Alzheimer's & Dementia
Diagnosis, Assessment
& Disease Monitoring

RESEARCH ARTICLE

Depression screening in cognitively normal older adults: Measurement bias according to subjective memory decline, brain amyloid burden, cognitive function, and sex

Louisa I. Thompson | Richard N. Jones

Department of Psychiatry, Alpert Medical School, Brown University, Providence, Rhode Island, USA

Correspondence
Louisa Thompson, Memory and Aging Program, Butler Hosptial, 345 Blackstone Blvd., Providence, RI 02906.
Email: louisa_thompson@brown.edu

Abstract

Introduction: Understanding the associations among depression, subjective cognitive decline, and prodromal Alzheimer's disease (AD) has important implications for both depression and dementia screening in older adults. The Geriatric Depression Scale (GDS) is a depression screening tool for older adults that queries memory concerns. To determine whether depression symptoms on the GDS (15-item version), including self-reported memory problems, differ by levels of brain amyloid beta ($A\beta$), a pathological hallmark of early stage AD, we investigated potential measurement bias with regard to $A\beta$ level. We also examined measurement bias attributable to level of cognitive functioning and sex as positive controls.

Methods: We examined 3961 cognitively normal older adults from the A4/LEARN Study. We used the MIMIC (multiple indicators, multiple causes) approach to detect measurement bias.

Results: We found measurement bias with small-to-moderate range effect sizes in several GDS-15 items with respect to $A\beta$ level, cognitive functioning, and sex. There was negligible impact of measurement bias attributable to $A\beta$ level on overall depressive symptom level.

Discussion: GDS-15 item responses are sensitive to $A\beta$ burden, cognitive functioning, and sex over and above what would be expected given the effect of those factors on depressive symptom severity overall. However, these direct effects for GDS item measurement bias are of small magnitude and do not appreciably impact the validity of inferences about depression based on the GDS-15.

KEYWORDS
Alzheimer's disease, amyloid, depression, subjective cognitive decline

1 | INTRODUCTION

Depression in older adults has been associated with worse subjective and objective cognitive performance,[1] higher rates of Alzheimer's disease (AD) and related dementias,[2] and neurodegenerative brain pathology.[3] It remains unclear whether depression is a risk factor for neurodegenerative dementia versus a consequence of it. Both lifetime history of depression[4] and late-life onset depression[3] are associated with elevated levels of brain amyloid beta ($A\beta$), a neuropathological hallmark of AD, suggesting that depression is a risk factor for AD. Con-

versely, higher baseline brain Aβ in cognitively normal older adults has been associated with up to a 4.5-fold increased likelihood of developing clinically significant depression symptoms over several years, even after adjusting for depression history,[5,6] suggesting Aβ deposition may be a risk factor for depression.

Subjective cognitive decline (SCD) is also a risk factor for subsequent AD and a recognized symptom of depression.[7–9] SCD, defined as cognitive complaints in the absence of objective cognitive impairment on standardized tests,[10] increases with age and is commonly reported in community samples of older adults.[11–13] Cognitively normal older adults with SCD may be twice as likely as their peers without SCD to develop dementia.[9] SCD has been associated with brain biomarkers of vascular and neurodegenerative disease and increased dementia risk before objective cognitive dysfunction can be detected.[10,14,15]

Taken together, the evidence for associations among depression, SCD, and AD is complex, but has important implications for depression screening in older adults. For example, it is unclear to what extent screening for depression in older adults could be biased by the presence of SCD or AD neuropathology. Successful detection and timely intervention for depression often rests heavily upon depression symptom questionnaires used in primary care settings. Such measures also play a critical role in screening and symptom tracking in clinical trials for the prevention and treatment of AD and related dementias. Given the co-occurrence of SCD and depression and their associations with AD, it is critical that we understand their interplay and potential impact on how older adults respond to common depression self-report measures, such as the 15-item version of the Geriatric Depression Scale (GDS-15), which includes a question about memory concerns ("Do you feel that you have more problems with memory than most?").[16,17]

This study investigated measurement bias in GDS-15 items in the assessment of depression with respect to an objective index of preclinical AD. We looked for evidence of variable item endorsement (differential item functioning [DIF]) on the basis of AD brain pathology (cerebral Aβ burden), as well as sex and cognitive function, other factors known to interact with self-reported depression, in the individual items comprising the GDS-15.[18] Sex was examined because older women face greater depression risk than men, and previous work has found at least one GDS item exhibited DIF when comparing men and women.[19] We hypothesized that positive endorsement of the GDS memory item (ie, "I have more problems with memory than most" = YES) would be predicted by brain cerebral Aβ over and above the relationship of cerebral Aβ to overall GDS depression severity and beyond the effects of sex and cognition. We conducted our investigation using a large sample of cognitively normal older adults.

2 | METHOD

2.1 | Participants

Pre-randomization (ie, screening) data from the Anti-Amyloid Treatment in Asymptomatic Alzheimer's Disease (A4)/Longitudinal Evaluation of Amyloid Risk and Neurodegeneration (LEARN) Study were

RESEARCH IN CONTEXT

1. **Systematic review**: The authors conducted a secondary data analysis using the A4/LEARN Study prerandomization data to detect measurement bias in the 15-item Geriatric Depression Scale (GDS-15) items according to subjective cognitive decline, cerebral amyloid beta (Aβ) burden, cognitive function, and sex in cognitively normal older adults.

2. **Interpretation**: GDS-15 item responses are sensitive to Aβ burden, cognitive functioning, and sex over and above what would be expected given the effect of those factors on depressive symptom severity overall. Most importantly, persons with high Aβ had a greater probability of endorsing memory concern on the GDS-15 than those with low Aβ.

3. **Future directions**: Our findings underscore the importance of considering SCD and cerebral amyloid burden in the context of depression screening in clinical settings and AD research trials. Further work should investigate whether the GDS-15 item measurement bias that we observed is consistent across more diverse samples, including community and clinical samples with higher rates of depression.

used for all analyses. The A4 study is a secondary prevention trial of an anti-amyloid antibody (solanezumab) in clinically normal individuals age 65 to 85 years with elevated brain Aβ on a positron emission tomography (PET) scan. The pre-randomization sample includes those who cleared all trial inclusion and exclusion criteria leading up to the Aβ PET scan, which was the final step to determine eligibility. Inclusion criteria consisted of having a Mini-Mental State Examination[20] score between 25 and 30, global Clinical Dementia Rating[21] scale score of 0, Logical Memory II[22] score between 6 to 18, and a study partner. Data from a total of 3961 participants with available GDS, Aβ PET, cognitive assessment, and demographic data were included in the present analysis. Sample demographics are included in Table 1.

2.2 | Depression assessment

Depressive symptomology was assessed using the 15-item version of the GDS (GDS-15), which uses a simple "yes/no" self-report format completed by the participant. The GDS is the most commonly used screening self-report measure for depression in individuals over the age of 65 years.[23] The 15 items used for this short form of the GDS were chosen due to their high correlation with depressive symptoms in previous validation studies.[16] The GDS-15 is strongly correlated ($r = .84$, $P < .001$) with the original 30-item GDS. The short form has a sensitivity of 92% and a specificity of 81% using a cut-off of five.[16]

TABLE 1 Participant demographic and clinical characteristics

Characteristic	Overall		SUVR ≥1.1		SUVR <1.1	
	Mean or n	(SD or %)	Mean or n	(SD or %)	Mean or n	(SD or %)
Number of observations(N [%])	3961	(100)	1394	(100)	2567	(100)
Age (mean [SD])	71.4	(4.7)	72.1	(4.9)	71.0	(4.6)
Sex (n [%])						
Male	1565	(40)	556	(40)	1009	(39)
Female	2396	(60)	838	(60)	1558	(61)
Ethnicity (n [%])						
Hispanic or Latinx	140	(3)	45	(3)	95	(4)
Not Hispanic or Latinx	3786	(96)	1332	(96)	2454	(96)
Unknown/not reported	35	(1)	17	(1)	18	(<1)
Race (n [%])						
American Indian/Alaska Native	9	(<1)	3	(<1)	6	(<1)
Asian	64	(2)	15	(1)	49	(2)
Native Hawaiian/other Pacific Islander	2	(<1)	1	(<1)	1	(<1)
Black or African American	160	(4)	41	(3)	119	(5)
White	3673	(93)	1316	(95)	2357	(93)
Unknown/not reported	26	(1)	10	(1)	16	(1)
Education (years) (mean [SD])	16.7	(2.7)	16.7	(2.7)	16.7	(2.8)
Total SUVR (range 0.7, 2.1) (mean [SD])	1.10	(0.20)	1.31	(0.18)	0.98	(0.06)
PACC Total Score Item Normalized Result Numeric (range −12,7, higher better) (mean [SD])	0.1	(2.5)	-0.3	(2.6)	0.3	(2.5)
GDS score (0-15, higher worse) (mean [SD])	1.0	(1.5)	1.0	(1.3)	1.0	(1.5)
GDS score of 5 or higher (n [%])	125	(3.2)	30	(2.2)	95	(3.7)

Abbreviations: GDS, Geriatric Depression Scale; PACC, Preclinical Alzheimer's Cognitive Composite; SD, standard deviation; SUVR, standardized uptake value ratio from amyloid positron emission tomography scan.

The A4 study excluded participants with a history of major depression based on Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria during the past 2 years. GDS-15 scores of 10 or higher at screening were also exclusionary. As such, there is a relatively low level of clinically significant symptom endorsement on the GDS-15 in this sample, consistent with other large AD observational studies and clinical drug trials.

2.3 | Cognitive assessment

Cognitive functioning was assessed using the ADCS-PACC (Alzheimer's Disease Cooperative Study-Preclinical Alzheimer Cognitive Composite)[24] total score, a composite of standardized z scores from tests that assess episodic memory, timed executive function, and global cognition. The PACC was developed for use in AD prevention clinical trials and has demonstrated sensitivity to detect cognitive decline associated with Aβ deposition.[24] PACC composite score decreases with worse performance.

2.4 | Amyloid PET neuroimaging

Cerebral Aβ level was determined by post-processed Aβ tracer florbetapir (18F-AV-45) PET imaging data. PET data processing was conducted by Invicro LLC. Our analysis used the continuous composite total cerebral amyloid tracer standardized uptake value ratio (SUVR) calculated using whole cerebellum as the reference region. For computing model implied characteristic curves, we used an established florbetapir cut-off of 1.10 to distinguish between Aβ– and Aβ+ participants, as previously described in the literature.[25] The cut-off used for determining Aβ status in the A4 study is not published.

3 | DATA ANALYSES

All statistical analyses were conducted using Mplus version 8.0 (Muthén & Muthén, Los Angeles, California). We used the MIMIC (multiple indicators, multiple causes) model to evaluate the GDS-15

TABLE 2 Item response proportions by sex, and differences in proportion endorsing expressed as crude standardized effect sizes (Cohen's h) and direct effects from MIMIC model (DIF model direct effect, k)

| | Proportion endorsing | | | Effect size | | |
| | | | | Crude | DIF model | |
Item	Total	Men	Women	h	k	P
(Not) satisfied with your life	.030	.029	.030	+0.01		
Dropped many activities	.059	.058	.060	+0.01		
Feel life is empty	.021	.018	.023	+0.04		
Get bored	.071	.080	.066	−0.05	−0.20	.001
(Not) in good spirits	.041	.035	.044	+0.05		
Afraid something bad will happen	.057	.048	.063	+0.07		
(Not) feel happy	.054	.053	.054	+0.00	−0.21	.001
Feel helpless	.030	.023	.035	+0.07		
Prefer to stay at home	.185	.204	.172	−0.08	−0.22	<.001
Problems with memory	.133	.146	.124	−0.06	−0.11	.04
(Not) wonderful to be alive	.036	.027	.041	+0.08		
Feel worthless	.012	.008	.015	+0.07		
(Not) feel full of energy	.250	.258	.245	−0.03	−0.16	.002
Feel hopeless	.009	.008	.010	+0.02		
Most others are better off	.021	.015	.025	+0.07		
					d	P
Indirect effect in underlying depression (female)					+0.26	<.001

Abbreviations: d, Cohen's d standardized mean difference effect size; DIF, differential item functioning; MIMIC, multiple indicators, multiple causes.

for evidence of measurement bias—or DIF—according to SUVR, level of cognitive functioning (PACC total score), and sex. This approach has been used previously to evaluate depression measures for evidence of DIF.[26] DIF occurs when respondents from different groups (or at different levels of a continuous background variable) at the same latent trait respond differently to the same item.[27] This approach can be viewed as a kind of multilevel model, in which we model the dependency of an item response (eg, responding yes/no to *Are you basically satisfied with your life?*) on both the underlying but unobserved latent variable presumed to cause depressive item responses (eg, depression level) and on background variables (sex, SUVR, and PACC score). Simultaneously, we regress the latent depression level on background variables. We iteratively regress GDS-15 item responses on background variables that the model identifies as potential sources of misfit if not estimated. These direct effects are evidence of DIF. The DIF effect, κ, can be interpreted as an effect size statistic. More details are available in the supporting information appendix.

Continuous predictors were mean centered and standardized to two standard deviation (SD) units, in accordance with recommendations from Gelman,[28] so as to produce an analytic variable with an SD of 0.5, which is the same scaling as a balanced binary variable.[28] This means that when comparing predictor effects in a regression model with a mix of binary and continuous predictors, and the continuous predictors are standardized to 2SD, the magnitude of the coefficients will be on a similar scale as those of binary variables. Negative GDS-

15 items were reverse coded, and are identified by "(not)" in the item descriptions for Tables 2, 3, and 4.

4 | RESULTS

Participant characteristics by Aβ SUVR (elevated vs not elevated) are shown in Table 1. Total mean SUVR in this cognitively normal sample was at the elevated threshold of 1.10 (SD = 0.20), and values ranged from 0.7 to 2.1. PACC scores were significantly lower (indicating worse performance) among those with elevated SUVR compared to those with sub-threshold SUVR. Total mean GDS score was 1.0 (SD = 1.50, range 0–12), which is well below the established positive symptom-screening cut-off of five. In the overall sample, not feeling full of energy, preferring to stay at home, and memory concerns were the most commonly endorsed GDS indicators of depression. The overall sample item response frequencies for each GDS-15 item endorsed are shown alongside the DIF effects by sex in Table 2, by SUVR in Table 3, and by PACC score in Table 4.

4.1 | Detected differential item functioning

Tables 2, 3, and 4 summarize direct effects by predictor variables: sex, Aβ burden (SUVR), and level of cognitive functioning (PACC score). Overall, we detected 14 DIF effects, and 7 of these describe

TABLE 3 Item response proportions by SUVR, and differences in proportion endorsing expressed as crude standardized effect sizes (Cohen's h) and direct effects from MIMIC model (DIF model direct effect, k)

| | Proportion endorsing | | | Effect size | | |
| | | | | Crude | DIF model | |
Item	Total	SUVR ≥1.1	SUVR <1.1	h	k	P
(Not) satisfied with your life	.030	.025	.032	+0.04		
Dropped many activities	.059	.057	.060	+0.01		
Feel life is empty	.021	.019	.022	+0.02		
Get bored	.071	.063	.076	+0.05		
(Not) in good spirits	.041	.039	.041	+0.01		
Afraid something bad will happen	.057	.056	.057	+0.00		
(Not) feel happy	.054	.049	.057	+0.04		
Feel helpless	.030	.029	.030	+0.01		
Prefer to stay at home	.185	.177	.189	+0.03		
Problems with memory	.133	.162	.117	−0.13	−0.22	< .001
(Not) wonderful to be alive	.036	.040	.033	−0.04	−0.18	.02
Feel worthless	.012	.009	.014	+0.05		
(Not) feel full of energy	.250	.230	.261	+0.07		
Feel hopeless	.009	.012	.008	−0.04	−0.26	.03
Most others are better off	.021	.019	.023	+0.03		
					d	P
Indirect effect in underlying depression (SUVr < 1.1)					+0.13	.01

Abbreviations: d, Cohen's d standardized mean difference effect size; DIF, differential item functioning; MIMIC, multiple indicators, multiple causes; SUVR, standardized uptake value ratio.

trivial effects. We detected no DIF effects that can be considered medium or large. Item characteristic curves for all 15 items and all DIF effects are also shown in Figures S1 and S2, respectively, in supporting information.

Women had a higher mean level of underlying latent depression severity (d = 0.26, P < .001, Table 2). This would lead us to expect women to endorse symptoms more commonly than men. This is observed for most entries in Table 2, but for some symptoms the base rates of endorsement are similar by sex and many of these are flagged as having significant DIF effects (direct effects) indicating a lower probability of endorsement, conditional on underlying level of the latent trait, for women. Examples include the item "*Do you feel happy most of the time?*" (k = −0.21, P = .001). This is one of the positively worded items in the GDS, and was reversed before analysis (indicated by the "[not]" in Table 2). The negative coefficient implies that women were *less* likely to endorse this item than were men conditional on the underlying level of depressive symptom severity (Figure 1B). A similar effect is observed for *bored*: this item also had a significant direct effect (k = −0.20, P = .001). Also similarly, the base rate was higher among men for *prefer to stay at home* (k = −.22, P < .001).

We found that persons with low SUVR had a slightly greater level of depressive symptom burden (d = +0.13, P = .01, Table 3) relative to persons with higher SUVR, though the overall difference was not clini-

cally meaningful. This would lead to the expectation of slightly higher base rates for symptom endorsement for the low SUVR group relative to the comparison group. This was observed, with the exception of *Problems with memory* (k = −0.22, P < .001, Table 3, Figure 2), *(not) Wonderful to be alive* (k = −0.18, P = .02), and *Feel hopeless* (k = −0.26, P = .03). These effects are all around the level of what would be called small effects in Cohen's effect size taxonomy[29] (.2). Nonetheless, our hypothesis was confirmed, such that persons with higher Aβ SUVR values had a greater probability of endorsing memory problems relative to persons with a lower SUVR at the same level of underlying depressive symptom severity. We also repeated our analysis using the continuous SUVR variable as the predictor, and found measurement bias for the same items with similar effect sizes given a linear rescaling of SUVR, including the memory item (k = −0.16, P < .001), the Alive item (k = −0.14, P = .02), and the Hope item (k = −0.20, P = .03). These results suggest our results are not due to bias introduced by dichotomizing SUVR.

Some of the largest DIF effects were observed for the relationship between cognitive performance level (PACC score, Table 4) and the GDS items *Do you (not) feel happy most of the time* (k = 0.35, P < .001) and *Do you think it is (not) wonderful to be alive now* (k = 0.26, P = .001). The indirect effect according to low versus high PACC score was −.40, implying a small to medium effect size and overall lower

TABLE 4 Item response proportions by PACC score, and differences in proportion endorsing expressed as crude standardized effect sizes (Cohen's h) and direct effects from MIMIC model (DIF model direct effect, k)

	Proportion endorsing			Effect size		
				Crude	DIF model	
		PACC score				
Item	Total	Low	High	h	k	P
(Not) satisfied with your life	.030	.036	.023	−0.08		
Dropped many activities	.059	.074	.044	−0.13		
Feel life is empty	.021	.028	.014	−0.10		
Get bored	.071	.081	.061	−0.08		
(Not) in good spirits	.041	.045	.037	−0.04	+0.16	.04
Afraid something bad will happen	.057	.065	.048	−0.07		
(Not) feel happy	.054	.053	.055	+0.01	+0.35	<.001
Feel helpless	.030	.039	.021	−0.11		
Prefer to stay at home	.185	.192	.177	−0.04	+0.14	.005
Problems with memory	.133	.167	.097	−0.21	-0.18	<.001
(Not) wonderful to be alive	.036	.035	.037	−0.04	+0.26	.001
Feel worthless	.012	.016	.008	−0.07		
(Not) feel full of energy	.250	.261	.239	−0.05	+0.14	.005
Feel hopeless	.009	.010	.009	−0.01		
Most others are better off	.021	.024	.018	−0.04		
					d	P
Indirect effect in underlying depression (high PACC score)					−.40	<.001

Abbreviations: d, Cohen's d standardized mean difference effect size; DIF, differential item functioning; MIMIC, multiple indicators, multiple causes; PACC, Preclinical Alzheimer's Cognitive Composite.

Notes: Low and high group expected proportions, and effect size statistics, are estimated from a logistic regression model with the item response as the dependent variable, and the PACC score included as a linear predictor. The "low" expected proportion corresponds to the expectation at a PACC score 1 standard deviation below the mean, and the "high" expected proportion corresponds to that at 1 standard deviation above the mean.
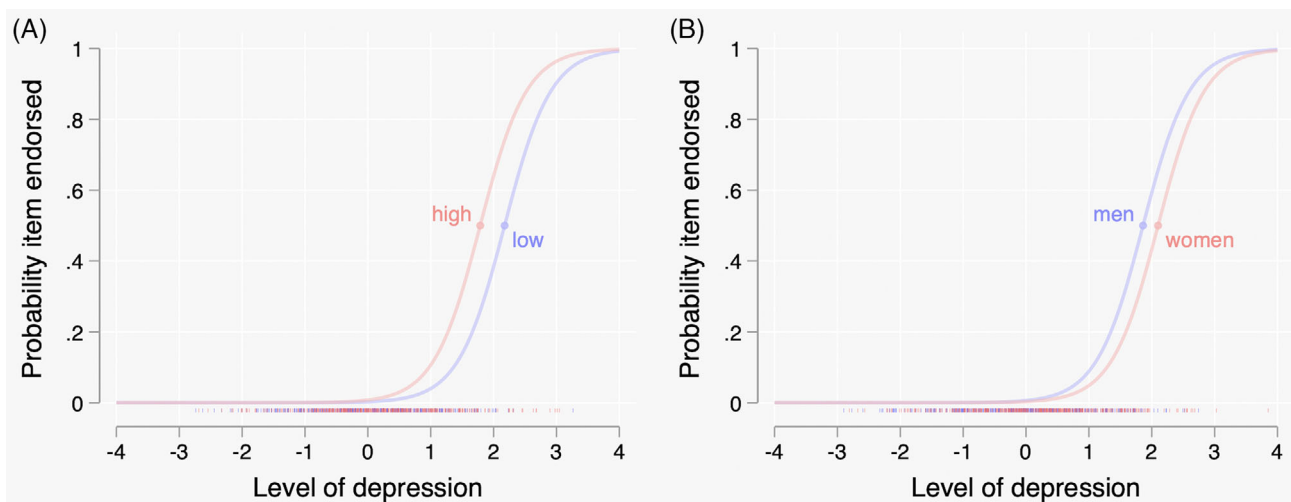
FIGURE 1 Model-implied item characteristic curve for Geriatric Depression Scale "happy" and for persons with "high" versus "low" cognitive performance on the Preclinical Alzheimer's Cognitive Composite (PACC) a continuous predictor (panel A) and sex as a categorical predictor (panel B). The y-axis indicates the expected proportion endorsing the item, and the x-axis indicates increasing levels of depression symptom severity. Individuals with higher cognitive performance on the PACC are more likely to endorse (not) being happy relative to those with lower cognition holding constant the underlying depressive symptom severity. Women are more likely than men to endorse (not) being happy
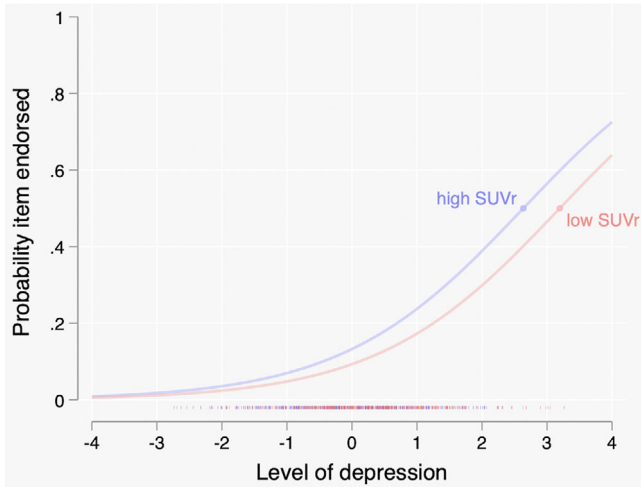
FIGURE 2 Model-implied item characteristic curve for Geriatric Depression Scale "memory" and for persons with "high" versus "low" amyloid beta (Aβ) standard uptake value ratio (SUVR) as the dichotomous predictor. The y-axis indicates the expected proportion endorsing the item, and the x-axis indicates increasing levels of depression symptom severity. Individuals with a higher Aβ SUVR are more likely to endorse having worse memory than most at a lower level of underlying depressive symptom severity than persons with a low SUVR

probability of endorsing depressive symptoms with better cognitive performance. The positive coefficients imply that individuals who achieve a high PACC score have a higher probability of endorsing these symptoms at a given underlying level of depressive symptom severity than do persons with a low PACC score (Figure 1A).

5 | DISCUSSION

Depression and SCD screening in older adults at risk for AD provides an opportunity for early intervention to improve emotional and cognitive health outcomes in older adults.[30,31] Using a large sample of cognitively normal older adults screening for an AD prevention trial, our analyses show that Aβ burden influences symptom endorsement on a widely used depression self-report measure.

As hypothesized, mild measurement bias attributable to Aβ burden was shown for the GDS-15 memory item indicating that, holding overall depression symptom level constant, persons with high Aβ had a greater probability of endorsing memory concern than those with low Aβ. Memory item measurement bias attributable to Aβ was greater than memory item measurement bias attributable to cognition or sex. Our finding is consistent with the literature, which suggests a linear relationship between SCD and cerebral Aβ deposition in persons with preclinical or prodromal AD.[14,32] It is possible this could mean that in some cases when GDS-15 symptomology endorsement is borderline clinically significant (eg, a score of 5), persons with higher levels of Aβ burden may endorse the GDS memory item pushing them into the clinically depressed range. However, the direct effect of measurement bias

attributable to Aβ burden in our sample was of small magnitude and did not appreciably impact the validity of inferences based on the GDS-15. Overall, our findings suggest that the GDS-15 is an appropriate self-report measure to use for screening for depression symptoms in older adults who may have subjective cognitive decline and those with preclinical AD.

Interestingly, lower Aβ burden was associated with slightly higher levels of depression symptomology overall, and the reason for this finding is unclear. In contrast, Aβ burden was also associated with direct effects in two items assessing hopelessness, such that increased Aβ was associated with greater endorsement of feelings of hopelessness, holding underlying depression severity constant. Hopelessness is a core symptom of depression and apathy and may be less likely to be elevated due to medical conditions in the elderly compared to more somatic-oriented symptoms (eg, reduced energy).

It is well established in the literature that depression is often strongly associated with cognitive decline. Our finding of measurement bias on GDS-15 hopelessness items related to Aβ burden lends support to the idea that, in addition to the relationship between depression and cognitive decline, a separate link may also exist between core depression symptoms and Aβ pathology. Others have failed to find associations between GDS-15 reported depression symptomology and Aβ pathology;[33] however, ours is the first analysis to consider this question at the level of individual scale items. It is possible that changes in brain connectivity shared between depression and Aβ deposition encompassing the default mode network, and areas such as the insula and cingulate may contribute to an association between depression and Aβ burden.[34,35] Our finding could also potentially be explained by an effect of family history, with the hypothesis being that older adults with a strong family history of dementia feel more hopeless about their own futures and are also more likely to have elevated Aβ. This may be particularly true in the A4 study screening sample representing individuals who are motivated to try an experimental medication for the prevention of AD.

We also found small magnitude measurement bias attributable to sex and cognitive function on some GDS-15 items. Women had a greater likelihood of endorsing depression symptoms overall, though men were more likely to endorse boredom and a desire to stay at home compared to women at the same underlying level of depression. Lower cognitive performance was associated with higher levels of depression symptomology overall, and this was also true for the memory item, indicating that lower levels of SCD were associated with better cognitive performance, as predicted. A few GDS-15 items demonstrated effects in the opposite direction (eg, items: happy, alive) of the overall association between depression and cognitive performance. Upon closer inspection, these items appeared to have complex, non-linear relationships to cognition, which could potentially be explored by different advanced analytic techniques, which are outside the scope of this article. However, the impact of measurement bias attributable to sex and cognitive function on overall depressive symptom level in this sample was negligible. These findings are consistent with recent studies that found small magnitude measurement bias attributable to age, sex, and cognitive functioning in a more representative

sample of Italian older adults.[18,19] While we found more items with measurement bias, this is likely because of our much larger sample size. Taken together, while some measurement bias in GDS-15 item functioning may exist in relation to sex and cognitive function, it does not seem to appreciably impact the validity of inferences based on the GDS-15.

While we observed measurement bias in memory item functioning related to both brain Aβ level and cognitive performance, it may be worth noting that memory item measurement bias related to Aβ level was greater than memory item measurement bias related to cognitive performance. This could in part be explained by the literature suggesting that SCD is associated with hypoconnectivity and overall better cognitive performance in cognitively normal older adults at risk for AD, despite lower immediate memory and global cognition when followed over time.[36] We would expect the strength of the association between SCD endorsement and cognitive performance to increase with the progression of AD from the preclinical to early symptomatic stages.

5.1 | Limitations and future directions

The main limitations to this investigation pertain to the sample used, which consisted of individuals who were in screening for an AD prevention drug trial (the A4 study), and were therefore not necessarily representative of older adults in the United States more broadly. Indeed, the sample was primarily White and educated beyond high school. Factors such as knowledge about AD, positive family history of AD, and subjective cognitive decline may be more prevalent in this sample relative to the general population, as these factors have been shown to be motivating factors for intentions to screen for AD or participate in AD research.[37,38]

The low level of depression symptomology reported in this sample limits our ability to draw conclusions about associations among depression symptoms, SCD, and Aβ levels, at least insofar as we were unable to examine them in a sample with higher rates of clinically significant symptom endorsement. Specifically, the analysis of measurement bias that we conducted is entirely dependent on the range of data in this sample. The DIF effects we observed for our variables of interest may be different in a clinical sample of depressed older adults. However, the mean GDS score of 1 and SD of 1.5 in our sample is consistent with other samples of cognitively normal older adults used in AD clinical research.[3,39] Additionally, we found similar results for measurement bias due to sex as Chiesi et al., who used a more representative community sample in their investigations of the GDS-15, suggesting that our findings may be generalizable.[18,19]

Our analysis was limited to cross-sectional data, and therefore we were unable to conduct any longitudinal analysis of associations between self-reported depression symptoms, including memory complaints and hopelessness, cognitive function, sex, and brain Aβ levels. Longitudinal investigations, such that of Harrington et al. discussed previously, are critical to understanding the temporal unfolding of relations among brain changes in AD, depression, and SCD.[3] A recent longitudinal analysis from the German Study on Aging, Cognition, and Dementia found that higher levels of SCD preceded both future objective memory decline and an increase in depression symptoms, suggesting that screening for SCD and depression symptoms may be particularly important in preclinical AD samples.[40]

5.2 | Concluding remarks

This study investigated whether self-reported depression symptoms on the GDS-15 differ by brain Aβ burden through an investigation of potential measurement bias with regard to cerebral Aβ level in a large sample of cognitively normal older adults from the A4/LEARN Study. After adjustment for depressive symptom severity (GDS total score), we found that GDS-15 item responses are sensitive to Aβ burden, cognitive functioning, and sex over and above what would be expected given the effect of those factors on depressive symptom severity overall. Persons with high Aβ had a greater probability of endorsing memory concern than those with low Aβ, but the impact of measurement bias attributable to Aβ level on overall depressive symptom level was negligible. While our findings raise important questions for future research on the relationship between Aβ and self-reported depression and memory concerns, the direct effects that were found in this study are of small magnitude and do not appreciably impact the validity of inferences based on the GDS-15.

CONFLICTS OF INTEREST

REFERENCES

1. Morimoto SS, Alexopoulos GS. Cognitive deficits in geriatric depression: clinical correlates and implications for current and future treatment. *Psychiatr Clin North Am*. 2013;36(4):517-531.

2. Byers AL, Yaffe K. Depression and risk of developing dementia. *Nat Rev Neurol*. 2011;7(6):323-331.

3. Harrington KD, Gould E, Lim YY, et al. Amyloid burden and incident depressive symptoms in cognitively normal older adults. *Int J Geriatr Psychiatry*. 2017;32(4):455-463.

4. Chung JK, Plitman E, Nakajima S, et al. Lifetime history of depression predicts increased amyloid-β accumulation in patients with mild cognitive impairment. *J Alzheimers Dis*. 2015;45(3):907-919.

5. Babulal GM, Ghoshal N, Head D, et al. Mood changes in cognitively normal older adults are linked to Alzheimer disease biomarker levels. *Am J Geriatr Psychiatry*. 2016;24(11):1095-1104.

6. Donovan NJ, Locascio JJ, Marshall GA, et al. Longitudinal association of amyloid beta and anxious-depressive symptoms in cognitively normal older adults. *Am J Psychiatry*. 2018;175(6):530-537.

7. Yates JA, Clare L, Woods RT. *Subjective memory complaints, mood and MCI: a follow-up study. Aging Ment Health*. 2015:1-9.

8. Yates JA, Clare L, Woods RT, Matthews FE. Subjective memory complaints are involved in the relationship between mood and mild cognitive impairment. *J. Alzheimer's Dis*. 2015;48:S115-S123.

9. Mitchell AJ, Beaumont H, Ferguson D, Yadegarfar M, Stubbs B. Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: meta-analysis. *Acta Psychiatr Scand*. 2014;130(6):439-451.

10. Rabin LA, Smart CM, Amariglio RE. Subjective cognitive decline in preclinical Alzheimer's disease. *Annu Rev Clin Psychol*. 2017;13(1):369-396.

11. Jonker C, Geerlings MI, Schmand B. Are memory complaints predictive for dementia? A review of clinical and population-based studies. *Int J Geriatr Psychiatry*. 2000;15(11):983-991.

12. Reid LM, MacLullich AMJ. Subjective memory complaints and cognitive impairment in older people. *Dement Geriatr Cogn Disord*. 2006;22(5-6):471-485.

13. Mendes T, Ginó S, Ribeiro F, et al. Memory complaints in healthy young and elderly adults: reliability of memory reporting. *Aging Ment Health*. 2008;12(2):177-182.

14. Amariglio RE, Becker JA, Carmasin J, et al. Subjective cognitive complaints and amyloid burden in cognitively normal older individuals. *Neuropsychologia*. 2012;50(12):2880-2886.

15. Snitz BE, Weissfeld LA, Cohen AD, et al. Subjective cognitive complaints, personality and brain amyloid-beta in cognitively normal older adults. *Am J Geriatr Psychiatry*. 2015;23(9):985-993.

16. Sheikh JI, Yesavage JA. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clin Gerontol*. 1986;5(1-2):165-173.

17. Yesavage J, Brink T, Rose T, Adey M. The Geriatric Depression Rating Scale: comparison with other self-report and psychiatric rating scales. In: Crook T, Ferris S, Bartus R, eds. *Assessment in Geriatric Psychopharmacology*. New Haven, CT: Mark Powles Associates, Inc.; 1983:153-167.

18. Chiesi F, Primi C, Pigliautile M, et al. Is the 15-item Geriatric Depression Scale a fair screening tool? A differential item functioning analysis across gender and age. *Psychol Rep*. 2018;121(6):1167-1182.

19. Chiesi F, Primi C, Pigliautile M, et al. Does the 15-item Geriatric Depression Scale function differently in old people with different levels of cognitive functioning. *J Affect Disord*. 2017;227:471-476.

20. Folstein MF, Folstein SE, McHugh PR. 'Mini-mental state.' A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198.

21. Morris JC. Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *Int Psychogeriatr*. 2005;9(S1):173-176.

22. Wechsler D. *Wechsler Memory Scale-Revised manual*. San Antonio: Psychological Corporation; 1987.

23. Lyness JM, Noel TK, Cox C, King DA, Conwell Y, Caine ED. Screening for depression in elderly primary care patients. *Arch Intern Med*. 1997;157(4):449-454.

24. Donohue MC, Sperling RA, Salmon DP, et al. The preclinical Alzheimer cognitive composite. *JAMA Neurol*. 2014;71(8):961.

25. Joshi AD, Pontecorvo MJ, Clark CM, et al. Performance characteristics of amyloid PET with florbetapir F 18 in patients with Alzheimer's disease and cognitively normal subjects. *J Nucl Med*. 2012;53(3):378-384.

26. Yang FM, Jones RN. Center for Epidemiologic Studies - Depression scale (CES-D) item response bias found with Mantel-Haenszel method was successfully replicated using latent variable modeling. *J Clin Epidemiol*. 2007;60(11):1195-1200.

27. Jones RN. Differential item functioning and its relevance to epidemiology. *Curr Epidemiol Rep*. 2019;6(2):174-183.

28. Gelman A. Scaling regression inputs by dividing by two standard deviations. *Stat Med*. 2008;27(15):2865.

29. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates; 1988.

30. Singh-Manoux A, Dugravot A, Fournier A, et al. Trajectories of depressive symptoms before diagnosis of dementia: a 28-year follow-up study. *JAMA Psychiatry*. 2017;74(7):712-718.

31. Pietrzak RH, Lim YY, Neumeister A, et al. Amyloid-β, anxiety, and cognitive decline in preclinical Alzheimer disease a multicenter, prospective cohort study. *JAMA Psychiatry*. 2015;72(3):284-291.

32. Zwan M, Villemagne VL, Doré V, et al. Subjective memory complaints in APOE ε4 carriers are associated with high amyloid-β burden. *J Alzheimers Dis*. 2016;49(4):1115-1122.

33. Perin S, Harrington KD, Lim YY, et al. Amyloid burden and incident depressive symptoms in preclinical Alzheimer's disease. *J Affect Disord*. 2018;229:269-274.

34. Mulders PC, van Eijndhoven PF, Schene AH, Beckmann CF, Tendolkar I. Resting-state functional connectivity in major depressive disorder: a review. *Neurosci Biobehav Rev*. 2015;56:330-344.

35. Grothe MJ, Teipel SJ. Spatial patterns of atrophy, hypometabolism, and amyloid deposition in Alzheimer's disease correspond to dissociable functional brain networks. *Hum Brain Mapp*. 2016;37(1):35-53.

36. Verfaillie SCJ, Binette AP, Vachon-Presseau E, et al. Archival report subjective cognitive decline is associated with altered default mode network connectivity in individuals with a family history of Alzheimer's disease. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2018;3(5):463-472.

37. Galvin JE, Fu Q, Nguyen JT, Glasheen C, Scharff DP. Psychosocial determinants of intention to screen for Alzheimer's disease. *Alzheimers Dement*. 2008;4(5):353-360.

38. Romero HR, Welsh-Bohmer KA, Gwyther LP, et al. Community engagement in diverse populations for Alzheimer disease prevention trials. *Alzheimer Dis Assoc Disord*. 2014;28(3):269-274.

39. Petersen RC, Aisen PS, Beckett LA, et al. Alzheimer's Disease Neuroimaging Initiative (ADNI): clinical characterization. *Neurology*. 2010;74(3):201-209.

40. Kleineidam L, Guski J, Miebach L, et al. O3-11-05: Disentangling the Temporal Sequence of Subjective Cognitive Decline, Depressive Symptoms and Memory Deterioration: A Trivariate Latent Difference Score Analysis. *Alzheimer's Association International Conference 2019.* 2019;15(7S_Part_17):P913-P913.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.