Dear Editor,

We thank Aubeeluck, Buchanan and Stupple (ABS) for their comments [23] regarding our recent paper on the psychometric properties of the Huntington’s Disease Quality of Life Battery for Carers (HDQoL-C) and the Alzheimer’s Carers Quality of Life Inventory (ACQLI) among carers of people with Huntington’s disease (HD) [1], and the Editor for providing us with this opportunity to clarify the issues raised in the letter.

It is true that our sample size was small. This was acknowledged [1], together with the fact that studies [2] have shown that samples of this size (or smaller; \( n \geq 20 \)) can be useful and provide stable interpretations relative to larger samples for basic traditional psychometric analyses. For example, the FACIT-F, a widely used fatigue scale, was initially tested with a sample of \( n = 49 \) [3]; subsequent larger-scale studies have not contradicted initial results.

Arguing that single-centre samples are problematic means that single-centre studies are inappropriate for the HDQoL-C. This is because traditional psychometric properties are not fixed scale characteristics but sample dependent, and should be supported within the sample in which a scale is used [4–6]. For example, the developers of the SF-36, one of the most widely used rating scales in the health sciences, argue that basic scaling assumptions should always be met before interpreting SF-36 data [7, 8].

We find the comment that our data are homogenous curious. First, and relevant also to the previous point, our sample characteristics were very similar to those previously reported in psychometric HDQoL-C studies [9, 10]. Secondly, all HDQoL-C scores ranged across their full or close to full possible ranges [1]. This aspect is important, since failure to do so limits the interpretability of psychometric data due to lack of sample representation across the latent continuum [11]. Had our HDQoL-C data been restricted to a narrow range, there would have been a problem related to “homogenous data”. This was not the case. Although our sample was recruited from a single centre, the recruitment area covered four English counties (a fact that unfortunately was unacknowledged in our paper).

We agree that factor analysis (FA) is common in psychometric studies. However, there are good arguments against factor analysing item-level data [12]. The prime reason for conducting item-level FA is to assess dimensionality [13]. Due to our limited sample size and inherent problems with such analyses, we took an alternative approach, multitrait/multi-item scaling analysis [11, 13]. Although addressing the same fundamental issue as FA, multitrait/multi-item scaling...
has fewer assumptions and sources of misinterpreta-
tions. Our results suggested dimensionality problems
with all HDQoL-C domain and total scores except
the Satisfaction with Life domain. As pointed out [1],
these results are in general agreement with previous
HDQoL-C studies [9, 10]. In those studies, no analy-
ses of the dimensionality of the full HDQoL-C were
presented but principal component analyses (a form of
FA) were applied to each proposed HDQoL-C domain
individually. This yielded 2-3 components per domain,
suggesting multidimensionality. While noted [9, 10],
this has been unacknowledged in the recommended use
of the HDQoL-C; instead, total domain scores and a
total HDQoL-C score are recommended [14, 15]. This
is in contrast to basic psychometric assumptions [12].

It is unclear what ABS mean by “standard tests
of validity”. In essence, there are three related steps
to traditional rating scale validation: (i) Defining and
operationalizing the target variable(s); (ii) testing scal-
ifying assumptions and internal validity; (iii) testing a
priori hypotheses regarding relationships between
scores and external criteria (external construct valid-
ity). We considered the first of these and focused on
the second [1]. This involved tests addressing whether
items within proposed scales appear to represent an
underlying latent variable, as well as tests of the legit-
immacy of calculating summed total scores according to
traditional (Likert) psychometric theory [7, 16]. Such
support is essentially a prerequisite for confident use
and interpretation of scores [4–8, 11]. Assuming that
“standard tests of validity” refers to external construct
validity, we did not address this due to general lack
of supported scaling assumptions/internal construct
validity for the HDQoL-C, and because such tests
provide generally weak and circumstantial evidence
[17, 18].

This leads us to the conceptual underpinnings of the
HDQoL-C. ABS argue that we are incorrect in say-
ing that the HDQoL-C lacks explicit definitions of its
target variables and that content validity was not
informed by carers. However, nowhere in the HDQOL-
C publications known to us are Practical aspects of
Caringgiving, Satisfaction with Life or Feelings about
Living with HD explicitly defined. As these are the pro-
posed HDQOL-C scores [14, 15], these are the target
variables. As for the total HDQol-C score, the tar-
get variable is supposedly quality of life (QoL), which
ABS imply is based on the WHO definition of QoL.
This is puzzling, since it is not mentioned in previous
publications [9, 10, 19]; instead, the guiding concep-
tual framework referred to is that of Cummins [20]. The
literature does state that the content of the HDQoL-C is
based on qualitative work involving HD carers [19] and
that the three HDQoL-C domains were identified as
pertinent to the QoL of spousal HD carers [9]. Indeed,
HD carers were involved in the development of the
HDQoL-C by (i) reviewing a generic QoL instrument
(based on Cummins’ QoL definition) for relevance;
(ii) documenting what compromises or enhances their
QoL (as defined by Cummins); (iii) describing how
caring affects their QoL; and (iv) commenting on the
appropriateness, difficulty and clarity of HDQoL-C
items [9, 19]. As for content validity of the HDQoL-C,
“two experts in the field of QoL and two experts in
the field of HD were asked to comment on the item
content” [9].

Most psychometric methods, including those
employed in our [1] and other HDQoL-C studies [9,
10] assume that items are observable manifestations of
the latent variable. That is, variations in item responses
should reflect variations in, e.g. QoL – not the other
way around [13]. Failure to meet basic psychometric
criteria and to provide evidence that scale generated
numbers represent measures rather than just numerals
suggests that the scale does not qualify as a measure-
ment tool [18, 21]. The development and psychometric
testing of the HDQoL-C [1, 9, 10, 19] provide reasons to
question its appropriateness as a tool for measure-
ment. This is concerning, particularly since it has been
suggested as an outcome measure [9, 10].

While carers may welcome reflecting on aspects
covered by the questionnaire, item response rates and
questionnaire completion time provide reasons to con-
sider if the HDQoL-C could be improved from a
respondent perspective. HDQoL-C completion time in
our study was very similar to that previously reported
[9], and was ×4 times longer than that of our compara-
tor questionnaire [1]. As for response rates, missing
item responses ranged from 0-11.5% [1]. It is unknown
how this compares to previous studies since response
rates were not reported [9, 10], although Cox (2002)
reported a 17.5% overall HDQoL-C response rate [22].

In our study [1], we also tested a carer QoL question-
naire developed for carers of people with Alzheimer’s
disease (the ACQLI) to further explore the notion that
carer impact has similarities across neurodegenerative
disorders. Given that the HDQoL-C failed to outper-
form the ACQLI we concluded that there is “tentative
support for the idea of a common cross-diagnostic carer
impact of neurodegenerative disorders” and that “the
ACQLI appears a promising candidate tool for further
work towards such cross-diagnostic QoL assessments”
[1]. This does not argue against disease specific
instruments.
ABS criticise the fact that the ACQLI is not freely available on the Internet. This is true and common practice for reasons including quality assurance through protection against questionnaire modifications leading to different (unsupported/untested) versions and translations. Sample ACQLI pages are available on the Internet and the full questionnaire can be obtained for non-commercial use following signed agreement and a nominal administrative fee (www.galen-research.com).

We have accepted the invitation from ABS to share our data, provided that data are subjected to rigorous state-of-the-art analyses using modern psychometric methods (Rasch analysis), and offered to assist in conducting and interpreting such analyses. We hope to receive a positive response. We also hope to see further independent psychometric analyses of the HDQoL-C, since scale validation is an iterative quantitative and qualitative process involving theoretical, substantive and empirical considerations [4, 18]. If we take our carers, patients and studies seriously, we also need to be serious about the outcome measures we use. Unless rating scales are treated with full scientific rigour, advances in the clinical sciences will be hampered and opportunities to improve care may be lost.

REFERENCES


[14] Aubeeluck A, Buchanan H. Huntington’s Disease Quality of Life Battery for Carers (HDQoL-C) - Validated for use with family Carers of Persons with Huntington’s Disease [user guide]. University of Nottingham. Nottingham, UK. 2007.


