

Brief Methodological Report



Assessment of Burden Among Family Caregivers of People With Parkinson's Disease Using the Zarit Burden Interview

Peter Hagell, PhD, Anette Alvariza, PhD, Albert Westergren, PhD, and Kristofer Årestedt, PhD

The PRO-CARE Group (P.H., A.W.), School of Health and Society, Kristianstad University, Kristianstad; Department of Health Care Sciences and Palliative Research Centre (A.A., K.Å.), Ersta Sköndal University College; Capio Palliative Care Unit (A.A.), Stockholm; Department of Health and Caring Science (K.Å.), Linnaeus University; Kalmar County Hospital (K.Å.), Kalmar, Sweden

Abstract

Context. Previous studies have supported the psychometric properties of the 22-Item Zarit Burden Interview (ZBI-22) scale among family caregivers of people with various disorders, including Parkinson's disease (PD). However, its short forms have not been psychometrically tested among PD family caregivers, and available psychometric analyses have not accounted for the ordinal nature of item-level data.

Objectives. To assess the psychometric properties of the ZBI-22 and its short forms among family caregivers of people with PD, while taking account for the ordinal nature of data.

Methods. Cross-sectional postal survey ZBI-22 data from 66 family caregiver members (59% women; mean age 69.6 years) of a local Swedish PD society branch were analyzed according to classical test theory methods based on polychoric/polyserial correlations.

Results. Missing item responses were $\leq 5\%$. Corrected item-total correlations were ≥ 0.42 and floor/ceiling effects were $<20\%$, besides for the briefest (4- and 1-item) short forms (20% and 40% floor effects, respectively). Reliability was good for all scales (ordinal alpha 0.89–0.95). External construct validity was in general accordance with a priori expectations. Short forms demonstrated good criterion-related validity (r_s 0.87–0.99) and discriminative ability (area under the curve, 0.91–0.98) relative to the full ZBI-22.

Conclusion. This study provides support for the reliability and validity of the ZBI-22 and its various short forms for use among PD family caregivers. In studies where caregiver burden is a central outcome, either ZBI-22 or ZBI-12 is suggested for use; other short forms can be used when caregiver burden is of less central focus or for clinical screening. *J Pain Symptom Manage* 2017;53:272–278. © 2016 American Academy of Hospice and Palliative Medicine. Published by Elsevier Inc. All rights reserved.

Key Words

Burden, family caregivers, Parkinson's disease, psychometrics, validation

Introduction

People with long-term progressive neurological disorders such as Parkinson's disease (PD) may be in great need of assistance, and family caregivers often provide practical as well as medical, emotional, and social supports.^{1,2} Family caregiver perspectives have therefore been increasingly recognized, and a central caregiver outcome is perceived burden,³ most commonly assessed using the generic 22-Item Zarit

Burden Interview (ZBI-22).^{2,4,5} In addition to the original ZBI-22,^{6,7} a number of short forms have been proposed as a means of simplifying its use, target the scale to different settings and needs, and to minimize respondent burden.^{4,8,9}

A Spanish study assessed the ZBI-22 regarding its basic psychometric performance among family caregivers of people with PD.¹⁰ The ZBI-22 was found to exhibit acceptable psychometric properties with, for example, 90% computable total scores, 3% floor/

Address correspondence to: Peter Hagell, PhD, The PRO-CARE Group, School of Health and Society, Kristianstad University, SE-291 88 Kristianstad, Sweden. E-mail: Peter.Hagell@hkr.se

Accepted for publication: September 6, 2016.

ceiling effects, item-total correlations between 0.31 and 0.78, a reliability estimate (coefficient alpha) of 0.93, and correlations between -0.33 and -0.50 with physical and -0.52 to -0.64 with mental health domains of the 36-Item Short Form Health Survey (SF-36).¹⁰ As with most classical test theory (CTT)-based psychometric studies, the core analyses (e.g., item-total correlations and coefficient alpha) were parametric and did therefore not account for the fact that item-level data are ordinal. However, polychoric-based correlation and reliability estimates take the ordinal nature of data into account and diminish the bias associated with applying parametric procedures with ordinal data.^{11–14} Furthermore, in the case of the ZBI, psychometric evidence regarding its various short forms is relatively limited⁴ and appears absent among PD family caregivers, which is limiting because short forms may provide more feasible clinical tools.

The aim of this study was to reassess the psychometric properties of the ZBI-22 and its short forms among Swedish family caregivers of people with PD using CTT approaches that take the ordinal nature of data into account.

Methods

The study was conducted in accordance with the Declaration of Helsinki and reviewed by the local

institutional ethics advisory committee. All participants provided informed consent.

Data were taken from an anonymous cross-sectional postal survey to all 107 registered family caregiver members of a local south Swedish branch of the Swedish PD Society.¹⁵

Instruments

Zarit Burden Interview

The ZBI-22 (Table 1) consists of 22 items with five ordered frequency-related response categories scored 0 (never) to 4 (nearly always), except for the final item, which has five ordered intensity-related response categories (0 = not at all; 4 = extremely). All 22 items are used to calculate a total score that can range between 0 and 88 (88 = more burden). A total score of 21 has been suggested as a burden cutpoint.^{4,16}

In addition, several short forms of the ZBI-22 have been proposed,^{4,8,9} including 12-, 8-, 7-, 6-, and 4-item versions and a single-item (ZBI-1) version to be used as a screening tool (Table 1). These are scored according to the same principle as the original ZBI-22, yielding total scores that range between 0 and 48 (ZBI-12), 32 (ZBI-8), 28 (ZBI-7), 24 (ZBI-6), 16 (ZBI-4), and 4 (ZBI-1) with higher scores representing more burden. Here, we used the Swedish version of the ZBI-22 (www.mapi-trust.org) but did also compute scores according to the six short forms, as embedded in the ZBI-22.

Table 1
The Zarit Burden Interview^a

ZBI-22 Items		Short-Form Items ^b					
No.	Content (Abridged)	ZBI-12	ZBI-8	ZBI-7	ZBI-6	ZBI-4	ZBI-1
1	Relative asks for more help than needed						
2	Not enough time for yourself	X		X	X	X	
3	Stressed between caring and other responsibilities	X		X	X	X	
4	Embarrassed over behaviors		X				
5	Angry when around your relative	X	X				
6	Relative affects your relationship with others	X	X	X	X		
7	Afraid of what the future holds for relative						
8	Your relative is dependent on you						
9	Strained when around your relative	X	X	X	X	X	
10	Your health has suffered because of caring	X		X	X		
11	Insufficient privacy because of your relative	X					
12	Social life has suffered because of caring	X	X				
13	Uncomfortable having friends over because of relative		X				
14	Relative seems to expect you to take care of him or her, as if you were the only one to depend on						
15	Do not have enough money to care for your relative						
16	Not able to take care of your relative much longer						
17	Lost control of your life since your relative's illness	X		X	X		
18	Wish you could leave the care to someone else		X				
19	Uncertain about what to do about relative	X	X			X	
20	You should be doing more for your relative	X					
21	Could do a better job caring for your relative	X					
22	Overall, how burdened do you feel in caring			X			X

ZBI-22 = 22-Item Zarit Burden Interview.

^aItem response categories: 0 = never, 1 = rarely, 2 = sometimes, 3 = quite frequently, 4 = nearly always (except Item 22: 0 = not at all, 1 = a little, 2 = moderately, 3 = quite a bit, and 4 = extremely).

^bTotal score ranges (higher = more burden): ZBI-12 = 0–48; ZBI-8 = 0–32; ZBI-7 = 0–28; ZBI-6 = 0–24; ZBI-4 = 0–16; and ZBI-1 = 0–4.

Additional Questionnaires

The survey also included demographic questions, the SF-36 version 1¹⁷ and the Sleep section of the Nottingham Health Survey (NHP).¹⁸ The SF-36 is a generic health status questionnaire covering eight aspects of health: physical functioning, role limitations/physical, bodily pain, general health, vitality, social functioning, role limitations/emotional, and mental health. Each of the eight SF-36 scales yields a total score of 0–100 (100 = better health). The five-item sleep section of the NHP (NHP-sleep) yields a total score of 0–100 (100 = worse sleep) and was used as an index of sleep problems.

Respondents also provided information on the person with PD who they cared for, including PD duration and caregiver-perceived PD severity according to the five-grade single-item PD Activities of Daily Living Scale (PADLS).¹⁹

Analyses

Data were analyzed using R version 3.2.2 (psych package version 1.5.8; www.r-project.org) and Stata version 14.1 (StataCorp, College Station, TX). The general approach was not only to replicate previous psychometric studies on the ZBI-22 in PD¹⁰ and of its seven short forms among carers of people with cancer, dementia, and brain injury⁴ but also to use procedures that take the ordinal nature of item data into account.^{12,13} Specifically, we conducted analyses regarding data completeness, scaling assumptions, targeting, reliability, external construct validity, and criterion-related validity.^{20–23}

Data completeness was studied by calculating the percentage of missing data for items and total scores, which should be <10% to be considered acceptable among responders. Scaling assumptions according to CTT criteria regarding the legitimacy of summing item scores into a total score assume that each item should contribute substantially to the total score (as indicated by corrected item-total correlations >0.3) and that items within the scale represent a common variable (as suggested by corrected item-total correlations exceeding 0.3–0.4). Item-total correlations were computed based on polychoric correlations (accounting for the ordinal nature of item-level data). In addition, we report traditional Pearson-based statistics to allow for comparisons with previous studies.

Score distributions, skewness, and floor and ceiling effects were assessed as indices of targeting, that is, how well scale scores accord with the range of burden in the sample. A well-targeted scale should have an average score close to the scale midpoint and span most of its potential range, without excess skewness (preferable between –1 and +1). Floor and ceiling

effects refer to the proportions of participants with the lowest (floor) and highest (ceiling) possible scores, respectively. Up to 15%–20% floor/ceiling effects can be considered acceptable.

Reliability was assessed by the polychoric-based ordinal version of coefficient alpha.^{12,13} The influence on alpha of deleting each item one at a time was also explored; an increased coefficient after item deletion suggests problems, for example, with construct conceptualization or multidimensionality. In addition to ordinal alpha, the traditional coefficient alpha was also calculated.²⁴ Alpha coefficients above 0.7 are acceptable. The SEM was calculated ($SD \times \sqrt{1 - \text{reliability}}$) as an estimator of score precision and expressed in the same unit as the respective scales, as well as a percentage of the highest possible total scores (to facilitate comparisons between various ZBI versions). SEM was calculated based on ordinal and traditional alphas.

External construct validity was assessed by Spearman correlations between ZBI scores and SF-36 and NHP-sleep scores, as well as between ZBI scores and caregiver-reported patient PD duration and PADLS scores. Based on Martínez-Martín et al.,¹⁰ moderate correlations (around 0.3–0.6) were expected between ZBI scores and patient PD duration and PADLS scores. Similar correlations were expected with NHP-sleep and SF-36 scores but with generally stronger correlations for aspects of mental (i.e., vitality, social functioning, role limitations/emotional, and mental health) than physical health according to the SF-36. Family caregivers who reported to have a disease of their own were expected to experience greater burden than those who did not, according to Mann-Whitney *U* tests. Effect sizes (*r*) of differences according to the Mann-Whitney *U* tests were calculated;^{25,26} medium effect sizes (i.e., around 0.3)²⁷ were expected.

Criterion-related validity of the ZBI short forms was assessed by Spearman correlations between short-form total scores and the ZBI-22 total score (here considered as the gold standard), which should be >0.7 to be acceptable.⁴ The coefficient of determination (r^2) was computed to assess the amount of variance in the ZBI-22 that may be explained by the variance in the respective short forms.

The discriminative performance of the short forms relative to the cutpoint of 21 on the ZBI-22 was assessed and compared using receiver operating characteristic (ROC) curves.⁴ The closer the ROC curve is to the upper left corner (100% sensitivity and 100% specificity) of an ROC plot, the higher the accuracy. The area under the curve (AUC) represents the ability to correctly classify people with and without burden according to the ZBI-22 cutpoint (i.e., the discriminative ability). AUCs can range between 0 and 1; an AUC

< 0.5 indicates performance worse than chance, whereas an AUC of 1 indicates perfect discrimination. AUCs ≥ 0.7 and >0.9 are considered acceptable and high, respectively.²⁸ Cutpoints for the ZBI short forms were estimated based on the Youden index.²⁸

Results

Participants ($n = 66$; 61% response rate) had a mean age of 69.6 (8.2), were mainly women ($n = 46$; 70%), and 39 (59%) reported having a disease on their own. They had been caring for a person with PD for up to 18 (median 3; q1–q3 2–5) years, and they spent a median (q1–q3; minimum–maximum) of 22.5 (10–75; 1–168) hours/week caring (Table 2).

Table 2
Characteristics of Family Caregivers ($n = 66$) and the Person With PD Who They Cared for

Characteristics	Findings
Family caregivers	
Age (years), mean (SD; minimum–maximum)	69.6 (8.2; 44–86)
Female gender, n (%)	46 (70)
Relationship to person with PD, n (%)	
Partner	63 (95)
Child	1 (2)
Other	2 (3)
Highest education, n (%)	
Primary	26 (40)
Secondary	20 (30)
Tertiary	20 (30)
Retired, n (%)	51 (77)
Own disease, n (%)	39 (59)
Care duration (years), md (q1–q3; minimum–maximum)	3 (2–5; 0–18)
Current caring (hours/week), md (q1–q3; minimum–maximum)	22.5 (10–75; 1–168)
SF-36, md (q1–q3; minimum–maximum) ^a	
Physical functioning	80 (70–90; 10–100)
Role—physical	100 (25–100; 0–100)
Bodily pain	62 (41–84; 0–100)
General health	67 (45–80; 10–100)
Vitality	57.5 (33–75; 0–100)
Social functioning	75 (50–100; 12.5–100)
Role—emotional	100 (33–100; 0–100)
Mental health	72 (48–88; 12–100)
NHP-sleep, md (q1–q3; minimum–maximum) ^b	20 (0–60; 0–100)
People with PD ^c	
Men/women, n (%)	45 (68)/21 (32)
Age (years), mean (SD; minimum–maximum)	71.5 (7.6; 54–87)
PD duration (years), mean (SD; minimum–maximum)	9.3 (6.0; 0–33)
PADLS, md (q1–q3; minimum–maximum) ^d	2.5 (2–4; 1–5)

PD = Parkinson's disease; md = median; SF-36 = the 36-Item Short Form Health Survey; NHP = Nottingham Health Profile; PADLS = Parkinson's Activity of Daily Living Scale.

^aScore range, 0–100 (100 = better).

^bScore range, 0–100 (100 = worse).

^cAs reported by carers.

^dScore range, 1–5 (1 = no difficulties with day-to-day activities; 5 = extreme difficulties with day-to-day activities).

Data completeness was good with $\leq 5\%$ missing item-level data and $\geq 92\%$ computable total scores across all ZBI forms (Table 3). There was also support for the scaling assumptions as assessed from corrected item-total correlations. Targeting was generally satisfactory with skewness values ≤ 0.36 across the various ZBI forms, although floor effects increased with decreasing numbers of items and failed to meet a priori criteria for the ZBI-4 and ZBI-1. Average total scores were consistently below scale midpoints, although the midpoint was within the third quartile and 1 SD for all ZBI formats but the ZBI-8. Reliability was good with ordinal alpha ranging 0.89–0.95 across the various ZBI formats. Apart from the ZBI-7 (Item 6; traditional coefficient alpha) and ZBI-6 (Item 6; ordinal and traditional coefficient alpha), reliability did not increase after item deletion. The SEM corresponded to 4.4%–9% of possible total scores and tended to increase by decreasing numbers of items per scale (Table 3).

External construct validity correlations were in general accordance with expectations and similar across the various ZBI forms (Table 4). As expected, carers who reported having a disease of their own had significantly higher ZBI scores ($P \leq 0.022$) than those who did not report a disease of their own, with effect sizes of 0.29–0.38. All ZBI short forms exhibited acceptable levels of criterion-related validity (Table 4), correlating ≥ 0.87 with the full ZBI-22 criterion (r^2 0.76–0.98).

Optimal cutpoints according to the Youden index relative to the gold standard cutpoint of 21 on the ZBI-22 are shown in Table 5. The ROC AUCs associated with the estimated cutpoints indicated good discriminative ability for all ZBI short forms.

Discussion

This study assessed the psychometric properties of the ZBI-22 and its proposed short forms among family caregivers of people with PD. In all instances, we found general support for the psychometric properties of the ZBI. This is in accordance with previous observations with the ZBI-22 among Spanish PD family caregivers.¹⁰ However, here we expanded those observations to include also the various ZBI short forms and used methodologies that take account of the ordinal nature of data.

Data completeness was good, and there was support for the legitimacy of summing item scores into total ZBI scores. Although ZBI scores were somewhat skewed toward lower levels, targeting was generally good as scale midpoints were within the interquartile ranges and SDs in all instances but for the ZBI-8. Furthermore, there were minimal ceiling effects, and floor effects were acceptable for all ZBI formats but

Table 3
Data Completeness, Scaling Assumptions, Targeting, and Reliability of the ZBI Among Carers of People With PD

Property	ZBI-22	ZBI-12	ZBI-8	ZBI-7	ZBI-6	ZBI-4	ZBI-1
Data completeness							
Missing item responses (minimum–maximum %) ^a	0–5	0–3	0–3	0–5	0–2	0–2	5
Computable scale scores (%) ^b	92	95	97	94	97	97	95
Scaling assumptions							
Corrected polyserial item-total correlation (minimum–maximum) ^c	0.42–0.80	0.49–0.76	0.56–0.78	0.64–0.89	0.63–0.87	0.73–0.89	NA
Corrected Pearson item-total correlation (minimum–maximum) ^c	0.36–0.88	0.47–0.85	0.47–0.75	0.55–0.86	0.54–0.84	0.68–0.85	NA
Targeting							
Possible score range (midpoint)	0–88 (44)	0–48 (24)	0–32 (16)	0–28 (14)	0–24 (12)	0–16 (8)	0–4 (2)
Mean (SD) score ^d	28.3 (18.0)	15.3 (11.0)	7.7 (6.2)	9.4 (7.1)	7.9 (6.0)	5.5 (4.2)	1.5 (1.3)
Median (q1–q3) score ^d	26 (14–46)	14 (6–25)	7 (2–13)	8.5 (4–15)	8 (3–12.5)	5 (2–9)	1 (0–3)
Minimum–maximum score ^e	0–60	0–37	0–20	0–23	0–20	0–13	0–4
Floor/ceiling effects (%) ^f	2/0	5/0	14/0	12/0	17/0	20/0	30/6
Skewness ^g	0.27	0.30	0.36	0.26	0.23	0.14	0.28
Reliability							
Ordinal α ^h	0.95	0.91	0.89	0.94	0.92	0.91	NA
Ordinal α when item deleted (minimum–maximum) ⁱ	0.94–0.95	0.90–0.91	0.86–0.88	0.92–0.94	0.90–0.93	0.85–0.90	NA
SEM, ordinal α based (% of total score) ^j	4.0 (4.6)	3.3 (6.9)	2.0 (6.4)	1.7 (6.2)	1.7 (7.1)	1.3 (7.8)	NA
Coefficient α ^h	0.95	0.93	0.88	0.92	0.90	0.88	NA
Coefficient α when item deleted (minimum–maximum) ⁱ	0.95–0.95	0.91–0.93	0.85–0.88	0.90–0.93	0.87–0.91	0.81–0.87	NA
SEM (% of total score) ^j	4.0 (4.6)	2.9 (6.0)	2.1 (6.7)	2.0 (7.2)	1.9 (7.9)	1.4 (9.0)	NA

ZBI = Zarit Burden Interview; q1–q3 = first to third quartile (25th–75th percentile); NA = not applicable.

^aShould be <10%.

^bShould be close to 100%.

^cShould be >0.3 to support summation of raw item scores and >0.3–0.4 to support a single underlying variable.

^dShould be close to scale midpoint.

^eShould span most of the scale's score range.

^fShould be <15%–20%.

^gShould be between –1 and +1.

^hShould be ≥ 0.80 .

ⁱShould not increase compared with α for the total score.

^jShould be less than half of the total score SD.

Table 4
External Construct Validity of the ZBI Among Carers of People With PD

Property	ZBI-22	ZBI-12	ZBI-8	ZBI-7	ZBI-6	ZBI-4	ZBI-1
Construct validity correlations ^a							
Patient PD duration ^b	0.43	0.41	0.31	0.39	0.40	0.35	0.30
Patient PADLS score ^b	0.50	0.42	0.43	0.49	0.43	0.42	0.45
Physical Functioning (SF-36) ^c	–0.36	–0.34	–0.30	–0.39	–0.38	–0.37	–0.32
Role limitations/physical (SF-36) ^c	–0.42	–0.42	–0.40	–0.46	–0.43	–0.48	–0.47
Bodily pain (SF-36) ^c	–0.53	–0.48	–0.51	–0.52	–0.53	–0.51	–0.49
General health (SF-36) ^c	–0.56	–0.55	–0.54	–0.55	–0.55	–0.56	–0.53
Vitality (SF-36) ^c	–0.61	–0.58	–0.57	–0.62	–0.61	–0.60	–0.60
Social functioning (SF-36) ^c	–0.69	–0.66	–0.65	–0.69	–0.68	–0.67	–0.65
Role limitations/emotional (SF-36) ^c	–0.49	–0.48	–0.44	–0.51	–0.49	–0.51	–0.44
Mental health (SF-36) ^c	–0.69	–0.69	–0.64	–0.71	–0.72	–0.69	–0.66
NHP sleep	0.46	0.45	0.44	0.45	0.45	0.48	0.38
Known groups' validity							
Carers without own disease, md (q1–q3)	16.5 (8–34)	9 (2–18)	3 (1–10)	4 (1–12)	4 (0–9)	2 (0–6)	1 (0–2)
Carers with own disease, md (q1–q3)	30 (17–49)	17 (10–28)	8 (5–14)	11 (6–18)	9 (5–15)	7 (4–10)	2 (1–3)
<i>P</i>	0.011	0.008	0.022	0.010	0.005	0.003	0.011
Effect size ^d	0.32	0.34	0.29	0.33	0.35	0.38	0.32
Criterion-related validity							
ZBI-22 correlation ^{a,e}	NA	0.99	0.93	0.97	0.96	0.95	0.87
ZBI-22 coefficient of determination (r^2)	NA	0.98	0.86	0.94	0.92	0.90	0.76

ZBI = Zarit Burden Interview; PD = Parkinson's disease; PADLS = Parkinson's Activity of Daily Living Scale; SF-36 = the 36-Item Short Form Health Survey; NHP = Nottingham Health Profile; md = median; q1–q3 = first to third quartile (25th–75th percentile).

^aSpearman correlations.

^bCaregiver reported.

^cNegative correlations because of opposite scoring directions.

^d0.1 = small, 0.3 = medium, and 0.8 = large.

^eShould be >0.7.

Table 5
Cutpoint Estimations and ROC Curve Analyses of ZBI Short Forms Relative to the Suggested ZBI-22 Cutpoint of 21 Among Carers of People With PD

Property	ZBI-12	ZBI-8	ZBI-7	ZBI-6	ZBI-4	ZBI-1
Cutpoint	11	5	7	7	4	1
Youden index (SE)	0.96 (0.04)	0.90 (0.06)	0.91 (0.05)	0.91 (0.05)	0.84 (0.07)	0.82 (0.06)
Area under ROC curve at cutpoint ^a	0.98	0.95	0.96	0.96	0.92	0.91

ROC = receiver operating characteristic; ZBI = Zarit Burden Interview; PD = Parkinson's disease.

^aValues ≥ 0.7 and > 0.9 are considered acceptable and high, respectively.

the briefest. This is expected because shorter forms have fewer items and therefore a more restricted outcome range. This was also reflected by generally reduced precision of scores with decreasing numbers of items, as manifested by increasing SEM values (expressed as a percentage of the total score range). However, reliability was good, although coefficient alpha tended to increase when Item 6 was deleted from the ZBI-6 and ZBI-7. This may suggest that this item does not work in harmony with the other items in those ZBI forms.²¹ However, given the relatively small increases and limited sample size, these observations should not be overinterpreted but rather serve as a prompt in future studies.

Associations with other variables were as expected and generally comparable across the various ZBI forms, suggesting similar external construct validity. Similarly, correlations between ZBI short-forms and the ZBI-22 were high and comparable across short-forms, although it was significantly higher for the ZBI-12 than for the ZBI-8, -4 and -1 and lower for the ZBI-1 than for the ZBI-12, -7 and -6 (data available on request). Analyses of cutpoints of the short forms relative to that of the ZBI-22 rendered cutpoints that were very close to those previously identified among family caregivers of people with cancer, brain injury, and dementia.⁴ In fact, identified cutpoints were identical for the ZBI-7, ZBI-4, and ZBI-1 and differed by only one point for the other short forms as compared with the cutpoints suggested by Higginson et al.,⁴ who identified 11 as the optimal cut-point for the ZBI-12 and 6 as the optimal cut-point for the ZBI-8 and -6. Furthermore, the discriminative performance of the short forms was high, with AUCs corresponding closely with those reported by Higginson et al.⁴

Our sample size is relatively small, and participants consisted of members of a national PD society, which may limit the study's representativeness. However, samples of this and smaller sizes have been found sufficient in producing stable reliability and validity indices according to CTT.²⁹ Nevertheless, evaluations based on larger samples are warranted to allow for firmer conclusions, particularly regarding dimensionality and response category functioning.^{30,31} However, to the best of our knowledge, this is the first study to address the psychometric performance of the ZBI

and its short forms using CTT methods appropriate for ordinal level data.

Our observations provide initial support for the psychometric properties of the ZBI and its various short forms for use among family caregivers of people with PD. The demonstration of the appropriateness of the ZBI short forms opens up for more flexible caregiver burden assessment. For example, in studies where caregiver burden is a primary or central outcome, investigators could use either the full ZBI-22 or the ZBI-12, whereas the ZBI-4 and ZBI-1 can be useful in circumstances where caregiver burden is a less central variable or for clinical screening purposes. These results are considered valuable given the importance to take family caregivers' perspectives into account in the holistic management of long-term neurological disorders such as PD. The ZBI short forms can open up for this without adding unnecessary respondent burden.

Disclosures and Acknowledgments

The authors thank the participants for their collaboration, the Swedish Parkinson Association for assistance with data collection, and the Swedish House in Kavalla, Greece, for inspiration. The study was supported by the Swedish Research Council, Swedish Parkinson Foundation, Swedish Parkinson Academy, Capho Palliative Care Unit, Stockholm, Sweden, Kristianstad University, Kristianstad, Sweden, and Linnaeus University, Kalmar, Sweden.

The authors declare no conflicts of interest.

References

1. Greenwell K, Gray WK, van Wersch A, van Schaik P, Walker R. Predictors of the psychosocial impact of being a carer of people living with Parkinson's disease: a systematic review. *Parkinsonism Relat Disord* 2015;21:1–11.
2. Martínez Martín P, Rodríguez-Blázquez C, Forjaz MJ. Quality of life and burden in caregivers for patients with Parkinson's disease: concepts, assessment and related factors. *Expert Rev Pharmacoecon Outcomes Res* 2012;12:221–230.
3. Adelman RD, Tmanova LL, Delgado D, Dion S, Lachs MS. Caregiver burden: a clinical review. *JAMA* 2014;311:1052–1060.

4. Higginson IJ, Gao W, Jackson D, Murray J, Harding R. Short-form Zarit Caregiver Burden Interviews were valid in advanced conditions. *J Clin Epidemiol* 2010;63:535–542.
5. Mosquera I, Vergara I, Larranaga I, et al. Measuring the impact of informal elderly caregiving: a systematic review of tools. *Qual Life Res* 2016;25:1059–1092.
6. Zarit SH, Anthony CR, Boutselis M. Interventions with care givers of dementia patients: comparison of two approaches. *Psychol Aging* 1987;2:225–232.
7. Zarit SH, Reeve KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist* 1980;20:649–655.
8. Arai Y, Tamiya N, Yano E. The short version of the Japanese version of the Zarit Caregiver Burden Interview (J-ZBI_8): its reliability and validity [in Japanese]. *Nihon Ronen Igakkai Zasshi* 2003;40:497–503.
9. Bedard M, Molloy DW, Squire L, et al. The Zarit Burden Interview: a new short version and screening version. *Gerontologist* 2001;41:652–657.
10. Martinez-Martin P, Forjaz MJ, Frades-Payo B, et al. Caregiver burden in Parkinson's disease. *Mov Disord* 2007;22:924–931. quiz 1060.
11. Carroll JB. The nature of the data, or how to choose a correlation coefficient. *Psychometrika* 1961;26:347–372.
12. Gadermann AM, Guhn M, Zumbo BD. Estimating ordinal reliability for Likert-type and ordinal item response data: a conceptual, empirical, and practical guide. *Pract Assess Res Eval* 2012;17:1–13.
13. Zumbo BD, Gadermann AM, Zeisser C. Ordinal versions of coefficients alpha and theta for Likert rating scales. *J Mod Appl Stat Methods* 2007;6:21–29.
14. Olsson U. Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* 1979;44:443–460.
15. Olsson Y, Claren L, Alvariza A, Arestedt K, Hagell P. Health and social service access among family caregivers of people with Parkinson's disease. *J Parkinsons Dis* 2016;6:581–587.
16. Cummings JL, Frank JC, Cherry D, et al. Guidelines for managing Alzheimer's disease: part I. *Am Fam Physician* 2002;65:2263–2272.
17. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473–483.
18. Hunt SM, McKenna SP, McEwen J, et al. A quantitative approach to perceived health status: a validation study. *J Epidemiol Community Health* 1980;34:281–286.
19. Hobson JP, Edwards NI, Meara RJ. The Parkinson's Disease Activities of Daily Living Scale: a new simple and brief subjective measure of disability in Parkinson's disease. *Clin Rehabil* 2001;15:241–246.
20. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess* 2009;13:iii. ix-x, 1–177.
21. Nunnally JC, Bernstein IH. *Psychometric theory*, 3rd ed. New York: McGraw-Hill, Inc., 1994.
22. Saris-Baglama RN, Dewey CJ, Chisholm GB, et al. *SF health outcomes™ scoring software user's guide*. Lincoln, RI: Quality Metric, Inc., 2004.
23. Ware JE Jr, Gandek B. Methods for testing data quality, scaling assumptions, and reliability: the IQOLA Project approach. *International Quality of Life Assessment*. *J Clin Epidemiol* 1998;51:945–952.
24. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
25. Field A, Miles J, Field Z. *Discovering statistics using R*. London: Sage Publications Ltd., 2012.
26. Rosenthal R. *Meta-analytic procedures for social research*, 2nd ed. Newbury Park, CA: Sage, 1991.
27. Cohen J. A power primer. *Psychol Bull* 1992;112:155–159.
28. Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003;29:1043–1051.
29. Hobart JC, Cano SJ, Warner TT, Thompson AJ. What sample sizes for reliability and validity studies in neurology? *J Neurol* 2012;259:2681–2694.
30. Ballesteros J, Santos B, Gonzalez-Fraile E, et al. Unidimensional 12-item Zarit Caregiver Burden Interview for the assessment of dementia caregivers' burden obtained by item response theory. *Value Health* 2012;15:1141–1147.
31. Siegert RJ, Jackson DM, Tennant A, Turner-Stokes L. Factor analysis and Rasch analysis of the Zarit Burden Interview for acquired brain injury carer research. *J Rehabil Med* 2010;42:302–309.