Diagnostic accuracy for major depression in multiple sclerosis using self-report questionnaires

Anja Fischer[1,2], Marcus Fischer[3], Robert A. Nicholls[4], Stephanie Lau[2], Jana Poettgen[2], Kostas Patas[2], Christoph Heesen[2,5] & Stefan M. Gold[2,6,7]

[1]Department of Health Psychology, King's College London, SE1 9RT, London, UK
[2]Center for Molecular Neurobiology, Institute of Neuroimmunology and Multiple Sclerosis (INIMS), University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[3]Department of Pharmaceutical Chemistry, University of California San Francisco, San Francisco, California 94158
[4]Structural Studies Division, MRC Laboratory of Molecular Biology, Cambridge, CB2 0QH, UK
[5]Department of Neurology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[6]Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[7]Department of Psychiatry and Psychotherapy, Campus Benjamin Franklin (CBF), Charité Universitätsmedizin, Berlin, Germany

Abstract

Objective: Multiple sclerosis and major depressive disorder frequently co-occur but depression often remains undiagnosed in this population. Self-rated depression questionnaires are a good option where clinician-based standardized diagnostics are not feasible. However, there is a paucity of data on diagnostic accuracy of self-report measures for depression in multiple sclerosis (MS). Moreover, head-to-head comparisons of common questionnaires are largely lacking. This could be particularly relevant for high-risk patients with depressive symptoms. Here, we compare the diagnostic accuracy of the Beck Depression Inventory (BDI) and 30-item version of the Inventory of Depressive Symptomatology Self-Rated (IDS-SR$_{30}$) for major depressive disorder (MSS) against diagnosis by a structured clinical interview. **Methods:** Patients reporting depressive symptoms completed the BDI, the IDS-SR$_{30}$ and underwent diagnostic assessment (Mini International Neuropsychiatric Interview, M.I.N.I.). Receiver-Operating Characteristic analyses were performed, providing error estimates and false-positive/negative rates of suggested thresholds. **Results:** Data from $n = 31$ MS patients were available. BDI and IDS-SR$_{30}$ total score were significantly correlated ($r = 0.82$). The IDS-SR$_{30}$ total score, cognitive subscore, and BDI showed excellent to good accuracy (area under the curve (AUC) 0.86, 0.91, and 0.85, respectively). **Conclusion:** Both the IDS-SR$_{30}$ and the BDI are useful to quantify depressive symptoms showing good sensitivity and specificity. The IDS-SR$_{30}$ cognitive subscale may be useful as a screening tool and to quantify affective/cognitive depressive symptomatology.

Significant Findings

1 IDS-SR$_{30}$ and the BDI are valid measures for multiple sclerosis-associated major depression.

2 The IDS-SR$_{30}$ cognitive subscale may be suitable as a screening tool in MS depression.

3 The IDS-SR$_{30}$ covers all diagnostic criteria of MDD and is available in numerous languages and free of

charge, making it a particularly useful tool for depression screening in MS.

Limitations

1 Small sample size
2 Depression self-report questionnaires are not suitable for distinction between *different affective disorders*.

Introduction

Multiple sclerosis (MS) is an inflammatory, demyelinating disease of the central nervous system and is regularly accompanied by psychiatric symptoms such as depression (Feinstein et al. 2014). With a lifetime risk of up to 50% and a point prevalence of up to 25%, major depressive disorder (MDD) is a frequent comorbidity of MS (Patten et al. 2003). Multiple sclerosis-associated depression has a substantial negative impact on patients' quality of life, cognition, and psychosocial functioning (Hakim et al. 2000; Sa 2008). Higher levels of depressive symptoms are also linked to poorer treatment compliance (Ivanova et al. 2012), and thus can affect long-term health outcomes. If left untreated, depressive symptoms in MS may worsen over time (Ensari et al. 2014). Despite the high clinical relevance of depression in MS, it remains frequently underdiagnosed and undertreated.

The diagnostic criteria for MDD include a number of somatic and vegetative symptoms that overlap with typical symptoms of MS (e.g., fatigue, sleep disturbance, impaired concentration), which can make accurate MDD diagnosis particularly difficult in this patient population. Therefore, valid and reliable, easy-to-use diagnostic tools taking into account the potential confounding of MS symptoms are needed. Adjustment of cutoff scores may be required to prevent false diagnoses due to somatic-symptom-related score inflation. This is particularly important in patients who might be at risk for a comorbid mood disorder, for example, patients with elevated self-reported depressive symptoms.

A wide range of self-rated questionnaires are available for quantification of depression. Some of these have been validated and used in MS patients (see Avasarala et al. 2003; Benedict et al. 2003; Moran and Mohr 2005; Mohr et al. 2007; Honarmand and Feinstein 2009; Quaranta et al. 2012). Guidelines published by the American Academy of Neurology recommended only the BDI as well as a two-question tool to screen for depressive disorders with a weak level of evidence and did not find sufficient evidence for other instruments (Minden et al. 2014).

Importantly, only a few studies to date (Sullivan et al. 1995; Pandya et al. 2005; Honarmand and Feinstein 2009; Quaranta et al. 2012; Patten et al. 2015) have used a structured clinical interview to establish MDD diagnosis, and only the most recent ones also included Receiver-Operating characteristics (ROC) analysis, the gold standard to verify diagnostic accuracy. The Hospital Anxiety and Depression Scale (HADS) showed good diagnostic accuracy (Honarmand and Feinstein 2009), however, it only covers some of the diagnostic criteria of MDD. Moreover, it is copyrighted and may not be easily available, particularly for clinics or research groups in developing countries. A clinician-based, MS-specific depression scale (MSDRS (Quaranta et al. 2012)) also achieved good accuracy overall, however, it has relatively poor sensitivity (38%) and so far has only been used in Italian patients. Finally, a very recent paper demonstrated good accuracy of the patient health questionnaire PHQ-9, the Center for Epidemiologic Studies Depression rating scale (CES-D), and the HADS in MS (Patten et al. 2015). However, there is still paucity of data directly comparing different self-report questionnaires head-to-head and against structured interviews. No study to date has addressed this question in German-speaking MS patients.

The 30-item self-rated Inventory of Depressive Symptomatology (IDS-SR$_{30}$) was developed as part of the STAR*D trial (Rush et al. 1996) and has been validated for several patient populations with physical illness so far but not for MS. In contrast to most self-rated questionnaires for depression such as the Beck Depression Inventory (BDI) or the HADS, it assesses all symptom domains for MDD as designated in the DSM-IV and is available both in patient self-rating as well as clinician-based rating form. Moreover, it has been validated in more than 30 languages and is freely available (http://www.ids-qids.org/) without licensing charges. It also offers a self-rated validated 16-item short version (QIDS-SR) and subscales providing separate scores for cognitive and somatic symptoms that have been derived (Duivis et al. 2013). It might therefore be a promising tool to screen for and quantify depressive symptomatology in MS.

Aims of the Study

Here, we compare diagnostic accuracy of the BDI, the IDS, its subscales, and its short form (QIDS) in a group of German MS patients who reported elevated depressive symptoms. This sample might therefore model a clinical situation where detection of MDD is particularly important. We aim to establish meaningful threshold values based on a structured clinical interview.

Methods

Subjects

MS patients ($n = 31$) were recruited via the MS clinic of the University Medical Center Hamburg Eppendorf using our patient database and written consent prior to inclusion in the study was obtained. We contacted patients by mail if the scores from their last clinical visit recorded in the database indicated elevated depressive symptoms as measured by the Mood subscale of the Hamburg Quality of Life Questionnaire for MS (HAQUAMS) (Gold et al. 2001).

Diagnosis of major depression

Patients underwent structured diagnostic interviews by trained raters (A.F., S.L.) (The Mini International Neuropsychiatric Interview, M.I.N.I.) (Ackenheil et al. 1999). Several approaches have been proposed to implement DSM-IV diagnostic criteria in patients with physical illnesses: "aetiological" (case-by-case exclusion of somatic symptoms judged likely to be due to the comorbid medical illness), "inclusive" (use all symptoms regardless of etiology), and "substitutive" (substitution of most or all somatic symptoms with additional cognitive or affective symptoms). For the current study, we used the inclusive approach, that is, MDD diagnosis was made if a patient met at least five of the nine criteria which must include "depressed mood" or "loss of interest/anhedonia."

Self-report measures of depression

All patients completed the Beck Depression Inventory (BDI) (Hautzinger et al. 1995) and the IDS-SR$_{30}$ (Trivedi et al. 2004). Subscore calculation of the IDS-SR$_{30}$ included a somatic and a cognitive subscale as published by Duivis et al. (2013). The cognitive scale contains 10 IDS-SR$_{30}$ items, one for each of the following symptoms domains: Feeling sad or irritable, the quality of mood, concentration/decision making, self-perception, suicidal ideation, general interest, as well as capacity for pleasure excluding and including sexuality. The somatic subscale includes items on sleep, appetite, weight, energy level, psychomotor retardation/restlessness, and leaden paralysis/physical energy.

Statistics

Major depressive disorder diagnosis was established based on the M.I.N.I. (criterion). Receiver-Operating characteristics curves were created using MatLab and MedCalc software, giving an overview of sensitivity and specificity combinations for possible thresholds in each questionnaire. Error estimates and confidence intervals were calculated by bootstrapping using 1000 replications. Using MedCalc, the BDI, IDS-SR$_{30}$ total and somatic and cognitive subscore ROC curves were compared statistically using the method of DeLong et al. (1988) for the calculation of the Standard Error of the Area Under the Curve (AUC) and of the difference between two AUCs . This algorithm is particularly useful because it adjusts the AUCs for the expected frequency of the condition (MDD in this case) in the population of interest (in this case MS). Based on available epidemiological research (Patten et al. 2003), we estimated the MDD point prevalence in the MS population at 25%. AUC values were interpreted according to the following guidelines: 0.9–1 excellent, 0.8–0.9 good, 0.7–0.8 fair, 0.6–0.7 poor.

Cutoff values were established with the (0, 1) minimum distance method giving equal weight to sensitivity and specificity. Distributions of the thresholds as well as the false-positive and false-negative rates were determined to estimate uncertainty and control for the small sample size. Finally, BDI and IDS-SR$_{30}$ scores were correlated using Pearson correlation coefficients. All values are given as mean $\pm$ SEM. P-values of <0.05 were considered significant.

Results

Demography

Patients were aged 22–66 years old (M = 49.06 ± 1.89). About 75% of the participants were female ($n = 25$). Clinical and demographic characteristics can be found in Table 1.

Depression frequency and severity

Twenty-one of the 31 patients enrolled fulfilled the criteria of MDD according to M.I.N.I. interviews. As expected, most patients had also psychiatric comorbidities including other mood disorders (dysthymic disorder, $n = 4$; lifetime mania or hypomania, $n = 5$), anxiety disorders (generalized anxiety disorder, $n = 11$; agoraphobia with and without panic disorder, $n = 5$; social phobia $n = 5$, posttraumatic stress disorder, $n = 1$; OCD, $n = 1$), or substance abuse ($n = 2$).

As expected, patients with MDD scored well over usual cutoffs for clinical depression in the IDS-SR$_{30}$ as well as the BDI (Table 1). In addition, due to screening criteria for this patient group, IDS-SR$_{30}$ depression scores were also slightly elevated in the patients not meeting diagnostic criteria for MDD (Table 1). BDI and IDS-SR$_{30}$ showed a highly significant intercorrelation ($r = 0.82$, $P < 0.0001$, 95% CI 0.67–0.91).

Table 1. Clinical and demographic characteristics. Mean (M) and standard error of the mean (SEM) for the Beck Depression Inventory (BDI), total 30-item self-report Inventory of Depressive Symptomatology (IDS-SR$_{30}$), somatic (IDS som), and cognitive IDS subscale (IDS-SR$_{30}$cog) as well as the Quick IDS (QIDS-SR) 16-item short version. Multiple sclerosis patients were diagnosed as depressed major depressive disorder (MDD) and nondepressed (no MDD) via structured clinical interviews (The Mini International Neuropsychiatric Interview, M.I.N.I.).

	Depressed (meeting MDD criteria M.I.N.I) $n = 21$	Nondepressed (not meeting MDD criteria M.I.N.I) $n = 10$
Age	50.60 ± 2.15	48.41 ± 3.95
Education (years)	10.95 ± 0.34	11.10 ± 0.53
Sex (male/female)	4/17	2/8
Disease course (CIS/RRMS/SPMS/ PPMS/unclear)	3/12/2/3/1	0/3/6/1/0
BDI	21.8 ± 1.90	9.6 ± 2.13
IDS-SR30	37.5 ± 10.58	21.2 ± 3.29
IDS som	8.76 ± 0.62	5.5 ± 1.26
IDS cog	13.8 ± 10.96	5.4 ± 1.23
QIDS	15.47 ± 1.39	8.67 ± 1.96

CIS, clinically isolated syndrome; PPMS, primary progressive multiple sclerosis; RRMS, relapsing-remitting multiple sclerosis; SPMS, secondary progressive multiple Sclerosis. Data are presented as mean ± SEM.

Table 2. IDS-SR$_{30}$ sensitivity and specificity.

Cutoff	Sensitivity	95% CI	Specificity	95% CI
≥9	100.00	88.4–100.0	0.00	0.0–24.7
>15	100.00	88.4–100.0	38.46	13.9–68.4
>19	96.67	82.8–99.9	46.15	19.2–74.9
>21	96.67	82.8–99.9	61.54	31.6–86.1
>25	90.00	73.5–97.9	61.54	31.6–86.1
>26	83.33	65.3–94.4	69.23	38.6–90.9
>27	80.00	61.4–92.3	76.92	46.2–95.0
>33	53.33	34.3–71.7	76.92	46.2–95.0
>37	46.67	28.3–65.7	84.62	54.6–98.1
>38	40.00	22.7–59.4	92.31	64.0–99.8
>39	40.00	22.7–59.4	100.00	75.3–100.0
>59	0.00	0.0–11.6	100.00	75.3–100.0

Predictive value of the self-rated Inventory of Depressive Symptomatology (IDS-SR$_{30}$) for major depressive disorder: sensitivity, specificity and their 95% confidence intervals (CI) for potential cutoff values.

ROC analyses

All ROC-derived sensitivity and specificity values are shown in Table 2, and ROC curves are depicted in Fig. 1. The AUC derived from the ROC for the IDS-SR$_{30}$ indicated good accuracy (AUC = 0.86 ± 0.08). A cutoff of 28 (SD (IDS-SR$_{30}$_total) = 3.66) provides a sensitivity of
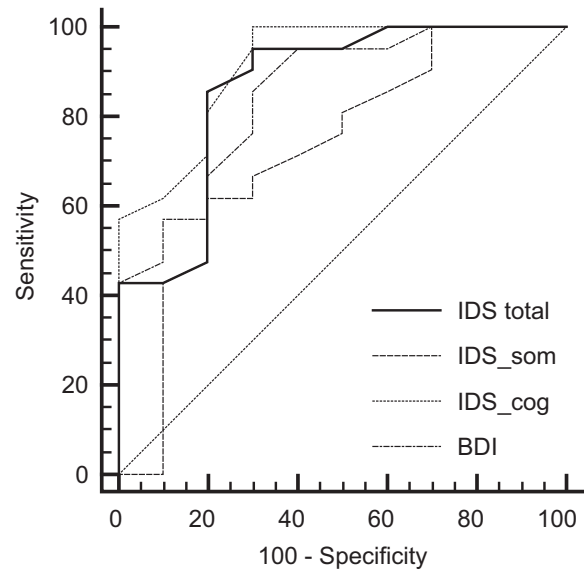

Figure 1. Receiver-Operating Characteristic (ROC) curves for the Beck Depression Inventory (BDI), the self-rated Inventory of Depressive Symptomatology (IDS-SR30) total score (IDS total) and the cognitive (IDS cog) and somatic (IDS som) subscores for predicting major depressive disorder based on structured clinical interviews (M.I.N.I).

80% and specificity of 77% (Table 2). The false-positive (negative) rate for IDS-SR$_{30}$ total when using 28 as the cutoff was estimated as 19.9 ± 7.3% (23.0 ± 12.0%). This results in a positive likelihood ratio of 5.67 and a negative likelihood ratio of 0.38. Furthermore, we determined diagnostic accuracy of the IDS-SR$_{30}$ cognitive and somatic subscales. The cognitive subscale reached excellent accuracy (AUC = 0.91 ± 0.06). For the IDS-SR$_{30}$ cognitive scale, the analysis yielded a cutoff value of 10 (sd(IDS-SR$_{30}$_cog) = 3.15, Table 3). The false-positive (negative) rate for the cognitive IDS-SR$_{30}$ subscale cutoff was estimated as 19.30 ± 7.74% (30.69 ± 13.32%), leading to a positive likelihood ratio of 4.25 and a negative likelihood ratio of 0.25. In contrast, the IDS-SR$_{30}$ somatic scale only showed fair accuracy (AUC = 0.72 ± 0.1). The QIDS-SR had good accuracy AUC of 0.80 ± 0.08 (CI 0.669–0.997) with a suggested cutoff of 13 (Sensitivity 66.67, Specificity 90.00).

Receiver-Operating characteristics analysis for the BDI revealed good accuracy (AUC = 0.85 ± 0.07) and a cutoff value of 12 (SD (BDI) = 3.69, Table 4). This cutoff yields Sensitivity of 88% and Specificity of 70%. The false-positive (negative) rate for the BDI with this cutoff was estimated as 12.48 ± 6.72% (30.15 ± 15.03%). For the BDI, we thus determined a positive likelihood ratio of 6.00 and a negative likelihood ratio of 0.43.

Comparison of AUC values for the IDS-SR$_{30}$ total score, IDS-SR$_{30}$ cognitive subscore, IDS-SR$_{30}$ subscale and

Table 3. IDS-SR30 cognitive subscale sensitivity and specificity.

Cutoff	Sensitivity	95% CI	Specificity	95% CI
≥1	100.00	86.8–100.0	0.00	0.0–24.7
>4	100.00	86.8–100.0	53.85	25.1–80.8
>5	96.15	80.4–99.9	61.54	31.6–86.1
>8	92.31	74.9–99.1	61.54	31.6–86.1
>9	80.77	60.6–93.4	69.23	38.6–90.9
>10	69.23	48.2–85.7	69.23	38.6–90.9
>11	53.85	33.4–73.4	76.92	46.2–95.0
>12	50.00	29.9–70.1	92.31	64.0–99.8
>14	34.62	17.2–55.7	92.31	64.0–99.8
>15	30.77	14.3–51.8	100.00	75.3–100.0
>24	0.00	0.0–13.2	100.00	75.3–100.0

Predictive value of the self-rated cognitive Inventory of Depressive Symptomatology subscale for major depressive disorder: sensitivity, specificity, and their 95% confidence intervals (CI) for potential cutoff values.

Table 4. BDI sensitivity and specificity.

Cutoff	Sensitivity	95% CI	Specificity	95% CI
≥0	100.00	85.8–100.0	0.00	0.0–30.8
>5	100.00	85.8–100.0	30.00	6.7–65.2
>7	95.83	78.9–99.9	40.00	12.2–73.8
>8	95.83	78.9–99.9	60.00	26.2–87.8
>11	87.50	67.6–97.3	70.00	34.8–93.3
>13	79.17	57.8–92.9	70.00	34.8–93.3
>14	70.83	48.9–87.4	80.00	44.4–97.5
>17	58.33	36.6–77.9	80.00	44.4–97.5
>19	58.33	36.6–77.9	90.00	55.5–99.7
>20	50.00	29.1–70.9	90.00	55.5–99.7
>21	45.83	25.6–67.2	100.00	69.2–100.0
>44	0.00	0.0–14.2	100.00	69.2–100.0

Predictive value of the Beck Depression Inventory (BDI) for major depressive disorder: sensitivity, specificity, and their 95% confidence intervals (CI) for potential cutoff values.

BDI yielded significant differences between IDS-SR30 total and IDS-SR30 somatic ($P = 0.02$) as well as IDS-SR30 cognitive and IDS-SR30 somatic ($P = 0.04$) while the difference between the BDI and IDS-SR30 somatic subscore failed to reach statistical significance ($P = 0.09$). There were no significant differences between the IDS-SR30 total score and the cognitive subscore ($P = 0.38$) as well as the BDI ($P = 0.80$).

Discussion

Our results indicate that two widely used patient-based instruments, the IDS-SR30 and the BDI, yield good accuracy for depression in MS when compared to a structured clinical interview. Moreover, we provide first evidence for validity of the IDS-SR30 total score, IDS-SR30 cognitive subscale, and the QIDS-SR short from for assessment of depression in MS.

Several studies have previously investigated psychometric properties of self-report depression questionnaires in MS. For the most part, analyses have been restricted to measures of reliability (such as internal consistency), correlational analyses with questionnaires measuring related concepts, or response to therapy (Nyenhuis et al. 1995; Sullivan et al. 1995; Avasarala et al. 2003; Benedict et al. 2003; Moran and Mohr 2005; Mohr et al. 2007; Honarmand and Feinstein 2009; Quaranta et al. 2012; Wang and Gorenstein 2013). However, a few have assessed diagnostic accuracy against a structured clinical interview: Mohr et al. (2007) demonstrated that two questions covering the two core symptoms of MDD (anhedonia and depressed mood) yield 99% sensitivity and 87% specificity. This approach is, therefore, highly accurate as a screening tool, although a more recent study reported lower estimates of specificity and sensitivity for this instrument (Patten et al. 2015). Moreover, it does not provide a quantitative score of depression severity. The 8-item depression subscale of the HADS (Honarmand and Feinstein 2009) was previously found to provide a sensitivity of 90% and a specificity of 87% for MDD in MS (as determined by the SCID). In this study, the authors also conducted a ROC analysis, which yielded an AUC of 0.94, which can be considered excellent. A recent study explored the diagnostic accuracy of the BDI in Italian MS patients against the SCID (Quaranta et al. 2012). Here, the AUC was 0.83 (good accuracy). The results from our study confirm the good accuracy of the BDI (AUC = 0.85), although we obtained markedly better sensitivity. We also provide first evidence that a comparatively new depression questionnaire, the IDS-SR30, also provides good accuracy when validated against a structured clinical interview.

The very recent study by Patten and colleagues provided the first available head-to-head comparison of self-report scales of depression in MS (Patten et al. 2015) and showed good accuracy for the CES-D, the PHQ (9 and 2), and the HADS. Since the PHQ is available free of charge, it might therefore be particularly interesting. With our study, there is now another freely available instrument (IDS) available for screening in MS depression. Moreover, our results also provide a direct comparison to the BDI, the only instrument that reached a sufficient level of evidence in the AAN guidelines.

Taken together, there are now several reliable and valid strategies for interested researchers and clinicians to screen for and quantify depression in MS, each with specific advantages and disadvantages. All scales evaluated to date (BDI, IDS-SR30, HADS, 2-question screen, PHQ, CES-D) show good sensitivity and specificity around 80%

or higher. The QIDS-SR, however, appears to be less sensitive but more specific. As noted in the AAN guidelines (Minden et al. 2014), "valid screening tools might improve identification of individuals who could benefit from further evaluation and treatment." If this is the goal, a low false-negative rate is required. In our study, the IDS-SR$_{30}$ had a markedly better false-negative rate (23%) compared to the BDI (30%). However, this still means that 23% of cases will be missed.

Clinically, a high false positive rate is less of a concern; it does however increase the administrative burden and may waste resources in particular settings such as primary care. For maximum sensitivity, specificity, and cost-effectiveness, the two-question approach proposed by David Mohr and colleagues might be the ideal choice. However, it does not yield a quantitative score of depression severity, which may be necessary in a research setting or to monitor treatment response in clinical care. The HADS provides a middle ground of a comparatively short scale offering both good accuracy for MDD diagnosis as well as a quantitative score. Generally, the HADS is a good measure for symptom severity in somatic, psychiatric, primary care patients and in the general population (Bjelland et al. 2002) and is therefore widely used. However, more recent work has revealed that it lacks consistent differentiation between symptoms of anxiety and depression (Cosco et al. 2012) and it does not cover all symptom domains of MDD.

The IDS-SR$_{30,}$ validated for the first time in MS patients in the current report, in our opinion has a number of features that make it a good option for measuring depression in MS: (1) it covers all DSM-IV criteria (and only those) (2) it offers parallel patient- and clinician-rated versions; (3) it was translated in many languages and is increasingly used; and (4) subscales for cognitive and somatic symptoms can be constructed (Duivis et al. 2013) as we have done in our present analysis and an algorithm for identification of DSM-assigned melancholic depression based on the items of the IDS-SR$_{30}$ is available (Khan et al. 2006). This might be particularly relevant for use in studies to explore novel biological substrates of depression in MS as these were found to differ between data-driven designations of melancholic and atypical idiopathic depression (Lamers et al. 2013). Similar dissociations between biological correlates and clinical features might also exist in MS-associated depression, as our previous research has indicated that affective and cognitive symptoms of depression in MS might be more closely related to neuroendocrine-limbic abnormalities (Gold et al. 2010, 2014) while vegetative/somatic aspects show closer correlations with markers of inflammation (Gold et al. 2011).

First applications in an RCT for a behavioral intervention (exercise) in MS also suggest that the IDS-SR$_{30}$ may

be responsive to detect changes in depressive symptomatology (Briken et al. 2014). Sensitivity to change remains an important issue for depression questionnaires in MS that have not systematically been addressed.

Some limitations have to be considered when interpreting the results from our present study. First of all, the sample size was small and all our patients were contacted because they had previously shown elevated depressive symptoms, that is, the sample was preselected for elevated levels of depression. On one hand, this sample might be a good model for clinical situations where accurate diagnosis is particularly important. On the other hand, in larger samples including many patients with very low or no depressive symptoms, diagnostic accuracy of IDS-SR$_{30}$ and BDI may be higher than reported here.

Despite finding the IDS-SR$_{30}$ somatic subscale to show only fair accuracy, the total IDS score was not found to perform significantly worse than the IDS-SR$_{30}$ cognitive subscale. This indicates that, while removal of somatic symptoms may be preferable, we found no evidence to suggest that it is strictly necessary for somatic symptoms to be removed from the IDS for diagnostic accuracy in MS. Future studies performed with a larger sample size will provide accurate/reliable estimates of the cutoff values. However, the specific values of the threshold estimates are not the most important results arising from this study. A far more meaningful and important result is the ability to provide estimates of the false-positive/negative rates for the various scores, given a particular score threshold. For example, we estimate the false-positive (negative) rate for IDS-SR$_{30}$_total as $19.9 \pm 7.3\%$ ($23.0 \pm 12.0\%$), noting that the provision of error estimates implicitly accounts for the small sample size. Pragmatically, these results are perhaps the most important results in the article, as they provide an estimate of the error rates that would be expected, should the particular cutoff value (in this case, 28) be used as the decision-making criterion.

Furthermore, the present study does not address the ability of the BDI or the IDS-SR$_{30}$ for differential diagnosis of MDD versus other affective disorders. In our sample, two patients with high scores on the BDI and IDS-SR$_{30}$ were found who did not meet diagnostic criteria for MDD according to the M.I.N.I. When looking at the M.I.N.I. data of these individuals, we observed that both met diagnostic criteria of dysthymia. This means that while the questionnaires have readily identified a mood disorder, they do not seem to be a means of distinguishing between MDD and dysthymia. This illustrates that distinction between *different affective disorders* may therefore be a particular challenge in MS that requires clinical interviews and cannot be achieved with general self-report questionnaires for depression.

In conclusion, both the IDS-SR$_{30}$ and the BDI are valid measures to quantify depressive symptoms and show good diagnostic accuracy. The IDS-SR$_{30}$ cognitive sub-scale may be useful as a screening tool and to quantify affective/cognitive depressive symptomatology.

Acknowledgments

Conflict of Interest

None declared.

References

Ackenheil, M., G. Stotz, and R. Dietz-Bauer. 1999. Mini International neuropsychiatric interview. German Version 5.0.0, DSM-IV. Psychiatrische Universitätsklinik München, München.

Avasarala, J. R., A. H. Cross, and K. Trinkaus. 2003. Comparative assessment of Yale Single Question and Beck Depression Inventory Scale in screening for depression in multiple sclerosis. Mult. Scler. 9:307–310.

Benedict, R. H., I. Fishman, M. M. McClellan, R. Bakshi, and B. Weinstock-Guttman. 2003. Validity of the Beck Depression Inventory-Fast Screen in multiple sclerosis. Mult. Scler. 9:393–396.

Bjelland, I., A. A. Dahl, T. T. Haug, and D. Neckelmann. 2002. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J. Psychosom. Res. 52:69–77.

Briken, S., S. Gold, S. Patra, E. Vettorazzi, D. Harbs, A. Tallner, et al. 2014. Effects of exercise on fitness and cognition in progressive MS: a randomized, controlled pilot trial. Mult. Scler. 20:382–390.

Cosco, T. D., F. Doyle, M. Ward, and H. McGee. 2012. Latent structure of the Hospital Anxiety And Depression Scale: a 10-year systematic review. J. Psychosom. Res. 72:180–184.

DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson. 1988. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845.

Duivis, H. E., N. Vogelzangs, N. Kupper, P. de Jonge, and B. W. Penninx. 2013. Differential association of somatic and cognitive symptoms of depression and anxiety with inflammation: findings from the Netherlands Study of Depression and Anxiety (NESDA). Psychoneuroendocrinology 38:1573–1585.

Ensari, I., R. W. Motl, E. McAuley, S. P. Mullen, and A. Feinstein. 2014. Patterns and predictors of naturally occurring change in depressive symptoms over a 30-month period in multiple sclerosis. Mult. Scler. 20:602–609.

Feinstein, A., S. Magalhaes, J. F. Richard, B. Audet, and C. Moore. 2014. The link between multiple sclerosis and depression. Nat. Rev. Neurol. 10:507–517.

Gold, S. M., C. Heesen, H. Schulz, U. Guder, A. Monch, J. Gbadamosi, et al. 2001. Disease specific quality of life instruments in multiple sclerosis: validation of the Hamburg Quality of Life Questionnaire in Multiple Sclerosis (HAQUAMS). Mult. Scler. 7:119–130.

Gold, S. M., K. C. Kern, M. F. O'Connor, M. J. Montag, A. Kim, Y. S. Yoo, et al. 2010. Smaller cornu ammonis 2-3/dentate gyrus volumes and elevated cortisol in multiple sclerosis patients with depressive symptoms. Biol. Psychiatry 68:553–559.

Gold, S. M., S. Kruger, K. J. Ziegler, T. Krieger, K. H. Schulz, C. Otte, et al. 2011. Endocrine and immune substrates of depressive symptoms and fatigue in multiple sclerosis patients with comorbid major depression. J. Neurol. Neurosurg. Psychiatry 82:814–818.

Gold, S. M., M. F. O'Connor, R. Gill, K. C. Kern, Y. Shi, R. G. Henry, et al. 2014. Detection of altered hippocampal morphology in multiple sclerosis-associated depression using automated surface mesh modeling. Hum. Brain Mapp. 35:30–37.

Hakim, E. A., A. M. Bakheit, T. N. Bryant, M. W. Roberts, S. A. McIntosh-Michaelis, A. J. Spackman, et al. 2000. The social impact of multiple sclerosis–a study of 305 patients and their relatives. Disabil. Rehabil. 22:288–293.

Hautzinger, M., M. Bailer, H. Worall, and F. Keller. 1995. Beck-depressions-inventar (BDI). Testhandbuch der deutschen Ausgabe. Huber, Bern.

Honarmand, K., and A. Feinstein. 2009. Validation of the Hospital Anxiety and Depression Scale for use with multiple sclerosis patients. Mult. Scler. 15:1518–1524.

Ivanova, J. I., R. E. Bergman, H. G. Birnbaum, A. L. Phillips, M. Stewart, and D. M. Meletiche. 2012. Impact of medication adherence to disease-modifying drugs on severe relapse, and direct and indirect costs among employees with multiple sclerosis in the US. J. Med. Econ. 15:601–609.

Khan, A. Y., J. Carrithers, S. H. Preskorn, R. Lear, S. R. Wisniewski, A. John Rush, et al. 2006. Clinical and demographic factors associated with DSM-IV melancholic depression. Ann. Clin. Psychiatry, 18:91–98.

Lamers, F., N. Vogelzangs, K. R. Merikangas, P. de Jonge, A. T. Beekman, and B. W. Penninx. 2013. Evidence for a differential role of HPA-axis function, inflammation and metabolic syndrome in melancholic versus atypical depression. Mol. Psychiatry 18:692–699.

Minden, S. L., A. Feinstein, R. C. Kalb, D. Miller, D. C. Mohr, S. B. Patten, et al. 2014. Evidence-based guideline:

assessment and management of psychiatric disorders in individuals with MS: report of the Guideline Development Subcommittee of the American Academy of Neurology. Neurology, 82:174–181.

Mohr, D. C., S. L. Hart, L. Julian, and E. S. Tasch. 2007. Screening for depression among patients with multiple sclerosis: two questions may be enough. Mult. Scler. 13:215–219.

Moran, P. J., and D. C. Mohr. 2005. The validity of Beck Depression Inventory and Hamilton Rating Scale for Depression items in the assessment of depression among patients with multiple sclerosis. J. Behav. Med. 28:35–41.

Nyenhuis, D. L., S. M. Rao, J. M. Zajecka, T. Luchetta, L. Bernardin, and D. C. Garron. 1995. Mood disturbance versus other symptoms of depression in multiple sclerosis. J. Int. Neuropsychol. Soc. 1:291–296.

Pandya, R., L. Metz, and S. B. Patten. 2005. Predictive value of the CES-D in detecting depression among candidates for disease-modifying multiple sclerosis treatment. Psychosomatics 46:131–134.

Patten, S. B., C. A. Beck, J. V. Williams, C. Barbui, and L. M. Metz. 2003. Major depression in multiple sclerosis: a population-based perspective. Neurology 61:1524–1527.

Patten, S. B., J. M. Burton, K. M. Fiest, S. Wiebe, A. G. Bulloch, M. Koch, et al. 2015. Validity of four screening scales for major depression in MS. Mult. Scler. pii: 1352458514559297.

Quaranta, D., C. Marra, M. Zinno, A. K. Patanella, M. J. Messina, C. Piccininni, et al. 2012. Presentation and validation of the multiple sclerosis depression rating scale: a test specifically devised to investigate affective disorders in multiple sclerosis patients. Clin. Neuropsychol. 26:571–587.

Rush, A. J., C. M. Gullion, M. R. Basco, R. B. Jarrett, and M. H. Trivedi. 1996. The Inventory of Depressive Symptomatology (IDS): psychometric properties. Psychol. Med. 26:477–486.

Sa, M. J. 2008. Psychological aspects of multiple sclerosis. Clin. Neurol. Neurosurg. 110:868–877.

Sullivan, M. J., B. Weinshenker, S. Mikail, and S. R. Bishop. 1995. Screening for major depression in the early stages of multiple sclerosis. Can. J. Neurol. Sci. 22:228–231.

Trivedi, M. H., A. J. Rush, H. M. Ibrahim, T. J. Carmody, M. M. Biggs, T. Suppes, et al. 2004. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol. Med. 34:73–82.

Wang, Y. P., and C. Gorenstein. 2013. Assessment of depression in medical patients: a systematic review of the utility of the Beck Depression Inventory-II. Clinics (Sao Paulo) 68:1274–1287.