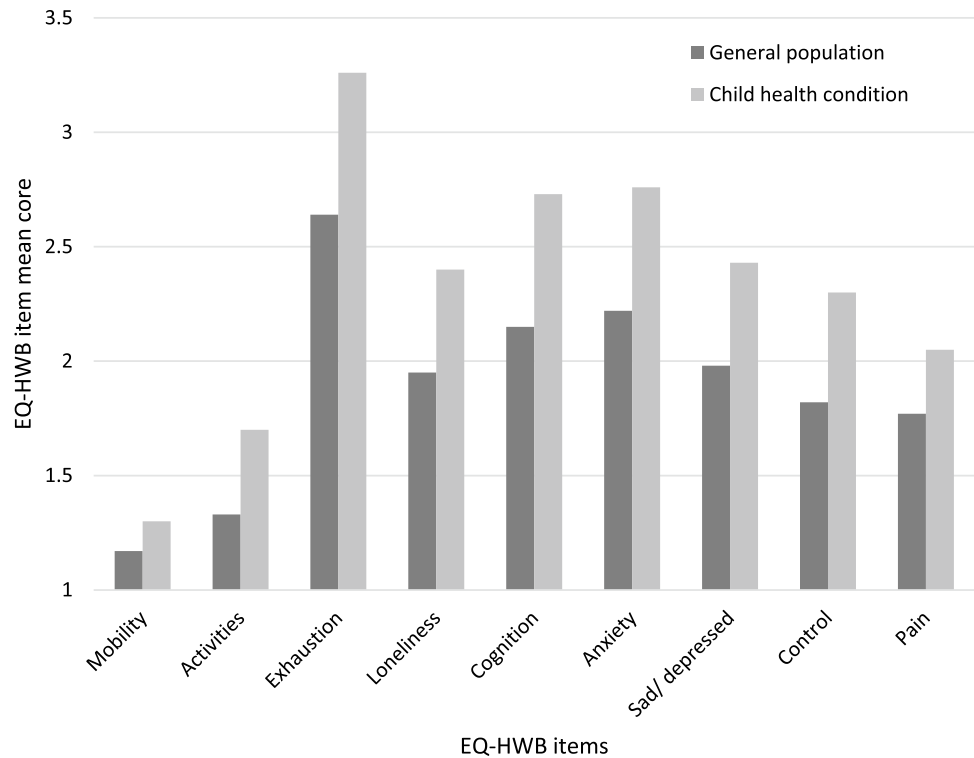
In this study, we investigated the psychometric performance of the EQ-HWB-S in parents of children with and without health conditions. Overall, the EQ-HWB-S performed well in this large sample of parents, suggesting that it is a valuable instrument for measuring spillover effects for economic evaluation. We found that there were significant differences between known groups measured

Table 6 Means of parent EQ-HWB-S item scores, sum-scores and preference-weighted scores for child health condition samples, ranked by EQ-HWB-S sum-score

Child's health condition	Number (%)	EQ-HWB-S items; M (SD)									Parent M (SD)	
		Mobility	Activity	Exhaustion	Loneliness	Cognition	Anxiety	Sad/depressed	Control	Pain	EQ-HWB-S preference-weighted score	EQ-HWB sum-score
ASD	476 (7.60)	1.42 (0.84)	2.04 (1.09)	3.62 (1.09)	2.56 (1.25)	3.11 (1.18)	3.20 (1.20)	2.81 (1.23)	2.61 (1.26)	2.23 (1.06)	0.606 (0.258)	23.93 (7.32)
Eating disorder	168 (2.68)	1.40 (0.86)	1.90 (1.10)	3.58 (1.13)	2.85 (1.33)	3.05 (1.20)	3.12 (1.16)	2.95 (1.29)	2.70 (1.35)	2.28 (1.04)	0.609 (0.270)	23.82 (7.54)
Sleep	442 (7.06)	1.22 (0.65)	1.76 (0.95)	3.60 (1.04)	2.55 (1.19)	3.03 (1.10)	3.00 (1.14)	2.53 (1.22)	2.49 (1.25)	2.08 (0.91)	0.672 (0.208)	22.27 (6.46)
Abdominal pain	377 (6.02)	1.26 (0.66)	1.70 (0.96)	3.49 (1.06)	2.41 (1.24)	2.92 (1.10)	2.86 (1.23)	2.52 (1.17)	2.36 (1.24)	2.34 (0.97)	0.672 (0.226)	21.83 (6.63)
ADHD	505 (8.07)	1.29 (0.72)	1.75 (0.97)	3.39 (1.13)	2.54 (1.24)	2.89 (1.11)	2.85 (1.22)	2.55 (1.19)	2.42 (1.24)	2.03 (0.92)	0.685 (0.215)	21.75 (6.86)
Anxiety or depression	452 (7.22)	1.39 (0.81)	1.65 (0.89)	3.27 (1.15)	2.55 (1.25)	2.77 (1.11)	2.94 (1.14)	2.62 (1.19)	2.44 (1.27)	2.10 (0.96)	0.682 (0.222)	21.69 (6.84)
Epilepsy	272 (4.34)	1.32 (0.71)	1.77 (0.97)	3.21 (1.26)	2.42 (1.29)	2.72 (1.18)	2.73 (1.25)	2.36 (1.15)	2.37 (1.29)	1.89 (0.93)	0.710 (0.225)	20.81 (7.39)
Hospital	807 (12.89)	1.24 (0.66)	1.64 (0.96)	3.12 (1.14)	2.14 (1.16)	2.54 (1.18)	2.51 (1.20)	2.18 (1.12)	2.14 (1.23)	1.91 (0.89)	0.748 (0.215)	19.40 (6.92)
Asthma	465 (7.43)	1.30 (0.78)	1.47 (0.85)	2.81 (1.09)	2.03 (1.12)	2.32 (1.13)	2.41 (1.17)	2.12 (1.09)	1.90 (1.10)	1.92 (0.96)	0.768 (0.217)	18.32 (6.69)
Dental	472 (7.54)	1.23 (0.67)	1.43 (0.78)	2.88 (1.06)	2.07 (1.11)	2.36 (1.08)	2.39 (1.17)	2.12 (1.13)	1.99 (1.14)	1.90 (0.86)	0.775 (0.193)	18.38 (6.22)
General population	1825 (29.15)	1.17 (0.58)	1.33 (0.71)	2.64 (1.10)	1.95 (1.10)	2.15 (1.07)	2.22 (1.15)	1.98 (1.07)	1.82 (1.07)	1.77 (0.83)	0.810 (0.187)	17.05 (6.29)
Total	6257 (100)	1.26 (0.69)	1.59 (0.91)	3.08 (1.17)	2.27 (1.21)	2.56 (1.17)	2.60 (1.23)	2.30 (1.18)	2.16 (1.22)	1.97 (0.93)	0.733 (0.223)	19.79 (7.08)

M mean, *SD* standard deviation

Fig. 2 Mean parent EQ-HWB-S scores by item for child health condition and general population samples



using level sum-scores and preference-weighted scores, all in expected directions. Effect sizes were moderate overall, and similar between female and male parents. The significant differences in EQ-HWB-S scores between the health condition sample and general population sample were stable when controlling for demographic variables. Items that were hypothesised to have larger mean differences when comparing parents with and without children with health conditions conformed to expectations. These results concord with the limited amount of previous research in caregiver populations [17, 30].

In terms of distribution, the mobility item was highly skewed with few mobility issues in the sample; however, significant mobility issues were not expected in this parent population. EQ-HWB-S items 3–8 (exhaustion, loneliness, cognition, anxiety, sad/depressed and control) had wider decrements in wellbeing than items 1, 2 and 9 (mobility, activities and pain) as was expected. There were consistent differences in response distributions between the health condition and general population samples, which were corroborated through further analysis using partial effects.

In known group validity testing, EQ-HWB-S sum-scores were different between groups across all tests in hypothesised directions, suggesting that the EQ-HWB-S preference-weighted scores and sum-scores were sensitive to group differences. The EQ-HWB-S preference-weighted scores and sum-scores gave similar results in these tests; however, the effect sizes for the preference-weighted scores were slightly lower, probably due to the pilot weights being stronger for

mobility and pain than items such as exhaustion, loneliness, cognition and control, which were higher in this population [12]. There were larger mean differences between groups for exhaustion, loneliness, cognition, mental health and control than for mobility, activities and pain, as expected in this population.

We found similar patterns between the child health condition sample and the general population sample, with exhaustion, cognition and anxiety scoring highest in both cohorts, but with parents of children with health conditions scoring lower on preference-weighted scores and higher on sum-scores (lower wellbeing) than the general population sample. The high levels of exhaustion found in this study are consistent with previous research on parents' quality of life [31], and particularly for parents of children with health conditions [32, 33]. The hospital sample was closer to the general population than most of the child health condition samples, possibly because their conditions were more likely to be acute and resolved more quickly or were being actively managed in a way that impacts were minimised.

The baseline comparison between parents with and without a child in one of the health condition samples (child health condition samples versus general population) indicated that there were significant demographic differences between groups. We therefore conducted a regression analysis to determine whether these differences between the groups on EQ-HWB-S scores remained when we controlled for demographic variables. Differences between child health condition samples versus general population were

maintained with this variable having the highest coefficient, suggesting that differences in EQ-HWB-S scores between parents with and without children with medical conditions remain when controlling for background demographic circumstance. The regression analysis confirmed our hypothesis that parents with a child with a health condition had lower EQ-HWB preference-weighted scores (lower wellbeing) than those in the general population sample, even when controlling for other factors such as holding a healthcare card (an indicator of adversity), which was also significant in the model.

Given recent evidence of the importance of separate analysis by sex in clinical research and with journals increasingly requiring analysis separately for female and male participants [34], it was important to determine how well the EQ-HWB-S instrument performed for male caregivers, who only made up 18% of the sample. We know that the burden of unpaid work is disproportionately carried by women, which exposes them to greater mental health risk [35]. Our sample accords with understood societal norms, where women (82% of this sample) may be expected to have a higher caregiving load than male caregivers. Through the analysis by gender, we found that the EQ-HWB-S appeared to be fit for purpose for both female and male parents. The effect size of the differences in EQ-HWB-S scores between caregivers of children with and without health conditions was higher for women than men in all three health condition known groups analyses, suggesting that, when caring for a child with a health condition, women may carry more of this load. The effect sizes for the COVID-19 variable were also larger for women, concordant with recent evidence [36].

4.1 Strengths and Limitations

The current study features a large dataset with responses from over 6000 children and families with data on specific child health conditions, as well as a general population sample, allowing detailed exploration of differences between groups, and supplying a valuable addition to the literature on the validity of the EQ-HWB-S in this population. A limitation of this study was that we did not have information on how much of the caregiving burden fell to each parent, or of other caring roles participants in this sample may undertake, such as caring for elderly parents. This is pertinent to the analysis by gender; it is possible that fathers do less overall caregiving than mothers, thus accounting for the slightly lower effect sizes. It would therefore be important for future studies to collect information on caregiving intensity and duration, to further examine differences between groups and explore caregiving impact. We note that we did not include the 0.34% of participants who did not specify either male or female as their gender in the analysis due to the very small

size of this portion of the sample. We did not have a variable for parent health status, which would have been useful to prevent confounding in the regression analysis. We were unable to assess EQ-HWB-S responsiveness to change or test-retest reliability due to the data collection being limited to one time point. Further, a head-to-head study of the EQ-HWB-S with other caregiver measures would aid in determining how well this instrument compares with other instrument options for measuring caregiver health and wellbeing.

The level sum-score has known limitations, as the same sum-score can have quite different profiles; thus, giving equal weight to each dimension makes assumptions about their relative importance [37]. To our knowledge, this is the first study to apply the pilot UK preference weights to EQ-HWB-S data. Here, we have demonstrated the use of these weights and compared them with the sum-scores throughout the paper. A limitation to the use of these weights is that different preference weights, especially if they were designed for an Australian population, could potentially lead to somewhat different results. Here, though, we see that the sum-scores and preference-weighted data are largely complementary, and we do not see differences in test significance between the utility and level sum-scores. There are several other caregiver HRQoL tools such as the CarerQoL, Caring Experience Scale and the ASCOT-Carer (adult social care), for which utility values are available. The key difference between these instruments and the EQ-HWB-S is that these measures focus only on carer quality of life, whereas the EQ-HWB-S is a generic instrument designed to capture carer, patient and social care quality of life, thereby enabling benefits to be measured in the units for both patients and their caregivers. In future studies we plan to compare the EQ-HWB-S with the CarerQoL to further understand the strengths and weaknesses of the EQ-HWB-S.

4.2 Conclusions

This study found that the EQ-HWB-S performed well and had good validity in comparing known groups in a population of parents of children with and without health conditions. The measure also appeared to be valid for both female and male parents. Due to the large sample size and the availability of different variables to perform known group validity, this paper adds to the literature building towards a greater understanding of the use of the EQ-HWB-S generally, and specifically for this population. Future work should focus on a more in-depth investigation of the EQ-HWB-S in this population to determine responsiveness to change and test-retest reliability, and whether the instrument can detect caregiver intensity and duration in known groups validity testing.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40273-024-01351-5>.

Declarations

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. QUOKKA is funded through the Medical Future Research Fund (MRFF) Grant number APP1200816.

Conflict of Interest The authors declare that they have no conflict of interest. Professor Devlin and Dr. Peasgood are members of the EuroQol group.

Availability of Data and Material N/A.

Ethics Approval Noted in text.

Consent to Participate Yes.

Author Contributions Conceptualization: CB, KD and TP. Methodology: CB and TP. Formal analysis and investigation: CB and TP. Writing—original draft preparation: CB. Writing—review and editing: All. Funding acquisition: ND, KD and HH. Resources: N/A. Supervision: N/A.

Disclosure This article is published in a journal supplement wholly funded by the Australian Government MRFF funded QUOKKA research grant 1200816, the University of Melbourne, the EuroQol Research Foundation, the University of Technology Sydney and Flinders University.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

References

- Brouwer WBF. The inclusion of spillover effects in economic evaluations: not an optional extra. *Pharmacoeconomics*. 2019;37(4):451–6. <https://doi.org/10.1007/s40273-018-0730-6>.
- Wittenberg E, James LP, Prosser LA. spillover effects on caregivers' and family members' utility: a systematic review of the literature. *Pharmacoeconomics*. 2019;37(4):475–99. <https://doi.org/10.1007/s40273-019-00768-7>.
- NICE; National Institute for Health and Care Excellence. 'Guide to the methods of technology appraisal 2013' [Online]; 2013. <https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781>. Accessed Sep. 27, 2023.
- Prosser LA, Lamarand K, Gebremariam A, Wittenberg E. Measuring family HRQoL spillover effects using direct health utility assessment. *Med Decis Making*. 2015;35(1):81–93. <https://doi.org/10.1177/0272989X14541328>.
- Dawoud D, et al. Capturing what matters: updating NICE methods guidance on measuring and valuing health. *Qual Life Res*. 2022;31(7):2167–73. <https://doi.org/10.1007/s11136-022-03101-6>.
- Wu Y, et al. Parental health spillover effects of paediatric rare genetic conditions. *Qual Life Res*. 2020;29(9):2445–54. <https://doi.org/10.1007/s11136-020-02497-3>.
- EuroQol Research Group. 'Extending the QALY Project' [Online]; 2022. <https://euroqol.org/extending-the-qaly-project-collaboration-with-nice/#:~:text=NICE%20is%20taking%20part%20in,social%20care%20and%20public%20health>.
- Brazier J, et al. Extending the QALY (E-QALY) Project: stages for developing a new generic measure. Sheffield: The University of Sheffield; 2020.
- Brazier J, et al. The EQ-HWB: overview of the development of a measure of health and wellbeing and key results. *Value Health*. 2022;25(4):482–91. <https://doi.org/10.1016/j.jval.2022.01.009>.
- Peasgood T, et al. 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*. 2021;22(7):1067–81. <https://doi.org/10.1007/s10198-021-01306-z>.
- Mukuria C, et al. Qualitative review on domains of quality of life important for patients, social care users, and informal carers to inform the development of the EQ-HWB. *Value Health*. 2022;25(4):492–511. <https://doi.org/10.1016/j.jval.2021.11.1371>.
- Mukuria C, Peasgood T, McDool E, Norman R, Rowen D, Brazier J. Valuing the EQ Health and Wellbeing Short (EQ-HWB-S) using time trade-off and a discrete choice experiment: a feasibility study. *Value Health*. 2023. <https://doi.org/10.1016/j.jval.2023.02.008>.
- EuroQol. 'Category C: Experimental version' [Online]. <https://euroqol.org/eq-5d-instruments/ip-protection/experimental-version/#:~:text=An%20Experimental%20Version%20is%20a,classified%20as%20a%20Beta%20Version>.
- Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21:271–92.
- Peasgood T, et al. Developing a new generic health and wellbeing measure: psychometric survey results for the EQ-HWB. *Value Health*. 2022;25(4):525–33. <https://doi.org/10.1016/j.jval.2021.11.1361>.
- Carlton J, et al. Generation, selection, and face validation of items for a new generic measure of quality of life: the EQ-HWB. *Value Health*. 2022;25(4):512–24. <https://doi.org/10.1016/j.jval.2021.12.007>.
- Monteiro AL, Kuharic M, Pickard AS. A comparison of a preliminary version of the EQ-HWB short and the 5-level version EQ-5D. *Value Health*. 2022;25(4):534–43. <https://doi.org/10.1016/j.jval.2022.01.003>.
- Isa SNI, et al. Health and quality of life among the caregivers of children with disabilities: a review of literature. *Asian J Psychiatry*. 2016;23:71–7. <https://doi.org/10.1016/j.ajp.2016.07.007>.
- Hoefman RJ, van Exel J, Brouwer WB. Measuring the impact of caregiving on informal carers: a construct validation study of the CarerQol instrument. *Health Qual Life Outcomes*. 2013;11(1):1–13.
- Dogruoz Karatekin B, Icagasioglu A. The effects of the functional levels of children with cerebral palsy on the quality of life of caregivers. *J Surg Med*. 2022;6(2):191–5. <https://doi.org/10.28982/josam.1061762>.
- Lindström C, Åman J, Norberg A. Increased prevalence of burnout symptoms in parents of chronically ill children. *Acta Paediatr*. 2010;99(3):427–32. <https://doi.org/10.1111/j.1651-2227.2009.01586.x>.

22. Morse CA, Messimeri-Kianidis V. Keeping it in the family: caregiving in Australian-Greek families. *Soc Work Health Care*. 2002;34(3–4):299–314. https://doi.org/10.1300/J010v34n03_04.
23. Florian V, Krulik T. Loneliness and social support of mothers of chronically ill children. *Soc Sci Med*. 1991;32(11):1291–6. [https://doi.org/10.1016/0277-9536\(91\)90045-E](https://doi.org/10.1016/0277-9536(91)90045-E).
24. Falk NH, Norris K, Quinn MG. The factors predicting stress, anxiety and depression in the parents of children with autism. *J Autism Dev Disord*. 2014;44(12):3185–203. <https://doi.org/10.1007/s10803-014-2189-4>.
25. Jones R, et al. Psychometric performance of HRQoL measures: an Australian paediatric multi-instrument comparison study protocol (P-MIC). *Children*. 2021. <https://doi.org/10.3390/children8080714>.
26. Jones R, et al. Australian Paediatric Multi-Instrument Comparison (P-MIC) Study: technical methods paper [Online]; 2023. <https://www.quokkaresearchprogram.org/project-1-1>.
27. Bethell C, Read D, Stein R, Blumberg S, Wells N, Newacheck P. Identifying children with special health care needs: development and evaluation of a short screening instrument. *Ambul Pediatr*. 2002;2(1):38–48.
28. Henry E, Cullinan J. Mental health spillovers from serious family illness: doubly robust estimation using EQ-5D-5L population normative data. *Soc Sci Med*. 2021;279: 113996. <https://doi.org/10.1016/j.socscimed.2021.113996>.
29. Cohen J. Quantitative methods in psychology: a power primer. *Psychol Bull*. 1992;112:1155–9.
30. Kuharic M, Monteiro A, Pickard AS. P37 comparison of the newly developed EQ-HWB-S to the ASCOT and EQ-5D-5L as an outcome measure for caregivers. *Value Health*. 2022;25(7):S294.
31. Roskam I, Raes M-E, Mikolajczak M. Exhausted parents: development and preliminary validation of the parental burnout inventory. *Front Psychol*. 2017;8:163. <https://doi.org/10.3389/fpsyg.2017.00163>.
32. Graaf G, Baiden P, Keyes L, Boyd G. Barriers to mental health services for parents and siblings of children with special health care needs. *J Child Fam Stud*. 2022;31(3):881–95. <https://doi.org/10.1007/s10826-022-02228-x>.
33. Giallo R, Wood CE, Jellett R, Porter R. Fatigue, wellbeing and parental self-efficacy in mothers of children with an autism spectrum disorder. *Autism*. 2013;17(4):465–80. <https://doi.org/10.1177/1362361311416830>.
34. Wainer Z, et al. Sex and gender in health research: updating policy to reflect evidence. *Med J Aust*. 2020;212(2):57. <https://doi.org/10.5694/mja2.50426>.
35. Ervin J, Taouk Y, Alfonzo LF, Hewitt B, King T. Gender differences in the association between unpaid labour and mental health in employed adults: a systematic review. *Lancet Public Health*. 2022;7(9):e775–86. [https://doi.org/10.1016/S2468-2667\(22\)00160-8](https://doi.org/10.1016/S2468-2667(22)00160-8).
36. Seedat S, Rondon M. Women’s wellbeing and the burden of unpaid work. *BMJ*. 2021. <https://doi.org/10.1136/bmj.n1972>.
37. Devlin N, Parkin D, Janssen B. Methods for analysing and reporting EQ-5D data. Cham: Springer International Publishing; 2020. <https://doi.org/10.1007/978-3-030-47622-9>.