



The optimal short version of the Zarit Burden Interview for dementia caregivers: diagnostic utility and externally validated cutoffs

Junhong Yu, Philip Yap & Tau Ming Liew

To cite this article: Junhong Yu, Philip Yap & Tau Ming Liew (2019) The optimal short version of the Zarit Burden Interview for dementia caregivers: diagnostic utility and externally validated cutoffs, *Aging & Mental Health*, 23:6, 706-710, DOI: [10.1080/13607863.2018.1450841](https://doi.org/10.1080/13607863.2018.1450841)

To link to this article: <https://doi.org/10.1080/13607863.2018.1450841>



© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 19 Mar 2018.



[Submit your article to this journal](#)



Article views: 6852



[View related articles](#)



[View Crossmark data](#)



Citing articles: 20 [View citing articles](#)

The optimal short version of the Zarit Burden Interview for dementia caregivers: diagnostic utility and externally validated cutoffs

Junhong Yu^a, Philip Yap^{b,c} and Tau Ming Liew^{d,e}

^aDepartment of Psychology, The University of Hong Kong, Hong Kong; ^bDepartment of Geriatric Medicine, Khoo Teck Puat Hospital, Singapore; ^cGeriatric Education and Research Institute, Singapore; ^dDepartment of Geriatric Psychiatry, Institute of Mental Health, Singapore; ^eSaw Swee Hock School of Public Health, National University of Singapore, Singapore

ABSTRACT

Objectives: Using a sample of dementia caregivers, we compared the diagnostic utility of the various short versions of the Zarit Burden Interview (ZBI) with the original scale to identify the most optimal one. Next, we established externally validated cutoffs for the various ZBI versions using probable depression cases as a reference standard.

Methods: Caregivers ($N = 394$; 236 males; Age_{mean} = 56 years) were administered the ZBI and a self-report depression measure. Participants who exceeded the cutoff for the latter were identified as probable depression cases. For each of the ZBI versions, a receiver operating characteristic (ROC) curve was plotted against probable depression cases. The area under these ROC curves between the short versions and the original were then compared using a non-parametric approach.

Results: Compared to the original ZBI, the AUROC were similar for the 6-item, 7-item, and two 12-item versions, but significantly worse for the other short variants. The sensitivity and specificity of the cutoffs for all ZBI versions ranged from 77.3% to 85.2% and 60.1% to 79.8%, respectively.

Conclusions: The original ZBI had good utility in identifying probable depression in caregivers, while the 6-item variant can be a useful alternative when short versions are preferred.

ARTICLE HISTORY

Received 6 December 2017
Accepted 26 February 2018

KEYWORDS

Dementia; caregivers; depression; Zarit Burden Interview; receiver operating characteristic

Introduction

As the prevalence of dementia soars due to the rapidly aging populations in developed countries, an increasing number of people will have to care for a loved one with dementia and undertake its associated burdens. Caregiver burden have been previously defined as the financial, physical and psychological consequence of caring for an adult with a disabling condition (George & Gwyther, 1986). Caring for a person with dementia (PWD) is a long-term commitment and may span up to 20 years after the initial dementia diagnosis (Karlín, Bell, Noah, Martichuski, & Knight, 1999). As a result of such long term burden, meta-analytic research have documented that caregivers, relative to non-caregivers, are more likely to suffer from depression and physical illnesses, as well as experience a decrease in self-efficacy and subjective well-being (Pinquart & Sörensen, 2003). Hence, the need to assess caregiving-related burden cannot be understated. Such assessments help inform the clinician if appropriate interventions are needed and ultimately the lived experience with dementia for both the caregiver and PWD.

The Zarit Burden Interview (ZBI; Zarit & Zarit, 1987) is one of the most widely used measures to assess caregiver burden. Since its inception, this 22-item scale has been translated into many languages and used in many countries across a diverse range of caregivers and patient populations; meta-analytic research have suggested the ZBI to be reliable across the diverse contexts in which it has been used in (Bachner & O'Rourke, 2007). To facilitate quicker administration, several short versions of the ZBI have been developed, using various methods, ranging from single-item to 18-item versions.

Whitlatch, Zarit, and von Eye (1991) carried out an exploratory factor analysis (EFA) and identified the two factors of personal strain and role strain, consisting of 18 items in total. Arai, Tamiya, and Yano (2003) similarly obtained the two factors of personal and role strain in their EFA. However their model consisted of eight items. In Knight, Fox, and Chou (2000) EFA, they identified the three factors of embarrassment/anger, patient's dependency and self-criticism which collectively consisted of 14 items. Bédard et al. (2001) created the 4-item and 12-item versions, by choosing four and 12 items, respectively, with the highest item-total correlation within a similar personal and role strain two-factor model that emerged as the optimal model in their EFA. Hébert, Bravo, and Prévaille (2000) obtained another 12-item version by carrying out an EFA based on the Whitlatch et al. (1991) two-factor model and selected items which constitute the most parsimonious structure of the two-factor model. These studies were carried out on dementia caregiver populations. Using data from caregivers of palliative care patients; Gort et al. (2005) produced their short version by having an expert committee select seven items. Higginson, Gao, Jackson, Murray, and Harding (2010) modified this 7-item version by removing the global item question (item 22) to obtain a 6-item version, which was validated in a mixed sample of dementia, cancer and brain injured patients' caregivers. Also within the same study, the authors examined a 1-item (global item question) version. The list of included items of these different short versions is presented in Table 1.

Nevertheless, the ZBI and their short variants are fraught with a few problems. First, having multiple short versions can

Table 1. Overview of all studied ZBI short variants.

Items in the original 22-item ZBI	ZBI variants								
	18-item ¹	14-item ²	12-item ³	12-item ⁴	8-item ⁵	7-item ⁶	6-item ⁷	4-item ³	1-item ⁷
1 Your relative asks for more help than he/she needs?	✓								
2 You don't have enough time for yourself?	✓	✓	✓	✓		✓	✓	✓	
3 Stressed between caring and meeting other responsibilities?	✓		✓	✓		✓	✓	✓	
4 Embarrassed over behaviors?	✓	✓			✓				
5 Angry when around your relative?	✓	✓	✓		✓				
6 Your relative affects your relationship with others in a negative way?	✓	✓	✓	✓	✓	✓	✓		
7 Afraid of what the future holds for relative?				✓					
8 Your relative is dependent on you?	✓	✓							
9 Strained when are around your relative?	✓	✓	✓	✓	✓	✓	✓	✓	
10 Your health has suffered because of your involvement with your relative?		✓	✓	✓		✓	✓		
11 You don't have as much privacy as you would like, because of your relative?	✓	✓	✓	✓					
12 Your social life has suffered because you are caring for your relative?	✓	✓	✓	✓	✓				
13 Uncomfortable about having friends over because of your relative?	✓	✓		✓	✓				
14 Your relative seems to expect you to take care of him/her, as if you were the only one he/she could depend on?	✓	✓							
15 You don't have enough money to care for your relative, in addition to the rest of your expenses?									
16 You will be unable to take care of your relative much longer?	✓								
17 You have lost control of your life since your relative's illness?	✓		✓	✓		✓	✓		
18 You could just leave the care of your relative to someone else?	✓	✓		✓	✓				
19 Uncertain about what to do about relative?	✓		✓		✓			✓	
20 You should be doing more for your relative?	✓	✓	✓						
21 You could do a better job in caring for your relative?	✓	✓	✓						
22 Overall, how burdened do you feel in caring for your relative?				✓		✓			✓

Note: 1 = Whitlatch et al. (1991); 2 = Knight et al. (2000); 3 = Bédard et al. (2001); 4 = Hébert et al. (2000); 5 = Arai et al. (2003); 6 = Gort et al. (2005); 7 = Higginson et al. (2010)

be problematic as it makes for comparison between them difficult, and it is unclear if these short versions are equivalent to the original in terms of diagnostic utility. It would also be useful to know which of these versions is the most optimal, that is – the least-item version that has comparable diagnostic utility to the original. Furthermore, these short variants and the original lack cutoffs that are validated against a clinically significant outcome. For instance, the cutoff in the original ZBI (Zarit & Zarit, 1987) and one of the 12-item (Bédard et al., 2001) versions were determined arbitrarily. While Higginson et al. (2010) reported the cutoffs for various short versions of ZBI in their study, they obtained these cutoffs by using the arbitrarily determined cutoff in the original as a reference standard. As such, these cutoffs lack external validity, as O'Rourke and Tuokko (2003) observed that the cutoff for the 12-item variant (Bédard et al., 2001) had low levels of sensitivity in classifying probable depression cases. The lack of validated cutoffs for the ZBI and its short variants makes it difficult for one to assess the clinical significance of the scores and, more importantly, the need for intervention.

Therefore, the current study aimed to address these issues and better inform clinicians on the use of the ZBI and their short versions. In the current study, we compared the diagnostic utility, via the area under the receiver operating characteristic (AUROC), between the various ZBI short versions and the original, to identify the most optimal ZBI short version. Next we sought to determine externally validated cutoffs for the various ZBI variants. Given that a specific clinical diagnosis does not exist to characterize excessive caregivers burden, similar to a previous validation study (O'Rourke & Tuokko, 2003) we opted to use a common consequence of such excessive burdens – depression (Alspaugh, Stephens, Townsend, Zarit, & Greene, 1999; Berger et al., 2005), as a reference standard. That is, if one becomes depressed, it would be reasonable to assume that such burdens are excessive.

Methods

Participants and procedures

Participants were recruited from the dementia services of two tertiary hospitals in the north-eastern part of Singapore. We used a consecutive sampling method and had a response rate of 87.8%. Participants were included if they are: (1) age 21 or above; (2) spouses or children of PWD; and (3) caring for PWD who is residing in the community. A total of 394 caregivers for persons at various stages of dementia were recruited. Table 2 presents the demographical characteristics of these caregivers and the person they are caring for. The PWDs were mostly females (Age_{mean} = 79.5; SD = 8.2) and at the moderate and severe stage of dementia. Their caregivers (Age_{mean} = 53.0; SD = 10.7) were mostly Chinese, married, children of PWD and the primary caregiver of the PWD. On average, they have cared for the PWD for 6.8 years (SD = 6.7).

At the respective clinics in the hospitals, these caregivers gave their informed consent before completing onsite, the self-administered demographic questionnaire, the ZBI, and Center for Epidemiological Studies Depression Scale (CES-D). This study had received the ethical approval from the Domain Specific Review Board of the National Healthcare Group, Singapore.

Measures

The ZBI (Zarit & Zarit, 1987) measures the perceived burden of caregivers via 5-point Likert scale items. These items were summed to produce a total score ranging from 0 to 88. According to the original test instructions, score range of 61–88 indicates high burden. The ZBI has demonstrated good reliability and validity for assessing caregiver burden in our local context (Yap, 2010). On top of the original 22-item total score, we also calculated the total scores for the different

Table 2. Demographic information of the caregivers and the persons with dementia ($n = 394$).

Participant characteristics	Mean (SD); range/Frequency
Caregivers	
Age	53.0 (10.7); 22–83
Females	236 (59.9)
Ethnicity	
Chinese	341 (86.6)
Malay	25 (6.3)
Indian	18 (4.6)
Others	10 (2.5)
Marital status	
Married	271 (68.8)
Single	94 (23.9)
Widowed/Divorced/Separated	29 (7.3)
Employment status	
Not working	123 (31.2)
Working part-time	52 (13.2)
Working full-time	219 (55.6)
Highest education	
Primary or no formal education	41 (10.4)
Secondary	228 (57.9)
Tertiary	125 (31.7)
Relationship with PWD	
Child	340 (86.3)
Spouse	54 (13.7)
Staying with PWD	264 (67.0)
Duration of caregiving in years	6.8 (6.7); 0.08–17.67
Frequency of caregiving	
Daily, for at least 4 h a day	211 (53.6)
Daily, but less than 4 h a day	79 (20.0)
At least once a week	84 (21.3)
Less than once a week	20 (5.1)
Primary caregiver role	279 (70.8)
ZBI score	34.8 (16.8); 0–80
CES-D score	15.7 (11.0); 0–54
Caregivers with probable depression (CES-D ≥ 16)	176 (44.7)
PWD	
Age	79.5 (8.2); 57–97
Females	278 (70.6)
Age at dementia diagnosis	75.6 (8.5); 49–95
Duration of dementia diagnosis in years	4.5 (3.5); 0.3–17.7
Stage of dementia ^a	
Mild	62 (15.7)
Moderate	163 (41.4)
Severe	169 (42.9)

Note: PWD = person with dementia; SD = standard deviation.

^a The stage of dementia was obtained using a brief measure based on the three levels of dementia severities described in the DSM-III-R. Participants chose among three options, corresponding to these three levels of severities, that best described the PWD – still capable of independent living (mild stage), needs some assistance with daily living (moderate stage), or needs round-the-clock supervision (severe stage). This brief measure is also nearly identical to the re-introduced dementia severity in the DSM-5.

short versions of the ZBI by summing up their respective items (see Table 1).

The CES-D is a 20-item, self-report scale which assessed depressive symptoms in the past one week (Radloff & Radloff,

Table 3. Correlation between scores of CES-D and the various ZBI versions

	Correlation with CES-D (r)
original 22-item (Zarit & Zarit, 1987)	0.71
12-item (Bédard et al., 2001)	0.71
12-item (Hébert et al., 2000)	0.72
7-item (Gort et al., 2005)	0.73
6-item (Higginson et al., 2010)	0.73
18 item (Whitlatch et al., 1991)	0.69
14-item (Longmire & Knight, 2011)	0.67
8-item (Arai et al., 2003)	0.62
4-item (Bédard et al., 2001)	0.67
1-item (Higginson et al., 2010)	0.56

Note: All correlation coefficients were statistically significant (Bonferroni adjusted $p < .001$)

1977). Each item is scored on a 4-point Likert scale to reflect the frequency of the described symptom. The total score for this scale ranges from 0 to 60. A cut-off score of ≥ 16 , as suggested by the original authors, was used in the current study to classify participant as having probable depression. This cut-off has been validated in the local context and has demonstrated high levels of sensitivity ($\geq 90.9\%$; Stahl et al., 2008), although its specificity was relatively low (≤ 67.6) in classifying depression cases. Such levels of specificity were similar to the pooled specificity statistic obtained via a meta-analysis (Vilagut, Forero, Barbaglia, & Alonso, 2016). Furthermore, in view of the trade-off between sensitivity and specificity, it may be better to opt for higher sensitivity than specificity, especially since false-negative diagnoses may have more costly consequences than those of false-positive in the clinical context.

Analyses

For each of the studied ZBI variants, we plotted a receiver operating characteristic (ROC) curve against probable depression cases and calculated the AUROC. Next, we assessed whether the AUROC of the various ZBI short versions were significantly different from that of the original via a non-parametric approach (DeLong & Carolina, 1988) to derive confidence intervals and standard errors of the differences in AUROC. Correlations between variables were examined using Pearson correlation coefficients. Statistical significance was set at $p < 0.05$. Given that multiple comparisons were carried out, we also computed bonferroni-adjusted p-values. All analyses were performed using the STATA software version 14.

Results

The results of the correlation between the total scores of the various ZBI version and CES-D are presented in Table 3. All

Table 4. AUROC statistics and comparison of the ZBI variants with the original 22-item version, and cutoffs related information

ZBI variants (first author)	AUROC (95% CI)	Compared with original 22-item			Optimal cutoff ^a	Sensitivity (%)	Specificity (%)
		P	Adj. P				
original 22-item (Zarit & Zarit, 1987)	0.859 (0.823, 0.895)	–	–	≥ 34	85.2	74.8	
12-item (Bédard et al., 2001)	0.859 (0.823, 0.895)	0.982	1	≥ 19	79.6	76.2	
12-item (Hébert et al., 2000)	0.862 (0.826, 0.898)	0.667	1	≥ 18	80.7	79.8	
7-item (Gort et al., 2005)	0.866 (0.831, 0.902)	0.411	1	≥ 11	80.7	79.4	
6-item (Higginson et al., 2010)	0.863 (0.827, 0.898)	0.710	1	≥ 9	81.3	77.5	
18 item (Whitlatch et al., 1991)	0.847 (0.809, 0.885)	0.001	0.009	≥ 28	82.4	73.9	
14-item (Longmire & Knight, 2011)	0.835 (0.796, 0.875)	<0.001	0.001	≥ 21	80.1	70.2	
8-item (Arai et al., 2003)	0.805 (0.762, 0.849)	<0.001	<0.001	≥ 10	77.3	73.4	
4-item (Bédard et al., 2001)	0.830 (0.790, 0.870)	0.005	0.044	≥ 7	83.0	70.6	
1-item (Higginson et al., 2010)	0.779 (0.735, 0.824)	<0.001	<0.001	≥ 2	84.7	60.1	

Note: ZBI = Zarit Burden Interview; AUROC = Area Under the Receiver Operating Characteristic; CI = confidence interval; Adj. = Bonferroni-adjusted. P values represent the statistical significance of the difference in AUROC between a ZBI variant and the original 22-item version.

^a The optimal cutoff is based on a balance between sensitivity and specificity, with a preference for slightly higher sensitivity to reduce the false negative rates.

correlations were statistically significant (Bonferroni adjusted $p > 0.001$). The 6-item and 7-item versions had the highest correlation ($r = 0.73$) with CES-D scores while the 1-item version had the lowest ($r = 0.56$). Table 4 presents the results of the AUROC analyses. The original ZBI had AUROC of 0.859. All ZBI variants yielded at least acceptable AUROC (ranging from 0.779 to 0.866). The 6-item (Higginson et al., 2010), 7-item (Gort et al., 2005) and the two 12-item (Bédard et al., 2001; Hébert et al., 2000) ZBI variants' AUROC were not significantly different from the original ZBI (all Bonferroni adjusted $p > 0.05$), while the rest had significantly worse AUROC relative to the original ZBI.

The derived cutoffs for all ZBI versions had acceptable levels of sensitivity and specificity ($>70\%$), except for the 1-item (Higginson et al., 2010) version (see Table 3). At the optimal cutoff of ≥ 34 , the original ZBI had a sensitivity of 85.2% and a specificity of 74.8%. Detailed results of the sensitivity and specificity statistics for all ZBI variants are reported in the Supplementary materials.

Discussion

The aim of the current study is twofold. First, we sought to identify the most optimal ZBI short version and derive externally validated cutoffs for the different versions of ZBI. We compared the AUROC between the ZBI short versions and the original. The 6-item version (Higginson et al., 2010) emerged as the most optimal short version in having the least number of items yet demonstrated comparable diagnostic utility as the original 22-item version. The cutoff for this version also had relatively high specificity and sensitivity in classifying probable depression cases. Interestingly, while this 6-item version had a comparable AUROC value to the original 22-item version, a few of the longer ZBI versions (Arai et al., 2003; Knight et al., 2000; Whitlatch et al., 1991) actually had significantly lower AUROC values than the original. One possible explanation for this is that the extra items beyond the 6-item version may be less relevant in the context of depression, hence their inclusion is unlikely to increase the scale's diagnostic utility as measured via the accuracy in classifying probable depression cases. Relatedly, in showing that ZBI scores correlate significantly and highly with depression scores, we have also demonstrated the excellent convergent validity of ZBI and provided evidence to confirm similar findings reported previously (Tang et al., 2016).

Next, using probable depression cases as a reference standard, we derived the cutoffs for the full ZBI and its shorter variants. It is interesting to note that the cutoffs reported previously (Bédard et al., 2001; Higginson et al., 2010; Zarit & Zarit, 1987) were lower than our validated cutoffs. Given that lower cutoffs generally correspond to lower specificity, this suggests that the previous cutoffs could suffer from reduced specificity in assessing caregiver burden and consequently result in false positive errors (i.e. identified as excessive burden when it is in fact not). Hence, one should exercise caution in interpreting the results of previous research that used these cutoffs.

These findings have two important implications. Firstly, the externally validated cutoffs for ZBI afford clinicians greater confidence in identifying caregivers who are experiencing high levels of caregiving-related burden, to the extent that they might be suffering from depression or at risk of developing it. This will facilitate early intervention for these caregivers

and efforts to improve the caregiving experience for both the caregiver and PWD. Secondly, by showing that the 6-item, 7-item and both 12-item versions of the ZBI are equivalent to the original 22-item version in terms of diagnostic utility, these shorter versions can be utilized to assess for caregiver burden with greater convenience and confidence.

Some limitations of the study are noteworthy. First, the participants were recruited from two hospitals in one region of Singapore, hence they may not be geographically representative. Second, the participants were recruited from dementia services in tertiary centers and may therefore not be fully representative of those in the community. However, this is less likely a problem considering that most PWDs in Singapore receive care from tertiary centers. Third, the caregivers sample included only spouses and children of the PWD, hence these findings may not apply to other caregivers such as the PWD's children-in-law or paid caregivers. Fourth, the sample was over-represented by Chinese; as such the results may only be applicable to the Chinese population. In relation to these generalizability issues, future studies may consider replicating these findings in other populations. Finally, given the cross-sectional nature of this study, it is possible that the depressive symptoms assessed among caregivers may not be a consequence of caregiving-related burden but rather the result of other unrelated stressors. As such, future studies may consider verifying the current findings with a longitudinal design.

Acknowledgment

The authors thank the participants and staff at Institute of Mental Health and Khoo Teck Puat Hospital, for their support of this research.

Disclosure statement

The authors have no conflicts of interest to declare.

Funding

This work was supported by the National Medical Research Council, Ministry of Health, Singapore [grant numbers NMRC/CG/004/2013, NMRC/Fellowship/0030/2016 to T.M.L, NMRC/CSSSP/0014/2017 to T.M.L] and pilot funding from the National University of Singapore.

References

- Alspaugh, M. E. L., Stephens, M. A. P., Townsend, A. L., Zarit, S. H., & Greene, R. (1999). Longitudinal patterns of risk for depression in dementia caregivers: Objective and subjective primary stress as predictors. *Psychology and Aging, 34*, 34–43. US: American Psychological Association. Retrieved from <https://doi.org/10.1037/0882-7974.14.1.34>
- Arai, Y., Tamiya, N., & Yano, E. (2003). The short version of the Japanese version of the Zarit caregiver burden interview (J-ZBI_8): Its reliability and validity. *Nihon Ronen Igakkai Zasshi. Japanese Journal of Geriatrics, 40*(5), 497–503.
- Bachner, Y. G., & O'Rourke, N. (2007). Reliability generalization of responses by care providers to the Zarit Burden Interview. *Aging & Mental Health, 11*(6), 678–685. Retrieved from <https://doi.org/10.1080/13607860701529965>
- Bédard, M., Molloy, D. W., Squire, L., Dubois, S., Lever, J. A., & O'donnell, M. (2001). The Zarit Burden Interview: A new short version and screening version. *Gerontologist, 41*(5), 652–657. Retrieved from <https://doi.org/10.1093/geront/41.5.652>
- Berger, G., Bernhardt, T., Weimer, E., Peters, J., Kratzsch, T., & Frolich, L. (2005). Longitudinal study on the relationship between symptomatology of dementia and levels of subjective burden and depression among family caregivers in memory clinic patients. *Journal of Geriatric*

- Psychiatry and Neurology*, 18(3), 119–128. Retrieved from <https://doi.org/10.1177/0891988704273375>
- Delong, E. R., & Carolina, N. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach author (s): Elizabeth R. DeLong, David M. DeLong and Daniel L. Clarke-Pearson published by: International biometric society stable. *Biometrics*, 44(3), 837–845.
- George, L. K., & Gwyther, L. P. (1986). Caregiver well-being: A multidimensional examination of family caregivers of demented adults. *The Gerontologist*, 26(3), 253–259.
- Gort, A. M., March, J., Gómez, X., de Miguel, M., Mazarico, S., & Ballesté, J. (2005). Short Zarit scale in palliative care. *Medicina Clinica*, 124(17), 651–653.
- Hébert, R., Bravo, G., & Prévile, M. (2000). Reliability, validity and reference values of the Zarit Burden Interview for assessing informal caregivers of community-dwelling older persons with dementia. *Canadian Journal on Aging / La Revue Canadienne Du Vieillessement*, 19(4), 494–507. Retrieved from <https://doi.org/10.1017/S0714980800012484>
- Higginson, I. J., Gao, W., Jackson, D., Murray, J., & Harding, R. (2010). Short-form Zarit caregiver burden interviews were valid in advanced conditions. *Journal of Clinical Epidemiology*, 63(5), 535–542. Retrieved from <https://doi.org/10.1016/j.jclinepi.2009.06.014>
- Karlin, N. J., Bell, P. A., Noah, J. L., Martichuski, D. K., & Knight, B. L. (1999). Assessing Alzheimer's support group participation: A retrospective follow-up. *American Journal of Alzheimer's Disease*, 14(6), 326–333. Retrieved from <https://doi.org/10.1177/153331759901400607>
- Knight, B. G., Fox, L. S., & Chou, C. -P. (2000). Factor structure of the burden interview. *J Clin Geropsychology*, 6(4), 249–258. Retrieved from <https://doi.org/10.1023/A:1009530711710>
- Longmire, C. V. F., & Knight, B. G. (2011). Confirmatory factor analysis of a brief version of the Zarit Burden Interview in black and white dementia caregivers. *Gerontologist*, 51(4), 453–462. Retrieved from <https://doi.org/10.1093/geront/gnr011>
- O'Rourke, N., & Tuokko, H. A. (2003). Psychometric properties of an abridged version of the Zarit Burden Interview within a representative Canadian caregiver sample. *The Gerontologist*, 43(1), 121–127. Retrieved from <https://doi.org/10.1093/geront/43.1.121>
- Pinquart, M., & Sörensen, S. (2003). Differences between caregivers and noncaregivers in psychological health and physical health: A meta-analysis. *Psychology and Aging*, 18(2), 250.
- Radloff, L. S., & Radloff, L. S. (1977). Center for epidemiologic studies depression scale (CES-D). *Applied Psychological Measurement* 1(1), 385–401.
- Stahl, D., Sum, C. F., Lum, S. S., Liow, P. S., Chan, Y. H., Verma, S.,... Chong, S. A. (2008). Screening for depressive symptoms: Validation of the CES-D scale in a multi-ethnic group of patients with diabetes in Singapore. *Diabetes Care*, 31, 1118–1119.
- Tang, J. Y., Ho, A. H., Luo, H., Wong, G. H., Lau, B. H., Lum, T. Y., & Cheung, K. S. (2016). Validating a cantonese short version of the Zarit Burden Interview (CZBI-Short) for dementia caregivers. *Aging & Mental Health*, 20(9), 996–1001. Retrieved from <https://doi.org/10.1080/13607863.2015.1047323>
- Vilagut, G., Forero, C. G., Barbaglia, G., & Alonso, J. (2016). Screening for depression in the general population with the center for epidemiologic studies depression (CES-D): A systematic review with meta-analysis. *PLoS One*, 11(5), e0155431.
- Whitlatch, C. J., Zarit, S. H., & von Eye, A. (1991). Efficacy of interventions with caregivers: A reanalysis. *The Gerontologist*, 31(1), 9–14. Retrieved from <https://doi.org/10.1093/geront/31.1.9>
- Yap, P. (2010). Validity and reliability of the Zarit Burden Interview in assessing caregiving burden. *Ann Acad Med Singapore*, 39, 758–763.
- Zarit, S. H., & Zarit, J. M. (1987). *Instructions for the burden interview*. University Park: Pennsylvania State University.