



## OPEN ACCESS

## EDITED BY

Marianna Mazza,  
Agostino Gemelli University Polyclinic (IRCCS),  
Italy

## REVIEWED BY

Yared Belay,  
Monash University, Australia  
Jeanette Melin,  
Research Institutes of Sweden (RISE), Sweden

## \*CORRESPONDENCE

Anthony L. Vaccarino  
✉ avaccarino@indocresearch.org

RECEIVED 09 February 2023

ACCEPTED 27 April 2023

PUBLISHED 02 June 2023

## CITATION

Vaccarino AL, Black SE, Gilbert Evans S,  
Frey BN, Javadi M, Kennedy SH, Lam B,  
Lam RW, Lasalandra B, Martens E, Masellis M,  
Milev R, Mitchell S, Munoz DP, Sparks A,  
Swartz RH, Tan B, Uher R and Evans KR (2023)  
Rasch analyses of the Quick Inventory of  
Depressive Symptomatology Self-Report in  
neurodegenerative and major depressive  
disorders.

*Front. Psychiatry* 14:1154519.  
doi: 10.3389/fpsy.2023.1154519

## COPYRIGHT

© 2023 Vaccarino, Black, Gilbert Evans, Frey,  
Javadi, Kennedy, Lam, Lam, Lasalandra,  
Martens, Masellis, Milev, Mitchell, Munoz,  
Sparks, Swartz, Tan, Uher and Evans. This is an  
open-access article distributed under the terms  
of the [Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction  
in other forums is permitted, provided the  
original author(s) and the copyright owner(s)  
are credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted which  
does not comply with these terms.

# Rasch analyses of the Quick Inventory of Depressive Symptomatology Self-Report in neurodegenerative and major depressive disorders

Anthony L. Vaccarino<sup>1\*</sup>, Sandra E. Black<sup>2,3</sup>, Susan Gilbert Evans<sup>1</sup>,  
Benicio N. Frey<sup>4,5</sup>, Mojib Javadi<sup>1</sup>, Sidney H. Kennedy<sup>6</sup>,  
Benjamin Lam<sup>3</sup>, Raymond W. Lam<sup>7</sup>, Bianca Lasalandra<sup>1</sup>,  
Emily Martens<sup>1</sup>, Mario Masellis<sup>3</sup>, Roumen Milev<sup>8,9</sup>, Sara Mitchell<sup>3</sup>,  
Douglas P. Munoz<sup>9</sup>, Alana Sparks<sup>1</sup>, Richard H. Swartz<sup>2,3</sup>,  
Brian Tan<sup>10</sup>, Rudolf Uher<sup>11</sup> and Kenneth R. Evans<sup>1</sup>

<sup>1</sup>Indoc Research, Toronto, ON, Canada, <sup>2</sup>Dr. Sandra Black Centre for Brain Resilience and Recovery, Hurvitz Brain Sciences Program, Sunnybrook Research Institute, Toronto, ON, Canada, <sup>3</sup>Department of Medicine (Neurology), Sunnybrook Health Sciences Centre, University of Toronto, Toronto, ON, Canada, <sup>4</sup>Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada, <sup>5</sup>Mood Disorders Program and Women's Health Concerns Clinic, St. Joseph's Healthcare Hamilton, Hamilton, ON, Canada, <sup>6</sup>Department of Psychiatry, University of Toronto, Toronto, ON, Canada, <sup>7</sup>Department of Psychiatry, University of British Columbia, Vancouver, BC, Canada, <sup>8</sup>Departments of Psychiatry and Psychology, Queen's University, Providence Care, Kingston, ON, Canada, <sup>9</sup>Centre for Neuroscience Studies, Queen's University, Kingston, ON, Canada, <sup>10</sup>Rotman Research Institute, Baycrest Health Sciences, Toronto, ON, Canada, <sup>11</sup>Department of Psychiatry, Dalhousie University, Halifax, NS, Canada

**Background:** Symptoms of depression are present in neurodegenerative disorders (ND). It is important that depression-related symptoms be adequately screened and monitored in persons living with ND. The Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR) is a widely-used self-report measure to assess and monitor depressive severity across different patient populations. However, the measurement properties of the QIDS-SR have not been assessed in ND.

**Aim:** To use Rasch Measurement Theory to assess the measurement properties of the Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR) in ND and in comparison to major depressive disorder (MDD).

**Methods:** De-identified data from the Ontario Neurodegenerative Disease Research Initiative (NCT04104373) and Canadian Biomarker Integration Network in Depression (NCT01655706) were used in the analyses. Five hundred and twenty participants with ND (Alzheimer's disease or mild cognitive impairment, amyotrophic lateral sclerosis, cerebrovascular disease, frontotemporal dementia and Parkinson's disease) and 117 participants with major depressive disorder (MDD) were administered the QIDS-SR. Rasch Measurement Theory was used to assess measurement properties of the QIDS-SR, including unidimensionality and item-level fit, category ordering, item targeting, person separation index and reliability and differential item functioning.

**Results:** The QIDS-SR fit well to the Rasch model in ND and MDD, including unidimensionality, satisfactory category ordering and goodness-of-fit. Item-person measures (Wright maps) showed gaps in item difficulties, suggesting poor

precision for persons falling between those severity levels. Differences between mean person and item measures in the ND cohort logits suggest that QIDS-SR items target more severe depression than experienced by the ND cohort. Some items showed differential item functioning between cohorts.

**Conclusion:** The present study supports the use of the QIDS-SR in MDD and suggest that the QIDS-SR can be also used to screen for depressive symptoms in persons with ND. However, gaps in item targeting were noted that suggests that the QIDS-SR cannot differentiate participants falling within certain severity levels. Future studies would benefit from examination in a more severely depressed ND cohort, including those with diagnosed clinical depression.

#### KEYWORDS

neurodegenerative disorder, Rasch measurement theory, depressive symptoms, validity, QIDS-SR, major depressive disorder

## Introduction

Symptoms of depression are present across a broad range of neurodegenerative disorders (ND), and can negatively impact quality of life, functioning and progression of disease (1–6). Given the comorbidity of depression in ND and its relation to poorer outcomes, it is important that depression-related symptoms be adequately screened and monitored in persons living with ND. Depression in ND can be a challenge to identify, however, as some symptoms of depression overlap with the manifestation of ND-related signs and symptoms, such as difficulties in concentration, fatigue, restlessness, sleep-and appetite-related problem and feeling of being slowed down (7–10).

The Quick Inventory of Depressive Symptomatology Self-Report (QIDS-SR) (11) is a widely used self-reported, symptom-based rating scale that is aligned with DSM-IV criteria for major depressive disorder (MDD) and can therefore be a useful tool to assess and monitor depressive severity across different patient populations. The QIDS-SR was originally developed to assess depressive severity in MDD (11), and since then has been used across a broad range of diverse patient populations (12–19). However, its psychometric properties in ND have not been investigated.

Given the overlap of symptoms of depression and ND, it is important that the measurement properties of the QIDS-SR also be assessed in persons living with ND. In this context, we used Rasch Measurement Theory (RMT) (20–26) to assess the psychometric properties of QIDS-SR in ND. For comparison, RMT was also applied to QIDS-SR in a MDD sample, as the QIDS-SR was originally developed in that population (11). RMT considers the probability of an item's score as a function of both the person's individual trait level (depressive severity) and the item's difficulty (severity level where 50% of respondents will endorse the item). RMT provides fundamental criteria for objective scale measurement and determines how well the observed data approximates the Rasch measurement model. Items that do not fit the Rasch model are indications that they are measuring more than one construct, and thus possibly confounded by the presence of ND-related signs and symptoms. Rasch-based criteria assessed included unidimensionality and item-level goodness of fit, category ordering, item targeting, person separation index and reliability and item bias (differential item functioning) (22, 25, 26).

## Methods

### Study population and datasets

The present study used de-identified data collected as part of the Ontario Neurodegenerative Disease Research Initiative (ONDRI, NCT04104373) and Canadian Biomarker Integration Network in Depression (CAN-BIND trial 1, NCT01655706). These programs are part of the Ontario Brain Institute's Integrated Discovery Programs, which supports collaborative research networks focused on various brain conditions, including neurodevelopmental disorders, cerebral palsy, epilepsy, mood disorders, and neurodegenerative diseases (27–29). As part of the overall Integrated Discovery Programs, common data elements (CDEs) are collected across all participating programs as means of supporting cross-disease research, including the demographic information and QIDS-SR assessments used in the present analyses (6).

ONDRI is a prospective multi-site research program designed to characterize and track progression of neurodegenerative and neurovascular disorders (30, 31). Five hundred and twenty male and female participants who met criteria for one of the following ND participated in the study: Alzheimer's disease or amnesic single or multidomain mild cognitive impairment (AD/MCI,  $n=126$ ), amyotrophic lateral sclerosis (ALS,  $n=40$ ), cerebrovascular disease (CVD,  $n=161$ ), frontotemporal dementia (FTD,  $n=53$ ) and Parkinson's disease (PD,  $n=140$ ) (Table 1). Please see Farhan et al. (30) and Sunderland et al. (31) for full protocol details, inclusion and exclusion criteria.

TABLE 1 Demographic and clinical characteristics.

	ND	MDD
N	520*	177
AGE, YEARS $\pm$ SD (RANGE)	68.59 $\pm$ 7.72 (40–88 years)	35.69 $\pm$ 12.52 (18–61 years)
SEX, % FEMALE	33.5%	62.3%
QIDS-SR $\pm$ SD (RANGE)	4.81 $\pm$ 3.39 (0–19)	9.59 $\pm$ 4.83 (0–24)

\*ADMCI ( $n=126$ ), ALS ( $n=40$ ), CVD ( $n=161$ ), FTD ( $n=53$ ), PD ( $n=140$ ).

CAN-BIND is a multi-site research program designed to identify biomarkers of antidepressant response in MDD (32, 33). Two-hundred and eleven male and female outpatients between 18 and 60 years old whose symptoms met criteria for a major depressive episode in the context of MDD, according to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR), as determined by the Mini International Neuropsychiatric Interview and had a MADRS score of  $\geq 24$ , participated in the first CAN-BIND study (CAN-BIND-1). As the MADRS entry criteria in CAN-BIND-1 would in effect restrict the range of scores and skew the data toward higher scores at baseline, Rasch analyses were applied to Week 8 MDD participant data ( $n=177$ ) (see Table 1). Please see Lam et al. (33) for full protocol details, inclusion and exclusion criteria.

Studies were carried out in accordance with the Declaration of Helsinki and the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) guidelines, and the study designs and procedures were reviewed by the appropriate ethics committees; informed consent was obtained from all participants after full explanation of the nature of the procedures.

## Assessments

Demographic and QIDS-SR items are CDEs that were collected across all OBI-funded programs to support cross-disease comparisons and were used in the present analyses (6). All assessments were captured electronically on REDCap (Research Electronic Data Capture)<sup>1</sup> within the Brain-CODE informatics platform<sup>2</sup> (29).

QIDS-SR is a 16-item self-report measure that assesses the severity of depressive symptoms based on DSM-5 criteria for major depressive episode, with items scored on a 4-point scale (0–3) (11). Scoring of the QIDS-SR converts the 16 items into nine DSM domains (sad mood, concentration, self-criticism, suicidal ideation, interest, energy/fatigue, sleep disturbances and changes in appetite/weight). Because of challenges in providing immediate follow-up with those expressing suicidal ideation, item #12 assessing suicidality was omitted from the ONDRI protocol and therefore removed from all analyses (8 domains with total score ranging from 0–24). This was not expected to impact the scale's ability to discriminate depression, as removal of suicidality items in similar scales [for example, PHQ-9 (with suicide item) vs. PHQ-8 (without suicide item)] does not impact psychometric properties (34).

## Analyses

Demographic and clinical characteristics were calculated and compared across all cohorts; ANOVA was used for comparison of continuous variables (age) and chi-squared for comparison of categorical variables (sex). Analyses were performed using SPSS V27. A level of  $p < 0.05$  was regarded as statistically significant. RMT was used to assess scale performance across cohorts. Rasch analyses were carried out using WinSteps Version 5.1.

## Unidimensionality

The Rasch model assumes unidimensionality and local independence of items. Unidimensionality was examined by principal component analysis of the residuals derived from the Rasch model. The scale was considered unidimensional if  $>40\%$  of variance was explained by the measurement variable and unexplained variance of the first contrast accounts for  $<10\%$  or has an eigenvalue  $<2$  (35–37). Local dependency was assessed by examining residual correlations between items (38), with correlations of  $r < 0.4$  suggesting no consequential response dependency (36, 37).

## Item fit

Item-level goodness fit statistics were calculated as an index of how much the observed score for a given item within the scale deviates from the expected score of the Rasch model (i.e., are the items measuring a single underlying construct?). Items that did not fit the model (misfits) do not contribute to measurement of the underlying construct and likely add unwanted noise to the scale. Item “infit” mean square (MNSQ) values provide a fit index for each item that are in near proximity to the person's severity level, and “outfit” MNSQ values for differences between observed and expected values for items that are far from the person's severity level. MNSQ values between 0.5 and 1.5 were considered acceptable fit, with values between 1.5 and 2 considered to underfit the model and do not contribute to measurement of the underlying construct but do not distort the results, and values greater than 2.0 flagged as misfits that can distort the scale (39).

## Category threshold ordering

Category probability curves were generated to test the ordering of the response categories to determine whether participants can discriminate between the ordered response options, with fit values between 0.5 and 1.5 considered to be acceptable fit, and values greater than 2.0 flagged as misfitting (39).

## Item targeting

Person-item maps were constructed that plot individual participants and items on a single continuum to compare the range and position of the person measure distribution to that of the item measure distribution. Both item “difficulty” (i.e., level of depressive severity that item assesses) and person “ability” (i.e., level of depressive severity) are visualized together on a logit scale (i.e., log of the odds); with the right side of the map displaying the items from most difficult (top) to least difficult (bottom) and the left side plotting the individual participants, with those at the top having the highest trait level (symptom severity) and those at the bottom the lowest. The targeting of the scale is assessed by comparing mean person and mean item logit locations, with good measurement targeting evidenced when mean persons and items difficulty are in close proximity to one another (within 1 logit) (40, 41). The clinical utility of summing individual items from a scale to form a total score of overall severity requires that the items be spread out across the severity level of a broad range of persons. Gaps between items are considered problematic, as persons falling within those gaps cannot be differentiated from one another.

## Differential item functioning

Ideally, a scale should perform similarly regardless of the subgroup being assessed (i.e., generalizability of the scale). To assess any

1 [www.project-redcap.org](http://www.project-redcap.org)

2 [www.braincode.ca](http://www.braincode.ca)

potential item biases, differential item functioning (DIF) was used to determine whether items have any differences in item difficulty between subgroups of interest. That is, whether subgroups with similar levels of depression have the same probability of endorsing a given item. The existence of DIF was assessed for dichotomous categories based on sex (male vs. female) and cohort (ND vs. MDD), with mean differences in person measures of  $>0.43$  logits representing slight-to-moderate DIF and  $>0.64$  logits as indications of moderate-to-large DIF (with  $p < 0.05$  in Rasch-Welch test statistic) (40, 41).

### Reliability and separation index

Reliability of the QIDS-SR was evaluated using the person separation index and reliability derived from the Rasch model. The person separation index evaluates the scale in terms of its ability to distinguish participants into distinct levels of severity (strata), with person separation indices of  $>1.5$  considered acceptable (representing a minimum required to separate sample into 2 distinct strata, i.e., low and high). Person reliability is analogous to Cronbach's alpha as a measure of internal constancy among all scale items, with values of  $>0.70$  as indication of acceptable model fit (39).

## Results

Demographic characteristics are shown in Table 1. The MDD cohort was predominately female (62.3%); consistent with the higher rates of depression reported by females (42–44). By contrast, the ND cohort was predominantly male (66.5%). The bias toward male participants in the ND cohort is likely related to ONDRI enrollment strategy that required presence of a spousal study partner that may have influenced sample demographics (6, 31). As expected, QIDS-SR total scores were higher in MDD than ND ( $9.59 \pm 4.83$  and  $4.81 \pm 3.39$ , respectively). Please see (31) and (32) for comprehensive demographic and clinical characteristics of ND and MDD cohorts, respectively.

### Unidimensionality

Unidimensionality of the QIDS-SR was supported by PCA with variances explained by the measurement variables of 46.4 and 57.5% and unexplained variance of the first residual of 11.1% (eigenvalue of 1.66) and 7.6% (eigenvalue of 1.42) in MDD and ND, respectively. No consequential response dependency was found, with all residual correlations  $<0.4$ .

### Item fit

Item difficulty estimates and fit statistics are shown in Table 2. Item difficulty estimates ranged from  $-2.83$  logits (SLEEP) and  $1.04$  logits (SELF-PERCEPTION) for the ND cohort and from  $-1.47$  logits (SLEEP) to  $0.53$  logits (PSYCHOMOTOR) in the MDD cohort. All items showed acceptable goodness-of-fit statistics in the ND cohort, with MNSQs ranging from 0.64 to 1.50. In the MD cohort, most items showed acceptable goodness-of-fit, with minor underfitting noted for APPETITE (infit MNSQ = 1.57; outfit MNSQ = 1.61) and SELF-PERCEPTION (infit MNSQ = 1.55) (See Table 2).

## Category threshold ordering

Category probability curves for the QIDS-SR showed that scale options functioned in sequential order in capturing increasing levels of severity, with participants with higher levels of depression endorsing higher QIDS-SR options for both ND and MDD cohorts (Figure 1). Underfitting (infit MNSQ = 1.71; outfit MNSQ = 1.94) was noted for the most severe level (option 3) in the ND cohort, which is likely to be related the low endorsement of this option (5%), as compared with MDD (27%).

### Item targeting

Person-item locations (Wright maps) for the QIDS-SR in ND and MDD are shown in Figure 2. The proximity of the mean item measure (logit = 0) to the mean person measure (logit =  $-1.89$ ) in the ND cohort suggests that QIDS-SR items target a more severe depression than experienced by the ND cohort. This is not surprising given that the participants with a diagnosis of MDD were excluded from the study, as per ONDRI protocol (30). Furthermore, the gap of items between  $-2.83$  logits and  $0.17$  logits suggest poor precision for persons falling between those severity levels (See Figure 2). In the MDD cohort, the proximity of the mean item measure (logit = 0) to the mean person measure (logit =  $0.56$ ) indicates satisfactory item targeting in MDD. However, further inspection of Figure 2 shows a gap of items targeting persons between  $-1.46$  logits and  $0.03$  logits, suggesting poor precision for persons falling within that range of severity (see Figure 2). Please see Supplementary Table S1 for QIDS-SR raw scores and corresponding person measure logit estimates.

### Differential item functioning

As part of the ONDRI protocol, individuals with “unstable psychiatric illness defined as psychosis (hallucinations or delusions), lifelong history of major depression, or untreated late-onset major depression within 90 days of the screening visit” were excluded from the study (30). As a result, individuals with more severe psychiatric symptoms or a clinical diagnosis of MDD were excluded from the ONDRI study. Therefore, to assess DIF in ND and MDD participants within similar ranges of depression, we compared those with total QIDS scores ranging from 6–15, which represents mild-to-moderate levels of depression (MDD =  $9.72 \pm 2.50$ ,  $n = 114$ ; ND =  $8.54 \pm 2.37$ ,  $n = 159$ ). DIF between ND and MDD cohorts was noted for 4 of the 8 items, with SADNESS (DIF contrast =  $0.44$ ,  $p < 0.05$ ) and SELF-PERCEPTION (DIF contrast =  $0.55$ ,  $p < 0.05$ ) more likely to be endorsed in MDD than ND, and SLEEP (DIF contrast =  $-0.77$ ,  $p < 0.05$ ) and PSYCHOMOTOR (DIF contrast =  $-0.49$ ,  $p < 0.05$ ) more likely to be endorsed in ND than MDD (see Table 2). These differences suggest that ND participants presenting with depressive symptoms are more likely to report somatic-related problems (sleep and psychomotor) and less likely to report non-somatic symptoms of sadness and negative self-perception. This latter finding is consistent with ONDRI's exclusion of persons with an MDD diagnosis (30); a diagnosis that requires the



TABLE 2 QIDS-SR item difficulty estimates, fit statistics, and differential item functioning.

	ND			MDD			DIF
	Item difficulty logit	Infit MNSQ	Outfit MNSQ	Item difficulty logit	Infit MNSQ	Outfit MNSQ	DIF contrast** (ND vs. MDD)
SADNESS	0.68	0.99	0.95	0.09	0.63	0.71	0.44*
SLEEP	-2.83	1.09	1.27	-1.47	1.00	1.05	-0.77*
APPETITE	-0.17	1.37	1.34	0.06	1.57	1.61	-0.13
CONCENTRATION	0.26	0.88	0.90	0.10	0.65	0.65	0.19
SELF-PERCEPTION	1.04	1.50	1.02	0.33	1.55	1.42	0.55*
INTEREST	0.90	0.87	0.64	0.34	0.93	0.84	0.35
ENERGY	-0.05	0.79	0.77	0.03	0.71	0.69	0.00
PSYCHOMOTOR	0.17	0.85	0.69	0.53	0.96	1.00	-0.49*

\* $p < 0.05$  (Rasch-Welch statistic); \*\*QIDS-SR total scores ranging from 6–15 were used for DIF.

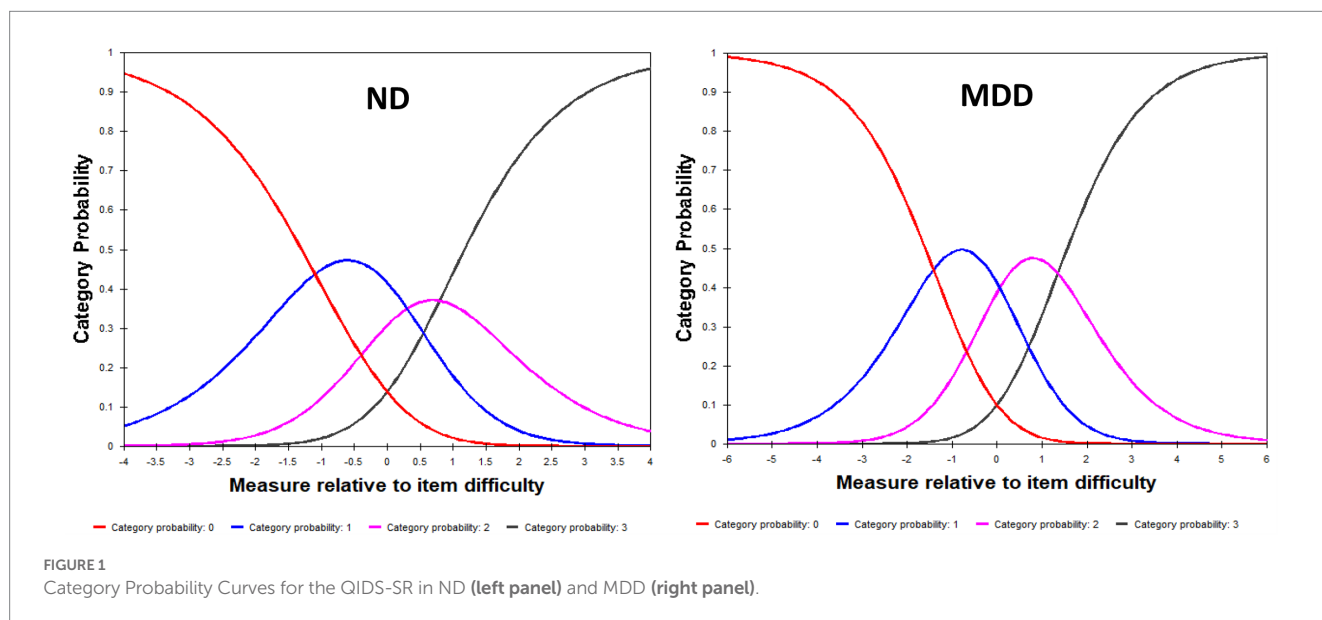


FIGURE 1 Category Probability Curves for the QIDS-SR in ND (left panel) and MDD (right panel).

presence of either depressed mood or loss of interest. No DIF was noted based on sex. As this was secondary use of data the impact of additional demographic and clinical factors was not examined, and limited sample sizes precluded subgroup analyses. Future studies would benefit from a better understanding of the generalizability of these results (i.e., ND-diagnosis) and impact of potential confounding factors (i.e., medication use, comorbidities, disease severity).

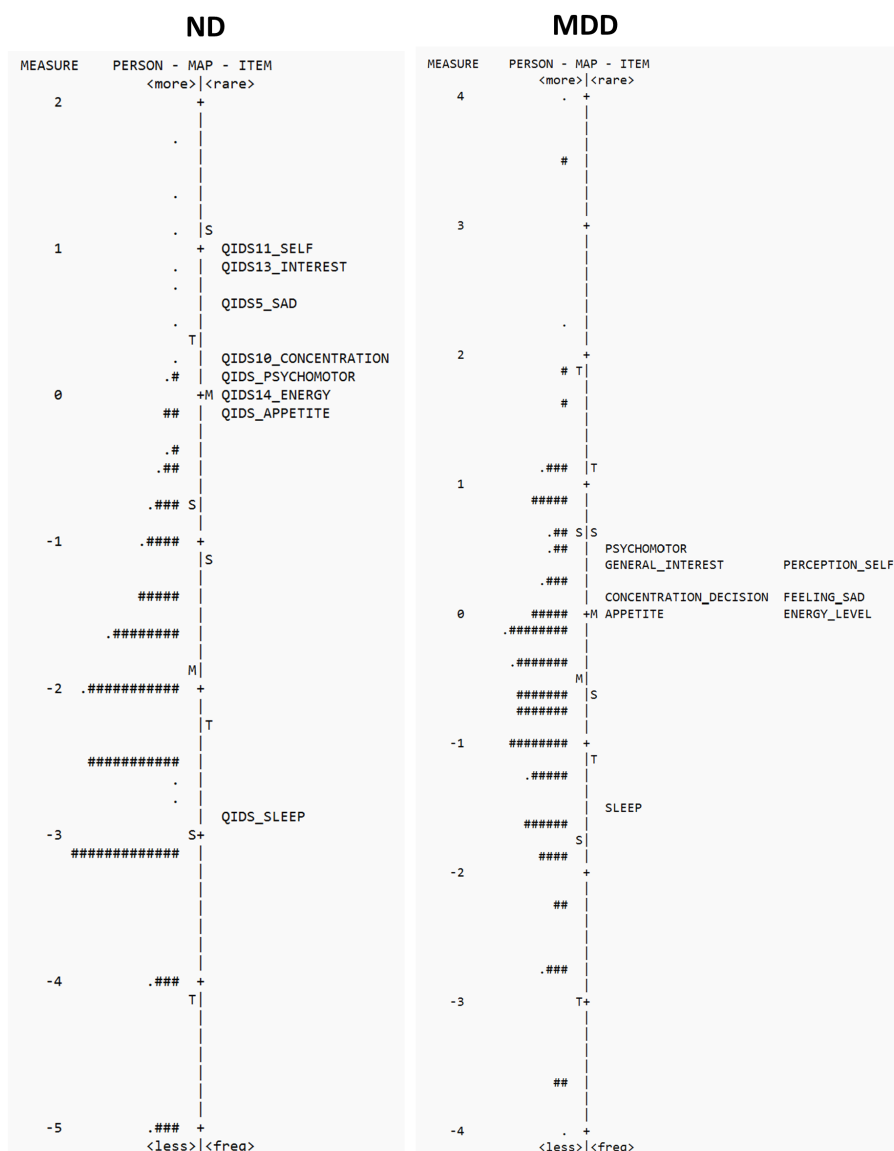
### Reliability and separation index

Person separation indices and reliability were acceptable in MDD, indicating that the QIDS-SR could stratify participants into two separate groups (1.86 and 0.78, respectively). By contrast, person separation indices and reliability were low in ND (1.28 and 0.62, respectively) indicating that the scale is not sensitive enough to separate the present ND cohort into different groups.

### Discussion

The present study used retrospective data pooled across two large cohort studies in persons with ND and MDD to evaluate the measurement properties of the QIDS-SR to assess symptoms of depression. Although the psychometric properties of the QIDS-SR have been well established in MDD and other populations (12–19), to our knowledge this is the first time its psychometric properties have been assessed in persons with ND.

In the ND cohort, the QIDS-SR showed favorable psychometric properties that support its construct validity, including unidimensionality and acceptable goodness-of-fit statistics, suggesting that all items contributed to the same underlying construct of depression. With respect to item difficulty, SLEEP was the least difficult/most endorsed item and SELF-PERCEPTION the most difficult/least endorsed item. The low person separation index and reliability, however, suggest that in its present form the QIDS-SR cannot distinguish different levels of severity in the ND sample used.



**FIGURE 2** Person-item location (Wright maps) for QIDS-SR items in ND (left panel) and MDD (right panel). The right side of the map displays the items from most difficult (top) to least difficult (bottom), and the left side plots the individual participants, with those at the top having the highest trait level (depression) and those that the bottom the least. M, mean difficulty; S, one standard deviation; T, two standard deviations.

This is further highlighted by the person-item map showing poor targeting at lower person severity levels (see Figure 2). This is not surprising given that the QIDS-SR is based on DSM symptom domain criteria for a major depressive episode and is therefore more targeted to a clinically depressed population. Additional studies are required therefore to assess the performance of the QIDS-SR in a broader ND population that includes individuals with a clinical diagnosis of depression.

Consistent with previous literature, the results of the present analysis provide further support for the utility of the QIDS-SR in assessing symptom severity in MDD (11, 16). With respect to item difficulty, SLEEP was the least difficult/most endorsed item and PSYCHOMOTOR the most difficult/least endorsed item. However, some shortcomings of the QID-SR in assessing depression in MDD were noted. Although person-item maps demonstrated satisfactory

item targeting between mean person and item measures in MDD, the lack of items between  $-1.46$  logits and  $0.03$  logits suggests poor measurement precision for persons falling between those lower severity levels (See Figure 2). The precision of the QIDS-SR would benefit, therefore, with the development of new items that target those severity levels (25). Incorporating Rasch analyses will aid in the development new items that target these gaps (25, 45). In addition, determining which symptoms might be relevant should also incorporate input from those affected and then tested in broader populations (46, 47) For example, previous studies have shown assessment of symptoms such as irritability, rumination and cognitive difficulties that are not assessed in the QIDS-SR can provide additional targeting across severity levels in MDD (25, 47). It is important to also note that the QIDS item #12 that assess thoughts of suicide was omitted from the analysis; and inclusion of this item in the assessment

provides additional targeting of those with higher levels of depression (25).

Underfitting was noted for the APPETITE item in MDD (see Table 2), thus calling into question the construct validity of this item and may add unwanted noise to the scale. Indeed, although changes in appetite are included in DSM-5 symptom criteria for MDD, studies have shown that appetite-related depression items show poor psychometric properties in MDD (25, 48), and appetite-related symptoms are variably expressed in MDD (49) and can be confounded by the presence of antidepressant side effects (50). It is also important to note that the QIDS-SR APPETITE item assesses both increases and decreases in appetite and weight as one item domain, whereas the MADRS, that was used to define the CAN-BIND-1 entry criteria, only assesses decreases in appetite. It is possible, therefore, that persons with these “atypical” reverse vegetative symptoms (increased appetite and weight gain) may have been underrepresented in the CAN-BIND cohort. All other items showed acceptable goodness-of-fit to the Rasch model, thus supporting their construct validity in MDD.

Person separation index and reliability (as well as Cronbach’s alpha) were satisfactory and suggest in the present MDD cohort two distinct levels of depression could be differentiated. Category probability curves were also favorable and demonstrate that QIDS-SR response option categories could be differentiated using the 4-point scale. With respect to item hierarchy and difficulty, SLEEP was considered the least difficult/most endorsed item and PSYCHOMOTOR the most difficult/least endorsed item.

DIF between ND and MDD is interesting and suggests that the presentation of depressive symptoms in ND may be more somatic in nature than MDD, with the ND cohort more likely to report PSYCHOMOTOR difficulties and problems with SLEEP, and less likely to report SADNESS and SELF-PERCEPTION (see Table 2). The reduced non-somatic symptoms in the ND cohort can be explained by ONDRI’s exclusion criteria of MDD diagnosis which would have effectively selected against non-somatic depression-related symptoms, such as sadness. It is also possible that non-motor symptoms can go underrecognized in ND, including blunted facial expression that could impact identification of mood-related symptoms (51). With respect to somatic symptoms of depression in the ND cohort, it could be argued that the assessment of depression in ND is confounded by the presence of overlapping ND-related signs and symptoms (8–10). However, the acceptable fit statistic and unidimensionality of the QIDS-SR in the ND cohort would argue against this interpretation and support its construct validity as a measure of depression in ND that is not confounded by the presence of ND-related signs and symptoms. Similarly, the PHQ-9, which is also based on DSM criteria for depression, has also been shown to be a valid scale in ND populations despite the presence of overlapping somatic symptoms (7, 10, 52). From a clinical perspective, these results suggest that the presentation of depression may be different between ND and MDD, and that clinicians explore somatic signs and symptoms in persons with ND as possible underlying depression.

It is important to note that age-matched controls were not included in this study, and thus we cannot rule out age-related changes in depression. Indeed, previous studies examining the relationship between age and presentation of depressive symptoms suggests that older adults with depression are more likely to report somatic

symptoms of depression than younger depressed adults (53, 54). Therefore, although the differences noted in the present study cannot exclude possible age-related effects, these results none-the-less highlight the importance of recognizing somatic presentation of depression in older adults with ND. This may be particularly important in developing therapeutic interventions in persons with ND and comorbid depression, as somatic symptoms of depression are associated with poorer QoL, worse treatment outcome and treatment resistance (55, 56).

In addition to differentiating ND- and depression-related symptoms, it’s also possible that age-related and cognitive difficulties could impact the ability to understand and elicit appropriate responses from participants. Although the present study did not assess the impact of ND-related difficulties on QIDS-SR responses, previous studies have shown that reliability and validity of patient-reported outcome measures are not impacted by cognition or age in PD (57) and the QIDS-SR has been well validated in aged populations (14). Furthermore, the ONDRI inclusion criteria required a minimum MOCA score, English fluency and visual ability, as part of the ONDRI protocol (30, 31), and participants were offered the option to have the QIDS-SR administered *via* a coordinator to allow clarification. However, although the ONDRI protocol would be expected to minimize the impact of age and ND-related symptoms, additional studies are required to better understand the impact ND-related difficulties on QIDS-SR responses.

In conclusion, the results of the present study are consistent with previous reports supporting the use of the QIDS-SR in MDD (11, 16) and suggest that the QIDS-SR may also be useful to screen for depressive symptoms in persons with ND. There are caveats, however, including item targeting issues that impacts scale’s ability to differentiate across different levels of severity. In addition to providing a framework to generate validation (i.e., unidimensionality and model fit), RMT also provides empirical evidence as to where scale performance can be improved to reduce noise and increase precision (25, 45). Visual examination of person-item maps, for example, shows that the QIDS-SR lacks precision across the full spectrum of severities, suggesting that more items are needed to target those falling within those gaps (25). Additional studies would also benefit from examination in a more severely depressed ND cohort, including those with diagnosed clinical depression.

## Data availability statement

Participants’ data used in this study are currently stored in the Brain-CODE Neuroinformatics Platform (<https://www.braincode.ca/>) managed by the Ontario Brain Institute. Requests to access these datasets should be directed to the Ontario Brain Institute at [info@braininstitute.ca](mailto:info@braininstitute.ca).

## Ethics statement

The studies involving human participants were reviewed and approved by all recruitment sites in accordance with the Governance Policy of Ontario Brain Institute as well the institutional policies. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

AV: study concept and design, statistical analysis, interpretation of data, and manuscript preparation. KE and RU: study design and interpretation of data. BE, SK, RWL, and RU: CAN-BIND recruiting site leads. SB, MM, DM, and RS: ONDRI recruiting site leads. SGE, MJ, B Lasalandra, EM, AS, and BT: data management and curation. All authors provided review of the manuscript for important intellectual content, read, and approved the final manuscript.

## Funding

This research was conducted with the support of the Ontario Brain Institute, an independent non-profit corporation, funded partially by the Ontario government. BEAM funding from Brain Canada, the Edwards Foundation and GE Healthcare for in kind support; funding was also received from Linda C. Campbell toward the BEAM study. CAN-BIND and ONDRI are Integrated Discovery Programs with support from the Ontario Brain Institute, with funding and/or in-kind support also provided by the investigators' universities and academic institutions. Additional funding for CAN-BIND was provided by CIHR, Lundbeck, Bristol-Myers Squibb, Pfizer, and Servier. The funders were not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

## Conflict of interest

RM has received consulting and speaking honoraria from AbbVie, Allergan, Eisai, Janssen, KYE, Lallemand, Lundbeck, Neonmind, Otsuka, and Sunovion, and research grants from CAN-BIND, CIHR, Janssen, Lallemand, Lundbeck, Nubiyota, OBI and OMHF. RWL has received honoraria for *ad hoc* speaking or advising/consulting, or

received research funds, from: Asia-Pacific Economic Cooperation, Bausch, BC Leading Edge Foundation, Brain Canada, Canadian Institutes of Health Research, Canadian Network for Mood and Anxiety Treatments, CAN-BIND Solutions, Carnot, Grand Challenges Canada, Healthy Minds Canada, Janssen, Lundbeck, Lundbeck Institute, Michael Smith Foundation for Health Research, MITACS, Myriad Neuroscience, Ontario Brain Institute, Otsuka, Pfizer/Viatris, Sanofi, Unity Health, Vancouver Coastal Health Research Institute, and VGH-UBCH Foundation. SK has received funding for Consulting or Speaking engagements from Abbvie, Boehringer-Ingelheim, Janssen, Lundbeck, Merck, Neurotorium, Otsuka, Pfizer, Sunovion, and Servier. He has received Research Support from Abbott, Brain Canada, CIHR (Canadian Institutes of Health Research), Janssen, Lundbeck, Ontario Brain Institute, Otsuka, SPOR (Canada's Strategy for Patient-Oriented Research). He has stock/stock options in Field Trip Health.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsy.2023.1154519/full#supplementary-material>

## References

- Galts CPC, Bettio LEB, Jewett DC, Yang CC, Brocardo PS, Rodrigues ALS, et al. Depression in neurodegenerative diseases: common mechanisms and current treatment options. *Neurosci Biobehav Rev.* (2019) 102:56–84. doi: 10.1016/j.neubiorev.2019.04.002
- Goodarzi ZS, Mele BS, Roberts DJ, Holroyd-Leduc J. Depression case finding in individuals with dementia: a systematic review and meta-analysis. *J Am Geriatr Soc.* (2017) 65:937–48. doi: 10.1111/jgs.14713
- Prisnie JC, Sajobi TT, Wang M, Patten SB, Fiest KM, Bulloch AGM, et al. Effects of depression and anxiety on quality of life in five common neurological disorders. *Gen Hosp Psychiatry.* (2018) 52:58–63. doi: 10.1016/j.genhospspsych.2018.03.009
- Roos E, Mariosa D, Ingre C, Lundholm C, Wirdefeldt K, Roos PM, et al. Depression in amyotrophic lateral sclerosis. *Neurology.* (2016) 86:2271–7. doi: 10.1212/WNL.0000000000002671
- Vaccarino AL, Sills T, Anderson KE, Endicott J, Giuliano J, Guttman M, et al. Assessment of day-to-day functioning in prodromal and early Huntington disease. *PLoS Currents.* (2011) 3:RRN1262. doi: 10.1371/currents.RRN1262
- Vaccarino AL, Beaton D, Black SE, Blier P, Farzan F, Finger E, et al. Common data elements to facilitate sharing and re-use of participant-level data: assessment of psychiatric comorbidity across brain disorders. *Front Psych.* (2022) 13:13. doi: 10.3389/fpsy.2022.816465
- Katzan IL, Lapin B, Griffith S, Jehi L, Fernandez H, Pioro E, et al. Somatic symptoms have negligible impact on patient health Questionnaire-9 depression scale scores in neurological patients. *Eur J Neurol.* (2021) 28:1812–9. doi: 10.1111/ene.14822
- Schrag A, Barone P, Brown RG, Leentjens AF, McDonald WM, Starkstein S, et al. Depression rating scales in Parkinson's disease: critique and recommendations. *Mov Disord.* (2007) 22:1077–92. doi: 10.1002/mds.21333
- Strober LB, Arnett PA. Assessment of depression in three medically ill, elderly populations: Alzheimer's disease, Parkinson's disease, and stroke. *Clin Neuropsychol.* (2009) 23:205–30. doi: 10.1080/13854040802003299
- Williams KG, Sanderson M, Jette N, Patten SB. Validity of the patient health Questionnaire-9 in neurologic populations. *Neurol Clin Pract.* (2019) 10:190–8. doi: 10.1212/CPJ.0000000000000748
- Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-item quick inventory of depressive symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry.* (2003) 54:573–83. doi: 10.1016/S0006-3223(02)01866-8
- Brown ES, Murray M, Carmody TJ, Kennard BD, Hughes CW, Khan DA, et al. The quick inventory of depressive symptomatology-self-report: a psychometric evaluation in patients with asthma and major depressive disorder. *Ann Allergy Asthma Immunol.* (2008) 100:433–8. doi: 10.1016/S1081-1206(10)60467-X
- Cameron IM, Crawford JR, Cardy AH, du Toit SW, Lawton K, Hay S, et al. Psychometric properties of the quick inventory of depressive symptomatology (QIDS-SR) in UK primary care. *J Psychiatr Res.* (2013) 47:592–8. doi: 10.1016/j.jpsychires.2013.01.019
- Doraiswamy PM, Bernstein IH, Rush AJ, Kyutoku Y, Carmody TJ, MacLeod L, et al. Diagnostic utility of the quick inventory of depressive symptomatology (QIDS-C16 and QIDS-SR16) in the elderly. *Acta Psychiatr Scand.* (2010) 122:226–34. doi: 10.1111/j.1600-0447.2009.01531.x
- Hedayati SS, Minhajuddin AT, Toto RD, Morris DW, Rush AJ. Validation of depression screening scales in patients with CKD. *Am J Kidney Dis.* (2009) 54:433–9. doi: 10.1053/j.ajkd.2009.03.016



16. Lamoureux BE, Linardatos E, Fresco DM, Bartko D, Logue E, Milo L. Using the QIDS-SR16 to identify major depressive disorder in primary care medical patients. *Behav Ther.* (2010) 41:423–31. doi: 10.1016/j.beth.2009.12.002
17. Liu R, Wang F, Liu S, Zhang Q, Feng Y, Sim K, et al. Reliability and validity of the quick inventory of depressive symptomatology—self-report scale in older adults with depressive symptoms. *Front Psych.* (2021) 12:12. doi: 10.3389/fpsy.2021.686711
18. Ma XR, Hou CL, Zang Y, Jia FJ, Lin YQ, Li Y, et al. Could the quick inventory of depressive symptomatology-self-report (QIDS-SR) be used in depressed schizophrenia patients? *J Affect Disord.* (2015) 172:191–4. doi: 10.1016/j.jad.2014.09.051
19. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The inventory of depressive symptomatology, clinician rating (IDS-C) and self-report (IDS-SR), and the quick inventory depressive symptomatology, clinician rating (QIDS-C) and self-report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med.* (2004) 34:73–82. doi: 10.1017/S0033291703001107
20. Andrich D. Rating scales and Rasch measurement. *Expert Rev Pharmacoecon Outcomes Res.* (2011) 11:571–85. doi: 10.1586/erp.11.59
21. Hendriks J, Fyfe S, Styles I, Skinner SR, Merriman G. Scale construction utilising the Rasch unidimensional measurement model: a measurement of adolescent attitudes towards abortion. *Australas Med J.* (2012) 5:251–61. doi: 10.4066/AMJ.2012.952
22. Johansson M, Preuter M, Karlsson S, Möllerberg M-L, Svensson H, Melin J. Valid and reliable? Basic and Expanded Recommendations for Psychometric Reporting and Quality Assessment. *OSF Preprints.* (2023). doi: 10.31219/osf.io/3htzc
23. Petrillo J, Cano SJ, McLeod LD. Using classical test theory, item response theory, and Rasch measurement theory to evaluate patient-reported outcome measures: a comparison of worked examples. *Value Health.* (2015) 18:25–34. doi: 10.1016/j.jval.2014.10.005
24. Tor E, Steketee C. Rasch analysis on OSCE data: an illustrative example. *Australas Med J.* (2011) 4:339–45. doi: 10.4066/AMJ.2011.755
25. Vaccarino AL, Kalali AH, Blier P, Gilbert Evans S, Engelhardt N, Foster JA, et al. The depression inventory development scale: assessment of psychometric properties using classical and modern measurement theory in a CAN-BIND trial. *Innov Clin Neurosci.* (2020) 17:30–40.
26. Boone WJ. Rasch analysis for instrument development: why, when, and how? *CBE life. Sci Educ.* (2016) 15:rm4. doi: 10.1187/cbe.16-04-0148
27. Stuss DT. The Ontario Brain Institute: completing the circle. *Canad J Neurol Sci.* (2014) 41:683–93. doi: 10.1017/cjn.2014.36
28. Stuss DT. From silos to systems: an integrated approach to neuroscience innovation. *Nat Rev Drug Discov.* (2015) 14:295–6. doi: 10.1038/nrd4615
29. Vaccarino AL, Dharsee M, Strother S, Aldridge D, Arnott SR, Behan B, et al. Brain-CODE: a secure neuroinformatics platform for management, federation, sharing and analysis of multi-dimensional neuroscience data. *Front Neuroinform.* (2018) 12:1–14. doi: 10.3389/fninf.2018.00028
30. Farhan SM, Bartha R, Black SE, Corbett D, Finger E, Freedman M, et al. The Ontario neurodegenerative disease research initiative (ONDRI). *Can J Neurol Sci.* (2017) 44:196–202. doi: 10.1017/cjn.2016.415
31. Sunderland KM, Beaton D, Arnott SR, Kleinstiver P, Kwan D, Lawrence-Dewar JM, et al. Characteristics of the Ontario neurodegenerative disease research initiative cohort. *Alzheimers Dement.* (2023) 19:226–43. doi: 10.1002/alz.12632
32. Kennedy SH, Lam RW, Rotzinger S, Milev RV, Blier P, Downar J, et al. Symptomatic and functional outcomes and early prediction of response to escitalopram monotherapy and sequential adjunctive aripiprazole therapy in patients with major depressive disorder: a CAN-BIND-1 report. *J Clin Psychiatry.* (2019) 80:18m12202. doi: 10.4088/JCP.18m12202
33. Lam RW, Milev R, Rotzinger S, Andreazza AC, Blier P, Brenner C, et al. Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. *BMC Psychiatry.* (2016) 16:105. doi: 10.1186/s12888-016-0785-x
34. Shin C, Lee SH, Han KM, Yoon HK, Han C. Comparison of the usefulness of the PHQ-8 and PHQ-9 for screening for major depressive disorder: analysis of psychiatric outpatient data. *Psychiatry Investig.* (2019) 16:300–5. doi: 10.30773/pi.2019.02.01
35. Mallinson T, Kozlowski AJ, Johnston MV, Weaver J, Terhorst L, Grampurohit N, et al. Rasch reporting guideline for rehabilitation research (RULER): the RULER statement. *Arch Phys Med Rehabil.* (2022) 103:1477–86. doi: 10.1016/j.apmr.2022.03.013
36. Kim S, Ouellet LJ, Mazza J, Spaulding AC. Rasch analysis and differential item functioning of a social support measure in jail inmates with HIV infection. *Eval Health Prof.* (2016) 40:33–60. doi: 10.1177/0163278716644954
37. Reckase MD. Unifactor latent trait models applied to multifactor tests: results and implications. *J Educ Stat.* (1979) 4:207–30. doi: 10.3102/10769986004003207
38. Yen WM. Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Appl Psychol Meas.* (1984, 1984) 8:125–45. doi: 10.1177/014662168400800201
39. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Meas Transact.* (1994) 8:370.
40. Boone WJ, Staver JR, Yale MS. *Rasch analysis in the human sciences.* Kindle ed. Berlin: Springer (2014).
41. Jiraniramai S, Wongpakaran T, Angkurawaranon C, Jiraporncharoen W, Wongpakaran N. Construct validity and differential item functioning of the PHQ-9 among health care workers: Rasch analysis approach. *Neuropsychiatr Dis Treat.* (2021) 17:1035–45. doi: 10.2147/NDT.S271987
42. Gater R, Tansella M, Korten A, Tiemens BG, Mavreas VG, Olatawura MO. Sex differences in the prevalence and detection of depressive and anxiety disorders in general health care settings: report from the World Health Organization collaborative study on psychological problems in general health care. *Arch Gen Psychiatry.* (1998) 55:405–13. doi: 10.1001/archpsyc.55.5.405
43. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry.* (1994) 51:8–19. doi: 10.1001/archpsyc.1994.03950010008002
44. Salk RH, Hyde JS, Abramson LY. Gender differences in depression in representative national samples: meta-analyses of diagnoses and symptoms. *Psychol Bull.* (2017) 143:783–822. doi: 10.1037/bul0000102
45. Peeters MJ, Augustine JM. Using Rasch measurement for instrument rating scale refinement. *Curr Pharm Teach Learn.* (2023) S1877-1297:00015–1. doi: 10.1016/j.cptl.2023.02.015. Epub ahead of print
46. United States Department of Health and Human Services. Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. *Fed Reg.* (2009) 74:65132–3.
47. Vaccarino AL, Evans KR, Kalali AH, Kennedy SH, Engelhardt N, Frey BN, et al. The depression inventory development workgroup: a collaborative, empirically driven initiative to develop a new assessment tool for major depressive disorder. *Innov Clin Neurosci.* (2016) 13:20–31.
48. Evans KR, Sills T, DeBrota DJ, Gelwicks S, Engelhardt N, Santor D. An item response analysis of the Hamilton depression rating scale using shared data from two pharmaceutical companies. *J Psychiatr Res.* (2004) 38:275–84. doi: 10.1016/j.jpsychires.2003.11.003
49. Maxwell MA, Cole DA. Weight change and appetite disturbance as symptoms of adolescent depression: toward an integrative biopsychosocial model. *Clin Psychol Rev.* (2009) 29:260–73. doi: 10.1016/j.cpr.2009.01.007
50. Fava M. Weight gain and antidepressants. *J Clin Psychiatry.* (2000) 61:37–41.
51. Poewe W, Luginger E. Depression in Parkinson's disease: impediments to recognition and treatment options. *Neurology.* (1999) 52:S2–6.
52. Chagas MHN, Crippa JAS, Loureiro SR, Hallak JEC, de Menezes-Gaya C, Machado-de-Sousa JP, et al. Validity of the PHQ-2 for the screening of major depression in Parkinson's disease: two questions and one important answer. *Aging Ment Health.* (2011) 15:838–43. doi: 10.1080/13607863.2011.569482
53. Drayer RA, Mulsant BH, Lenze EJ, Rollman BL, Dew MA, Kelleher K, et al. Somatic symptoms of depression in elderly patients with medical comorbidities. *Int J Geriatr Psychiatry.* (2005) 20:973–82. doi: 10.1002/gps.1389
54. Hegeman JM, Kok RM, van der Mast RC, Giltay EJ. Phenomenology of depression in older compared with younger adults: Meta-analysis. *Br J Psychiatry.* (2012) 200:275–81. doi: 10.1192/bjp.bp.111.095950
55. McIntyre RS, Konarski JZ, Mancini DA, Giacobbe P, Soczynska JK, et al. Improving outcomes in depression: a focus on somatic symptoms. *J Psychosom Res.* (2006) 60:279–82. doi: 10.1016/j.jpsychores.2005.09.010
56. Vaccarino AL, Sills TL, Evans KR, Kalali AH. Prevalence and association of somatic symptoms in patients with major depressive disorder. *J Affect Disord.* (2008) 110:270–6. doi: 10.1016/j.jad.2008.01.009
57. Gruber-Baldini A, Haesung K, Unick J, Shulman L. Impact of cognitive functioning and age on patient-reported outcomes in Parkinson's disease. *Innov Aging.* (2020) 4:156–7. doi: 10.1093/geroni/igaa057.510