Development and Validation of a Multidimensional Short Version Zarit Burden Interview (ZBI-9) for Caregivers of Persons With Cognitive Impairment

Lim Zhi-Xiang, BSocSci,*
Wee Shiong Lim, MBBS, MRCP (UK), MMed (Int Med), MHPEd, AGSF,
FAMS,†‡ and Ee-Yuee Chan, RN, MN, MSc, PhD§‖

Background: There is a lack of appreciation of the full dimensionality of the original 22-item Zarit Burden Interview (ZBI) in the development of short versions. Existing short versions are premised upon a 1-factor or 2-factor structure or statistical techniques for item selection. Thus, there is a need for ZBI short versions that considers the multidimensional constructs of role strain, personal strain, and worry about performance (WaP) during item selection to provide a more holistic and comprehensive evaluation.

Purpose: To develop and validate a short version of ZBI through a combined quantitative and qualitative approach that incorporates the validated 4-factor structure of role strain demands; role strain-control; personal strain, and WaP.

Patients: We studied 202 caregivers of patients with dementia (84.2%) or mild cognitive impairment (15.8%) attending a memory clinic in Singapore.

Methods: Confirmatory factor analysis and qualitative considerations from expert consensus were used for item selection. Confirmatory factor analysis fit statistics support the 4-factor structure. The 9-item ZBI-9 showed good internal consistency (Cronbach's $\alpha = 0.87$) and convergent validity with anxiety and depression (Pearson correlation: Hospital Anxiety and Depression sub-scales, $r \geq 0.56$, $P < 0.001$; ZBI-22, $r = 0.95$, $P < 0.001$). Using a cut-off score of ≥ 13, ZBI-9 displayed good discriminatory ability for depressive symptoms (area under curve $= 0.79$, $P < 0.001$; sensitivity $= 70\%$, specificity $= 75\%$). The ZBI-9 also displayed comparable performance to the 22-item full version and three 12-item short versions.

Conclusion: The ZBI-9 is a multidimensional short-version assessment tool for caregivers of persons with dementia and mild cognitive impairment that is reliable, valid, and discriminates depressive symptoms.

Key Words: Alzheimer disease, caregiver burden, Zarit Burden interview, factor analysis, short version

(Alzheimer Dis Assoc Disord 2023;37:59–65)

The Zarit Burden Interview (ZBI) 22-item is the most widely used tool for the assessment of burden among caregivers of persons with dementia. It has been used in clinical and research settings to evaluate and measure caregiver burden.[1] The ZBI has been translated into multiple languages and has shown to be reliable and valid in diverse settings.[2,3] Short versions that aim to assess the full scope of the burden construct, albeit with fewer items, have also been developed to reduce the burden of its administration.[4] These are appreciated in clinical settings as they enable healthcare providers to perform the assessment in a shorter time and reduce the inconvenience for caregivers, who are often time-pressed. Such benefits are also desirable in research studies where participants would likely be required to complete other assessment tests in addition to the ZBI.

Short versions should be distinguished from screening versions (typically comprised of fewer than 5 items) that are designed for rapid screening of caregivers.[5,6] Unlike brief screening tools, short ZBI versions should seek to retain the factor structure of the original burden construct. Recent studies that support the multidimensionality of ZBI beyond a unidimensional score highlight the importance of considering the full dimensionality of the original 22-item ZBI during the development of short versions.[7–9] However, several short versions omit this consideration, relying on expert opinion[6] or solely on statistical techniques for item selection.[10] While this can optimize psychometric properties and reduce bias during the selection process, it can also lead to a loss of content validity;[11] for instance, Ballesteros[10] described a unidimensional 12-item short version which was empirically developed based on item response theory without consideration of prior ZBI factor structures.

Another major challenge in the development of short versions is the lack of consensus about the essential factors that underpin the multidimensionality of ZBI. Earlier studies in Caucasian populations defer to the canonical dual-factor structure of ZBI defined by role strain and personal strain.[12] In contrast, more recent studies in diverse Asian and Caucasian populations consistently highlighted a distinct dimension of self-appraisal burden manifested by the desire to 'do more' (item 20) or 'do better' (item 21) in the caregiving role.[7,13] This factor represents a conceptual continuum of worry about caregiving performance (WaP), ranging from "inadequacy" and "worry" through to more severe degrees of "self-criticism" and "guilt".[14] WaP is a distinct dimension of burden among adult children caregivers, and unlike role and personal strain, it has been endorsed as early as the stage of mild cognitive impairment (MCI).[9] The viability of ZBI multidimensionality being premised on the 3 key dimensions of role strain, personal strain and WaP was corroborated by a recent confirmatory factor analysis, which revealed that 3 or 4 factor models were superior to 1-factor (total score) and 2-factor (role and personal strain) models.[11]

This lack of appreciation of the full dimensionality of the original 22-item ZBI is paralleled in the development of abridged ZBI versions. Existing short versions are premised upon a 1-factor or 2-factor structure, being informed either by the dual-factor structure (role and personal strain)[4] or statistical techniques for item selection.[10] There is thus a need for ZBI short versions that considers the multidimensional constructs of role strain, personal strain, and WaP during item selection to provide a more holistic and comprehensive evaluation. For instance, factors premised upon items 20 and 21, which were previously labeled as 'role strain', would have been more accurately classified as WaP burden.[15] This is not a moot point, as caregivers with an identical ZBI score may differ in the type of burden that they are experiencing as represented by the factors. In support of this, a recent validation study of the Cantonese version of the 12-item ZBI developed by Bédard[4] reported a 3-factor model comprised role strain, negative emotion, and self-criticism, which was supported by confirmatory factor analysis.[16] Moreover, there is increasing appreciation within the field that item selection is a complex process that would benefit from both quantitative and qualitative evaluations so that unique contextual considerations can be incorporated.[17] In particular, expert opinions from clinicians who routinely use the ZBI will help provide a qualitative understanding of how participants respond to individual items and ensure that the dimensions of caregiver burden are sufficiently represented.

The aim of this study is thus to develop and validate a short version through a combined quantitative and qualitative approach that incorporates the multidimensionality of the ZBI in a predominantly Chinese population of caregivers of persons with dementia or mild cognitive impairment attending a tertiary memory clinic in Singapore.

METHODS

Setting, Participants, and Instruments

We prospectively recruited patient-caregiver dyads of community-dwelling older adults and their family caregivers attending the memory clinic, Department of Geriatric Medicine, in Tan Tock Seng Hospital, Singapore, from September 2018 to October 2019. We included patients who were: (a) 65 years and above, (b) diagnosed with dementia or MCI, and (c) living in the community. We defined family caregivers as family members aged 21 years and above who was most involved in the provision of daily care and familiar with the patient's social and medical status. In the case of MCI patients, caregivers may not be heavily involved in the assistance with activities of daily living, but they still assist with arranging for, accompanying, and/or supervising medical appointments, social activities, and healthy lifestyle-related tasks (such as physical activity and nutrition).

Caregivers who were not family members or were unable to understand English or Mandarin were excluded. Among 204 eligible caregiver-patient dyads, 2 caregivers did not complete the questionnaire, yielding 202 dyads in the final analysis. Ethical approval was obtained from the Institutional Review Board of the National Healthcare Group.

We administered the survey in English or Mandarin. Survey items included socio-demographic characteristics such as age, gender, education level, marital status, work status, presence of domestic helper, living arrangement (living with or apart from the patient), and relationship to the patient.

The 22-item ZBI was used in this study. Each item is rated on a 5-point Likert scale ranging from 0 = "never" to 4 = "always". Item scores can be summated to provide a total score ranging from 0 to 88, where higher scores represent higher levels of burden. We previously validated the 4-factor structure of ZBI-22 that accounted for 62.2% of the variance, namely: Factor 1: demands of care and social impact on the caregiver (role strain 1, RS1); Factor 2: confidence or control over the situation (role strain 2, RS2); Factor 3: psychological impact on the caregiver (personal strain, PS); and Factor 4: worry about caregiving performance (WaP).[18] The 4-factor structure has been shown to be superior to the 2-factor and 3-factor structures.[11]

The Hospital Anxiety and Depression Scale (HADS) was used to assess symptoms of anxiety and depression in caregivers.[19] This scale has been used in caregivers and the community setting.[20] The 14-item scale contains 2 7-item subscale measuring anxiety (HADS-A) and depression (HADS-D). Total scores for each subscale ranged from 0 to 21, with higher scores representing higher levels of anxiety or depression symptoms. A cut-off of ≥ 8 cut-off was used in both HADS-D and HADS-A to denote significant depressive and anxiety symptoms.[20]

Short Version Development

We followed the rigorous methodological guidelines recommended by Goetz[17] to shorten the scale while maintaining the validity and psychometric properties of the original tool. We employed a 2-phase iterative approach, comprising quantitative analysis [confirmatory factor analysis (CFA)] followed by qualitative item selection by an expert panel.

In the quantitative phase, we first performed CFA to test the fit of the 4-factor model as specified by Cheah et al.[18] We additionally tested the fit of the 1-factor, 2-factor and 3-factor models and compared it with the 4-factor model. The weighted least squares means and variance estimation method was used for the CFA. As the χ^2 statistic is sensitive to sample size, a family of fit indices was reported to evaluate model fit. The following fit indices were used: χ^2 statistic, with $P > 0.05$; root mean square error of approximation (RMSEA) and standardized root mean square error of approximation (SRMR), with values of <0.05 and <0.08 signifying good model fit and acceptable fit, respectively; and comparative fit

index (CFI) and Tucker-Lewis index (TLI), whereby > 0.95 and > 0.90 indicate a good and acceptable fit, respectively.[21] Generally, models with smaller RMSEA and SRMR values, and larger CFI and TLI values are considered to have a better fit. In addition, the fit criteria for the χ^2 statistic using factor loadings from the CFA analysis, we shortlisted items with factor loadings > 0.6 for qualitative evaluation, ensuring that at least 2 items within each factor were retained for adequate representation.[22]

Shortlisted items underwent qualitative appraisal by an expert panel comprising 3 geriatricians, a nurse clinician from the memory clinic, a nurse researcher, and a research executive. As recommended by Goetz[23] qualitative evaluation allows unique contextual considerations to be incorporated in item selection beyond psychometric properties per se. Items were assessed based on coherence with factor construct, clinical relevance, item comprehension, and redundancy.

Statistical Analysis

Descriptive and inferential analyses were conducted using IBM SPSS version 24. CFA was performed using R statistical software version 4.0.2 "lavaan" package. The significance level was set at 5%.

To assess the psychometric properties of the short version, we first performed CFA to ascertain that fit statistics support the 4-factor structure. Next, we determined corrected item-total correlation and item-factor correlation using Pearson correlation; internal consistency using Cronbach's α; convergent validity through Pearson correlation analysis with HADS-D, HADS-A, and ZBI-22. To ascertain concurrent validity, we further examined the discriminatory ability of the scale to detect significant depressive symptoms (HADS-D ≥ 8) through area under the receiver operating curve (AUC), and determined optimal cut-off, sensitivity, and specificity using the Youden index.

Lastly, we performed an analysis to compare the relative performance of our newly created short version with previously developed short versions. Using our current sample, we calculated values of internal consistency, convergent validity, and AUC for the short versions described by Bédard,[4] Ballesteros,[10] and Hébert.[23]

RESULTS

Characteristics of Caregivers and Care Recipients

Caregivers were mainly females (67.8%) with a mean age of 57 years (SD = 12.5). The majority was adult children (64.9%), followed by spouses (28.7%), and most of them lived with their care recipients (69.38%). Care recipients had a mean age of 79 years (SD = 6.2) and were predominantly females (64.4%). The majority (84.2%) were diagnosed with dementia, while 15.8% were diagnosed with mild cognitive impairment. Almost one-third (31.7%) had a live-in domestic helper to assist with caregiving duties and household chores (Table 1). The ZBI total scores were 25.28 (SD = 14.54) and factor scores were F1 = 12.20 (6.96), F2 = 3.68 (3.46), F3 = 6.13 (4.62), and F4 = 3.27 (1.88), respectively.

Item Selection and Short Version Development

Quantitative Analysis (CFA)

CFA fit indices indicate a reasonable fit for the 4-factor model [$\chi^2(203) = 292.73$, $P < 0.001$; RMSEA = 0.047; SRMR = 0.061; CFI = 0.912; and TLI = 0.900], and

TABLE 1. Characteristics of Caregivers and Care-recipients (N = 202)

Caregiver characteristics	N (%) / Mean ± SD
Age in years	57.04 ± 12.5
Sex (Female)	137 (67.8)
Marital status (Married)	149 (73.8)
Education level	
No formal education	6 (3)
Primary	17 (8.4)
Secondary	53 (26.2)
Tertiary	126 (62.4)
Ethnicity	
Chinese	186 (92.1)
Malay	7 (3.5)
Indian	9 (4.5)
Working (full/ part-time)	123 (60.9)
Relationship to care-recipient	
Spouse	58 (28.7)
Child	131 (64.9)
Others	13 (6.4)
Living with care-recipient	141 (69.8)
ZBI Burden (range 0 to 88)	25.28 ± 14.54
Factor 1 (range 0 to 36)	12.20 ± 6.96
Factor 2 (range 0 to 20)	3.68 ± 3.46
Factor 3 (range 0 to 24)	6.13 ± 4.62
Factor 4 (range 0 to 8)	3.27 ± 1.88
Depression, HADS subscale (range 0 to 21)	4 ± 3.47
Anxiety, HADS subscale (range 0 to 21)	5.28 ± 3.68
Care Recipient Characteristics	
Age, year	78.89 ± 6.16
Female	130 (64.4)
Presence of live-in domestic helper	64 (31.7)
Diagnosis	
Dementia	170 (84.2)
Mild Cognitive Impairment	32 (15.8)

HADS indicates Hospital Anxiety and Depression Scale; ZBI, Zarit Burden Interview.

provided a better fit than the other models (Table 2). χ^2 test for all models were significant. Both the 1-factor and 2-factor models met the criteria for acceptable fit for RMSEA, while only the 1-factor model met the criteria for acceptable fit for RMSEA. Both models did not meet the fit criteria for CFI and TLI. The 3 and 4-factor models had a good fit for RMSEA, and an acceptable fit for SRMR and CFI. In addition, only the 4-factor model achieved an acceptable fit. Across all 4 models, the 4-factor model presented with the lowest RMSEA and SRMR values, and the highest CFI and TLI values suggesting that the 4-factor model is the superior model. These results support the retention of WaP (factor 4) in addition to role strain (factors 1 and 2) and personal strain (factor 3).

Standardized factor loadings in CFA ranged from 0.509 to 0.984. For each factor, we shortlisted items with factor loadings > 0.6 for qualitative adjudication by an expert panel. The exception was item 22, which was excluded as it is an overall burden question. Thus, we shortlisted items 2, 3, 11, and 12 for factor 1; items 13, 16, and 17 for factor 2; items 5, 6, 9, 10, and 19 for factor 3; and items 20 and 21 for factor 4 (Table 3).

Qualitative Adjudication

During qualitative appraisal by the expert panel, items were further reduced based on established criteria (Table 3).

TABLE 2. Factor Analysis Fit Indices

References	No. items	df	χ2	RMSEA	SRMR	CFI	TLI
1 Factor (Zarit and Reever)[1]	22	209	263.61*	0.060	0.074	0.849	0.833
2 Factor (Whitlatch et al)[12]	18	134	273.056*	0.072	0.080	0.828	0.804
3 Factor (Cheah et al)[18]	22	206	305.073*	0.049	0.064	0.903	0.891
4 Factor (Cheah et al)[18]	22	203	292.73*	0.047	0.061	0.912	0.900

*$P < 0.001$.

df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; CFI = Comparative Fit Index; NNFI = Non-normed Fit Index.

For factor 1 (RS1), while all 4 shortlisted items were representative of the factor, item 11 was excluded as it was felt that the term 'privacy' was potentially ambiguous and subject to different interpretations ('personal space' vs. 'private time'). In addition, the panel agreed that items 2, 3, and 12 had a relatively higher conceptual coherence to Burden and factor 1, than item 3. Thus, items 2, 3, and 12 were selected for factor 1. For factor 2 (RS2), item 13 pertains more to the social impact of caregiving rather than confidence or control and is, therefore, less representative of factor 2. Thus, items 16 and 17 were retained. For factor 3 (PS), item 19 was excluded as it was less representative of the factor. While both items 9 and 5 could fit under the 'psychological impact', we excluded item 5 as it was appraised as covering an overly narrow aspect of 'psychological impact'. In addition, we selected item 10 over item 6 as it covered a unique and clinically relevant dimension in caregiver burden. Thus, items 9 and 10 were retained. For Factor 4, the expert panel agreed that the 2 items (items 20 and 21) measured an important factor in caregiver burden that cannot be represented by the other factors. Thus, both items of factor 4 were retained.

Psychometric Properties of ZBI-9

CFA, Item Statistics, and Internal Consistency

The final short version consisted of 9 items (ZBI-9) with a total score ranging from 0 to 36. The mean total score was 10.29 (SD = 6.32), with corresponding factor scores of F1 = 3.75 (2.86), F2 = 1.38 (1.60), F3 = 1.89 (1.81), and F4 = 3.27 (1.88). CFA showed acceptable fit for ZBI-9 ($\chi2$ (21) = 27.21, P = 0.16, RMSEA = 0.038, SRMR = 0.032, CFI = 0.98, TLI = 0.97). Standardized factor loadings ranged from 0.80 to 0.86 for factor 1; 0.68 to 0.93 for factor 2; 0.73 to 0.79 for factor 3; and 0.67 to 0.96 for factor 4.

The corrected item-total correlation was good (0.55 to 0.77), with the exception of the 2 items from factor 4 (0.26 and 0.33). Item factor correlations ranged from 0.88 to 0.91. ZBI-9 demonstrated good internal consistency (Cronbach's α = 0.87); factors 1 and 4 had the highest internal consistency among the factors (Cronbach's α = 0.87 and 0.78, respectively) (Table 4).

Convergent Validity and Concurrent Validity

In terms of convergent validity, ZBI-9 showed a moderate correlation with HADS-A (r = 0.62, $P < 0.001$) and HADS-D (r = 0.57, $P < 0.001$) and a strong correlation with ZBI-22 (r = 0.95, $P < 0.001$). Importantly, it displayed good discriminatory ability for significant depressive symptoms (AUC = 0.79, $P < 0.001$) (Table 4). Using the Youden Index, a cut-off score of ≥ 13 on the ZBI-9 would provide the best balance of sensitivity and specificity to detect significant depressive symptoms (Sensitivity: 70%, Specificity: 75%, PPV: 38%, NPV: 92%).

Comparison With Other Short Versions

While the ZBI-9 shares 7 items with the ZBI-12 by Bédard,[4] it only shares 5 items with the version by Ballesteros[10] and 3 items with the version by Hebert.[23] Table 5 displays the internal consistency, convergent validity, and AUC of the three 12-item short versions, which were calculated using the current sample. Cronbach's α ranged from 0.90 to 0.92, and are slightly higher than ZBI-9 (α = 0.87). For convergent validity, the correlations for HADS-A (r = 0.63 to 0.65) and HADS-D (r = 0.58 to 0.60) were comparable to the ZBI-9. Similarly, the discriminative ability of the three 12-item short versions to detect significant depressive symptoms (AUC = 0.80–0.81) were comparable to the ZBI-9 (AUC = 0.79). Not surprisingly, ZBI-22 had higher internal consistency and AUC compared with the short versions (Table 5).

DISCUSSION

We developed a multidimensional short version (ZBI-9) that accounted for the prior factor structure of the original 22-item ZBI to provide a comprehensive assessment of the burden among caregivers of older persons with cognitive impairment. We adopted a rigorous approach combining quantitative analysis and qualitative content evaluation by experts to develop the multidimensional 9-item tool. To our knowledge, this is the first empirically derived short version of the ZBI that incorporates the unique dimension of WaP.[14,16] Taken together, ZBI-9 has good psychometric properties and facilitates a valid, reliable, and multidimensional assessment of caregiver burden that is less onerous and shorter to complete than the ZBI-22, enhancing its acceptability for both clinical and research usage. A novel feature of our study is the inclusion of caregivers for individuals with MCI, thereby extending the use of our tool to this population as well.

Our results revealed that the ZBI-9 performed well in terms of fit statistics in CFA with good factor loadings. This supports the factorial validity of the tool, which is predicated on the validated 4-factor solution of the 22-item ZBI,[18] namely role strain (2 factors), personal strain, and WaP. Notably, the multidimensionality of ZBI-9 encompasses WaP, which has been consistently reported in both Asian and Western literature.[7] Indeed, WaP appears to be a common dimension among adult children caregivers, likely due to the influence of prevailing societal norms and perceived expectations of caregiving[24,25] or familial relationships—love and the wish to return the best care to loved ones.[26]

Corrected item-total correlations were good for all items in the ZBI-9 with the exception of the 2 items from

TABLE 3. Item Means, CFA Factor Loadings, and Expert Panel Evaluation of ZBI-22 Items

No.	Item	Mean (SD)	CFA factor loadings	Expert panel's reasons for inclusion / exclusion
Factor 1 Demands of care and social impact on caregiver (RS1)		—	—	—
1.	Do you feel that your care recipient asks for more help than he/she needs?	1.42 (1.03)	0.52	Excluded due to: Low factor loading (≤ 0.6)
2.	Do you feel that because of the time you spend with your care recipient that you don't have enough time for yourself?	1.27 (1.08)	0.74	Included due to: High factor loading Expert Panel - representative of factor 1
3.	Do you feel stressed between caring for your care recipient and trying to meet other responsibilities for your family or work?	1.52 (1.08)	0.83	Included due to: High factor loading Expert Panel - representative of factor 1
4.	Do you feel embarrassed over your care recipient's behavior?	0.63 (0.83)	0.57	Excluded due to: Low factor loading (≤ 0.6)
7.	Are you afraid what the future holds for your care recipient?	1.51 (1.25)	0.59	Excluded due to: Low factor loading (≤ 0.6)
8.	Do you feel your care recipient is dependent on you?	2.19 (1.20)	0.56	Excluded due to: Low factor loading (≤ 0.6)
11.	Do you feel that you don't have as much privacy as you would like because of your care recipient?	0.95 (1.09)	0.77	Excluded (despite high factor loading) due to: Expert Panel - accuracy may be affected by different interpretations of 'privacy', for instance, 'invasion of personal space' versus 'private time'
12.	Do you feel that your social life has suffered because you are caring for your care recipient?	0.96 (1.04)	0.83	Included due to: High factor loading Expert Panel - representative of factor 1
14.	Do you feel that your care recipient seems to expect you to take care of him/her as if you were the only one he/she could depend on?	1.74 (1.28)	0.51	Excluded due to: Low factor loading (≤ 0.6)
Factor 2 Confidence or control over the situation (RS2)		—	—	—
13.	Do you feel uncomfortable about having friends over because of your care recipient?	0.48 (0.79)	0.65	Excluded (despite high factor loading) due to: Expert Panel – may allude more towards social impact of caregiving than confidence or control over situation
15.	Do you feel that you don't have enough money to take care of your care recipient in addition to the rest of your expenses?	0.99 (1.06)	0.56	Excluded due to: Low factor loading (≤ 0.6)
16.	Do you feel that you will be unable to take care of your care recipient much longer?	0.65 (0.86)	0.69	Included due to: High factor loading Expert Panel - representative of factor 2
17.	Do you feel that you have lost control of your life since the care recipient's illness?	0.73 (0.92)	0.87	Included due to: High factor loading Expert Panel - representative of factor 2
18.	Do you wish that you could leave the care of your care recipient to someone else?	0.83 (1.00)	0.60	Excluded due to: Low factor loading (≤ 0.6)
Factor 3 Psychological impact on caregiver (PS)		—	—	—
5.	Do you feel angry when you are around your care recipient?	1.08 (0.92)	0.66	Excluded (despite high factor loading) due to: Expert Panel - covers an overly narrow aspect of psychological impact
6.	Do you feel that your care recipient currently affects your relationships with other family members in a negative way?	0.79 (1.01)	0.67	Excluded (despite high factor loading) due to: Preference for item 10 over item 6
9.	Do you feel strained when you are around your care recipient?	1.12 (1.00)	0.79	Included due to: High factor loading Expert Panel - representative of factor 3
10.	Do you feel your health has suffered because of your involvement with your care recipient?	0.77 (1.04)	0.69	Included due to: High factor loading Expert Panel - covers a unique and clinically relevant dimension in caregiver burden
19.	Do you feel uncertain about what to do about your care recipient?	1.13 (1.06)	0.67	Excluded (despite high factor loading) due to: Expert Panel: less representative of factor 3
22.	Overall, how burdened do you feel in caring for your care recipient?	1.24 (1.04)	0.75	Excluded as it is an overall burden item
Factor 4 Worry about caregiving performance		—	—	—
20.	Do you feel you should be doing more for your care recipient?	1.66 (1.06)	0.98	Included due to: High factor loading Expert Panel - item represent a distinct factor "worry about caregiving performance" which cannot be represented by other factors.
21.	Do you feel you could do a better job in caring for your care recipient?	1.61 (1.02)	0.65	Included due to: High factor loading Expert Panel - item represent a distinct factor "worry about caregiving performance" which cannot be represented by other factors.

factor 4. This is consistent with the findings of earlier studies[11] and highlights the potential pitfalls of solely relying on psychometric indices for item selection in the development of short versions without taking into account the prior factor structure of the original scale. Nonetheless, it is reassuring that factor 4 demonstrated good internal

TABLE 4. Psychometric properties and CFA factor loadings for ZBI-9

			CFA Loadings			
Items	Corrected Item-total Correlation	Item-factor Correlation	Factor 1	Factor 2	Factor 3	Factor 4
1 Do you feel that because of the time you spend with your care recipient that you don't have enough time for yourself?	0.72	0.90	0.80	—	—	—
2 Do you feel stressed between caring for your care recipient and trying to meet other responsibilities for your family or work?	0.77	0.89	0.84	—	—	—
3 Do you feel that your social life has suffered because you are caring for your care recipient?	0.78	0.89	0.86	—	—	—
4 Do you feel that you will be unable to take care of your care recipient much longer?	0.55	0.90	—	0.68	—	—
5 Do you feel that you have lost control of your life since the care recipient's illness?	0.73	0.91	—	0.93	—	—
6 Do you feel strained when you are around your care recipient?	0.68	0.88	—	—	0.79	—
7 Do you feel your health has suffered because of your involvement with your care recipient?	0.59	0.89	—	—	0.73	—
8 Do you feel you should be doing more for your care recipient?	0.33	0.91	—	—	—	0.96
9 Do you feel you could do a better job in caring for your care recipient?	0.26	0.90	—	—	—	0.67
Cronbach's alpha			0.87	0.77	0.73	0.78

consistency and high factor loadings. As a 9-item short version, ZBI-9 demonstrated good psychometric properties in terms of internal consistency and convergent and concurrent validity, which is comparable to the original 22-item ZBI. Moreover, in comparison with the 12-item short-versions, which do not account for the multi-dimensionality of ZBI, the performance of the ZBI-9 is largely comparable despite having fewer number of items.

Taken together, ZBI-9 offers a reliable and valid alternative to existing short versions of the ZBI, with the advantage of incorporating WaP to maintain the multidimensional content validity of the burden construct. A multidimensional assessment is beneficial as caregivers with the same total burden score may differ in the types of burden they are experiencing, necessitating different target interventions. For instance, an earlier study that examined the relationship between 'doing more' (item 20) and 'doing better' (item 21) in the evaluation of WaP burden reported that the 'doing more' high WaP subgroup was associated with adult children caregivers and disruptive behaviors in the domain of depression/dysphoria, alluding to the importance of early identification and intervention to mitigate downstream consequences.[13] In addition, caregivers with WaP burden can benefit from psychoeducation with specific psychological interventions, such as psychotherapeutic and mindfulness-based interventions, to enhance the sense of personal mastery and self-efficacy, thus averting the slippery slope to more negative appraisals of one's caregiving performance which may result in guilt, overall caregiver stress, and ultimately burnout.[27]

This is especially salient in light of the observation that WaP, unlike role and personal strain, occurs early in mild cognitive impairment and very mild dementia,[18] and in the trajectory of multidimensional ZBI burden.[28]

There are some limitations to our study. Firstly, participants were only recruited from a single tertiary outpatient clinic. Nevertheless, the sample is comparable with earlier studies of memory clinic populations and is fairly representative of the caregivers of older adults with mild-to-moderate dementia.[7] Secondly, because our study population is largely comprised of the caregivers of persons with dementia, results in the MCI population would need to be confirmed in other studies. Thirdly, being a cross-sectional study, we were unable to validate the predictive validity of ZBI-9 compared with ZBI-22. Fourthly, current literature has yet to come to a consensus regarding the factor structure of the ZBI. The present study adopted a 4-factor structure which is validated and reflects the multidimensional nature of the ZBI. Although the WaP factor did not perform as well as the other 3 factors in some areas, the 4-factor structure is corroborated by CFA fit statistics. Lastly, while there may be contention that the use of an expert panel for item selection may be inherently subjective, this was part of a rigorous process whereby the qualitative appraisals complement the use of quantitative methods (CFA and EFA); furthermore, it is increasingly recognized that qualitative appraisals are an important aspect of scale development.[17] Further studies are needed to establish if the

TABLE 5. Validity Statistics of ZBI Short Versions

ZBI versions	No. items	Cronbach's α	ZBI-22*	HADS-A*	HADS-D*	AUC†
ZBI-9	9	0.87	0.95	0.62	0.57	0.79
ZBI-22	22	0.93	–	0.65	0.59	0.81
Other short versions						
Bedard et al, 2001[4]	12	0.90	0.97	0.63	0.58	0.81
Ballesteros et al, 2012[10]	12	0.92	0.97	0.64	0.59	0.80
Hebert et al, 2000[23]	12	0.92	0.97	0.65	0.60	0.81

*Represents Pearson's correlations. All correlations are statistically significant ($P < 0.001$).
†AUC = area under curve analysis.

validity and reliability of ZBI-9 can be generalizable to more severe stages of dementia or to other clinical settings, for example, in more acute care settings.

CONCLUSION

Using a 2-step approach of quantitative psychometric analysis and qualitative content evaluation, we developed and validated the ZBI-9, a multidimensional 9-item short version of ZBI. ZBI-9 retains the validated 4-factor structure of the original 22-item ZBI and has good psychometric properties that are comparable with previous 12-item short versions. The ZBI-9 provides a comprehensive assessment of caregiver burden, including WaP, and has the advantage of being quicker to administer compared with the 22-item ZBI.

ACKNOWLEDGMENTS

REFERENCES

1. Zarit SH, Reever KE, Bach-Peterson J. Relatives of the impaired elderly: correlates of feelings of burden. *Gerontologist.* 1980;20:649–655.
2. Ko K-T, Yip P-K, Liu S-I, et al. Chinese version of the Zarit Caregiver Burden Interview: a validation study. *Am J Geriatr Psychiatry.* 2008;16:513–518.
3. 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. *Jap J Geriatr.* 2003;40:497–503.
4. Bédard M, Molloy DW, Squire L, et al. The Zarit Burden Interview: a new short version and screening version. *Gerontologist.* 2001;41:652–657.
5. Liew TM, Yap P. A 3-item screening scale for caregiver burden in dementia caregiving: Scale development and score mapping to the 22-item Zarit Burden Interview. *J Am Med Dir Assoc.* 2019;20:629–633.e612.
6. Gort AM, Mingot M, March J, et al. Short Zarit scale in dementias. *Med Clin (Barc).* 2010;135:447–449.
7. Lau S, Ali N, Lim JP, et al. ZBI multidimensionality: The salience of worry about performance. *Int Psychogeriatr.* 2019;31:909–911.
8. Knight BG, Fox LS, Chou C-P. Factor structure of the burden interview. *J Clin Geropsychol.* 2000;6:249–258.
9. Lau S, Chong MS, Ali N, et al. Caregiver burden: Looking Beyond the Unidimensional Total Score. *Alzh Dis Assoc Dis.* 2015;29:338–346.
10. Ballesteros J, Santos B, González-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.
11. Li R, Chong MS, Chan PCM, et al. Worry About Caregiving Performance: A Confirmatory Factor Analysis. *Front Med.* 2018;5:79.
12. Whitlatch CJ, Zarit SH, von Eye A. Efficacy of interventions with caregivers: a reanalysis. *Gerontologist.* 1991;31:9–14.
13. Heyzer L, Ali NB, Chew AP, et al. Worry about Performance: Unravelling the Relationship between 'Doing More' and 'Doing Better'. *J Nutr Health Aging.* 2019;23:843–848.
14. Lim WS, Cheah WK, Ali N, et al. Worry about performance: a unique dimension of caregiver burden. *Int Psychogeriatr.* 2014;26:677–686.
15. Branger C, O'Connell ME, Morgan DG. Factor Analysis of the 12-Item Zarit Burden Interview in Caregivers of Persons Diagnosed With Dementia. *J Appl Gerontol.* 2016;35: 489–507.
16. Tang JY-m, Ho AH-y, Luo H, et al. Validating a Cantonese short version of the Zarit Burden Interview (CZBI-Short) for dementia caregivers. *Aging Ment Health.* 2016;20:996–1001.
17. Goetz C, Coste J, Lemetayer F, et al. Item reduction based on rigorous methodological guidelines is necessary to maintain validity when shortening composite measurement scales. *J Clin Epidemiol.* 2013;66:710–718.
18. Cheah WK, Han HC, Chong MS, et al. Multidimensionality of the Zarit Burden Interview across the severity spectrum of cognitive impairment: an Asian perspective. *Int Psychogeriatr.* 2012;24:1846.
19. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand.* 1983;67:361–370.
20. Bjelland I, Dahl AA, Haug TT, et al. The validity of the Hospital Anxiety and Depression Scale: an updated literature review. *J Psychosom Res.* 2002;52:69–77.
21. Marsh HW, Hau K-T, Grayson D. Goodness of Fit in Structural Equation Models. In: Maydeu-Olivares A, McArdle JJ, eds. *Contemporary psychometrics: A festschrift for Roderick P.* McDonald. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers; 2005:275–340.
22. Moran LA, Guyatt GH, Norman GR. Establishing the minimal number of items for a responsive, valid, health-related quality of life instrument. *J Clin Epidemiol.* 2001;54:571–579.
23. Hébert R, Bravo G, Préville M. Reliability, validity and reference values of the Zarit Burden Interview for assessing informal caregivers of community-dwelling older persons with dementia. *Can J Aging/La Revue Canadienne du vieillissement.* 2000;19:494–507.
24. Zarzycki M, Morrison V, Bei E, et al. Cultural and societal motivations for being informal caregivers: A qualitative systematic review and meta-synthesis. *Health Psychol Rev.* 2022;5:1–201.
25. Liu J, Bern-Klug M. I should be doing more for my parent: Chinese adult children's worry about performance in providing care for their oldest-old parents. *Int Psychogeriatr.* 2016;28:303.
26. Greenwood N, Smith R. Motivations for being informal carers of people living with dementia: a systematic review of qualitative literature. *BMC Geriatr.* 2019;19:1–18.
27. Cheng ST, Au A, Losada A, et al. Psychological Interventions for Dementia Caregivers: What We Have Achieved, What We Have Learned. *Curr Psychiatry Rep.* 2019;21:59.
28. Li D, Hu N, Yu Y, et al. Trajectories of Multidimensional Caregiver Burden in Chinese Informal Caregivers for Dementia: Evidence from Exploratory and Confirmatory Factor Analysis of the Zarit Burden Interview. *J Alzheimers Dis.* 2017;59:1317–1325.