Screening for Depression in Old Age With Very Short Instruments: The DIA-S4 Compared to the GDS5 and GDS4

Sonja Heidenblut, PhD[1] [ID] and Susanne Zank, PhD[1]

Abstract

Purpose of the Study: This article presents a short form of the Depression in old Age Scale with four items (DIA-S4). The diagnostic accuracy of the DIA-S4 was tested and compared to short forms of the Geriatric Depression Scale (GDS5, GDS4). **Methods:** Using the Montgomery and Asberg Depression Rating Scale (MADRS) as gold standard, the scales were validated with a sample of $N = 331$ geriatric inpatients. **Results:** The DIA-S4 had an internal consistency of .70, the GDS5 of .55, and the GDS4 of .58. The test efficiency considering ROC analyses for the DIAS-4 was AUC = .86, for the GDS5 AUC = .78, and for the GDS4 AUC = .74. The best cut-off score for the DIA-S4 was 1.5 with a sensitivity of 87% and a specificity of 68%, for the GDS4 1.5 with a sensitivity of 58% and a specificity of 81%, and for the GDS5 1.5 with a sensitivity of 88% and a specificity of 49%. **Conclusion:** Based on the data of this study, the DIA-S4 shows better psychometrical qualities than the GDS5 and the GDS4. It can be used as a very short screening scale for depression in old age in research and clinical practice.

Background

Depression is one of the most common psychiatric disorders in older adults (Horackova et al., 2019). In older populations, it is strongly associated with physical disability and a poor general health status (Xiang, 2018). For this reason, inpatients in geriatric healthcare units are especially at risk for depression, with prevalence rates ranging between 14% and 29% (Gantner et al., 2003; Helvik et al., 2012; McCusker et al., 2005). Thus, national guidelines recommend standard screening for this target group (DGPPN, 2015; NCCFMA, 2010).

Screening for Depression in Geriatric Healthcare Units

There are several challenges when it comes to detecting depression in geriatric healthcare units. The screening tools are intended to be used by medical staff who have no specific background knowledge in the field of psychiatric diagnoses. Therefore, the scales should have fixed response sets that make them easy to administer and to interpret. As geriatric inpatients often suffer from cognitive or functional impairments, screening instruments must also be suitable for verbal administration and be easy to understand and respond to. Thus, items should be short and simple and the response set should not be a rating scale but a simple "yes"/"no" alternative. Furthermore, somatic symptoms of depression like sleep disturbances or loss of appetite should not be included in the screening scales because these symptoms are very common among older people and not specifically related to depression (Hegeman et al., 2012). For this reason, many diagnostic instruments for depression are not suitable for the use with geriatric patients.

There are two depression scales which were especially designed for older people, namely the Geriatric Depression Scale (GDS) and the Depression in old Age Scale (DIA-S) (Heidenblut & Zank, 2010, 2014). Whereas the GDS is easy to use, it still contains some

[1]Chair of Rehabilitative Gerontology, University of Cologne, Cologne, Germany

Corresponding Author:
Sonja Heidenblut, Chair of Rehabilitative Gerontology, University of Cologne, Herbert-Lewin-Str. 2, Cologne 50931, Germany.
Email: Sonja.Heidenblut@uni-koeln.de

Table 1. Validation Studies on Very Short Versions of the Geriatric Depression Scale.

Authors	Year	Scale	Selected items	Validation sample	N	Cronbach's alpha	Validation standard	Best cut-off Score	Sensitivity	Specificity
D'Ath	1994	GDS4	1,3,6,7	Outpatients	120	.55	Geriatric Mental Status Schedule	0.5	93%	63%
D'Ath	1994	GDS4	1,3,6,7	Outpatients	120	.55	Geriatric Mental Status Schedule	1.5	61%	88%
van Marwijk	1995	GDS4	1,2,7,9	Outpatients	586	.64	Diagnostic Interview Schedule	1.5	67%	66%
Shah	1996	GDS4	1,3,6,7	Longterm care patients	42	—	—	0.5	85%	60%
Pomeroy	2001	GDS4	1,3,6,7	Rehabilitation patients –in and day care	87	.57	ICD-10 Diagnosis for depressive Episode	0.5	82%	67%
Rinaldi	2003	GDS5	1,4,8,9,12	Inpatients, outpatients, nursing home residents	181	—	DSMIV diagnosis	1.5	94%	81%
Chattat	2001	GDS5	1,4,8,9,12	Day care rehabilitation	126	.49	Clinical diagnosis	2.5	92%	77%

Note. Only studies with a diagnostic gold standard other than a longer version of the GDS and with interviewers blinded to test results are presented.

items that can lead to false positive test results when used in an inpatient setting (e.g., "Do you often get bored," "Do you prefer to stay at home rather than going out and doing new things"). The DIA-S was designed to avoid this kind of context bias. It consists of 10 items that focus mainly on the emotional, motivational, and cognitive symptoms of depression. In a first validation study, the scale is superior to the GDS15 in scale sensitivity, scale specificity, and internal consistency (Heidenblut & Zank, 2010). The DIA-S is recommended as a screening tool for depression within the context of geriatric assessments in Germany (Deutsche Gesellschaft für Geriatrie [DGG], 2019; Hautzinger et al., 2018). It has been translated into English, Korean, and Farsi (Heidenblut & Zank, 2014; Jung et al., 2019; Rashedi et al., 2016), and has been used in different clinical trials (Mätzold et al., 2019; Otto & de Wall, 2019; Sieske et al., 2019).

The Need for Very Short Screening Scales

Although the GDS15 and the DIA-S10 do not take a long time to administer, there are settings that require even shorter instruments. In geriatric health care units, the first step of a geriatric assessment is usually a multi-dimensional screening (DGPPN, 2015). In this step, depression and multiple other health conditions are assessed within one interview. As physicians must be able to conduct this assessment in a reasonable amount of time, each of the screening tools used within the interview must be very short. This also applies to research settings that measure depression among multiple other variables. Research interviews that take too much time can reduce the subjects' compliance considerably which can lead to higher dropout rates, missing values, and reduced validity of the data.

GDS Short Forms With Four or Five Items

The need for very short depression screenings in research and practice has led to the development of a number of different very short versions of the Geriatric Depression Scale (D'Ath et al., 1994; Hoyl et al., 1999; van Marwijk et al., 1995) (see Table 1 for details). Although these scales can be applied in a very short amount of time, their performance in test accuracy is considerably reduced because of their short format. There are two four-item versions and one five-item version of the scale. Each of these scales includes different items of the original GDS (D'Ath et al., 1994; Hoyl et al., 1999; van Marwijk et al., 1995) (see Table 1 for details). The items were selected for the respective scale due to their performance compared to different diagnostic gold standards. Theoretical aspects concerning the content of the scales were not considered. Table 1 shows an overview of validation studies of the scales. The internal consistency of the scales is low according to most studies (.49–.64), which indicates that the items on the scales do not measure the same constructs. Furthermore, according to most of the studies, the shortening of the scales has also led to either decreased scale sensitivity or decreased scale specificity (see Table 1 for details). For example, according to the study by Chattat et al. (2001), the scale efficacy of the GDS5 was good, but internal consistency

Table 2. The DIA-S and A Short Version of the Scale With Four Items.

Item number	Item	Answer	
1	**I am feeling down.**	*Yes*	**No**
2	I'm afraid that I might say or do the wrong thing.	Yes	No
3	I can relax easily.	Yes	*No*
4	My life seems to make little sense.	Yes	No
5	**It's hard to motivate myself.**	*Yes*	**No**
6	I'm anxious about the future.	Yes	No
7	**I can enjoy my life, even when things are sometimes more difficult.**	**Yes**	*No*
8	Difficulties tend to make me feel somewhat overwhelmed.	Yes	No
9	**I tend to worry a lot.**	*Yes*	**No**
10	Basically I am content with my life.	Yes	No

Note. The items in bold letters are part of the DIA-S4. The answer that scores with one point is printed in italics.

was only .49, which means that the outcomes of the scale were not reliable. According to the study by D'Ath et al. (1994), on the other hand, scale sensitivity was very good, but positive test outcomes were false in 37% of the cases, which reduces the usefulness in clinical practice considerably.

The Short Form of the Depression in Old Age Scale (DIA-S4)

For this study, we developed the DIA-S4 as an alternative to other very short screening instruments. Our aim was to create a tool that is as short and as easy to apply as the GDS4 or GDS5 versions, but that is more accurate in measuring depression in old age. The instrument is a shorter version of the Depression in old Age Scale (DIA-S) (Heidenblut & Zank, 2014) (see Table 2 for details).

Design and Methods

The Construction of the DIA-S4

For the DIA-S4, we selected items from the original scale considering their content and their results in the item analyses based on the original validation data (Heidenblut & Zank, 2010, 2014) (see Tables 2 and 4). We selected items with moderate difficulty, high discriminatory power and high prognostic effectiveness. To create a scale that is sensitive to different kinds of depression equally, we chose three items that focus on the key symptoms of depression (depressed mood, low energy, loss of pleasure) according to the ICD-10 (World Health Organization [WHO], 1992) and one item that focuses on rumination. We included one negatively worded item to avoid acquiescence bias.

Study Design and Data Collection

The study was approved by the institutional review board "Ethik Kommission Westfalen-Lippe und der Medizinischen Fakultät der Westfälischen Wilhelms-Universität Münster." Data for the validation of the scales was collected between 2007 and 2009 in three inpatient geriatric healthcare units in Germany. Only patients whose cognitive status was sufficient (MMSE score ≥15) and who did not suffer from aphasia, delirium, or psychotic disorders were allowed to participate in the study. All patients meeting the criteria were invited by staff members of the respective clinics to participate in the study. All participants provided written informed consent. In the first stage of data collection, all subjects who had given informed consents were included. Because participants with depression were more difficult to recruit than patients without depression, we carried out a second stage of data collection, during which staff members purposely recommended subjects with clinical signs of depression in order to increase the number of possibly ill subjects in the sample. In the end, we had a total of $n = 151$ subjects with depression subjects and $n = 181$ subjects without depression in our sample.

All subjects participated in two parts of the study. In one part, a clinical psychologist conducted a semi-structured psychiatric interview, referring to 10 depressive symptoms. Based on the outcome of this interview, the psychologist rated the subjects' depressive status on the Montgomery and Asberg Depression Rating Scale (MADRS). The subjects' MADRS scores were used as the gold standard for clinical depression. In the other part of the study, trained staff members applied the GDS15 and the DIA-S. In most cases, both screening scales were used within one interview, but in some cases, a second interview was required. In about half of the cases, the psychiatric interview was conducted first; in the other half, the screening scales were applied first. The interviewers who applied the screening scales were blind to the results of the psychiatric interview and vice versa in order to make sure that they were not biased by their assumptions about the subjects' depressive symptoms. The participating clinics provided data on the subjects' mental and functional statuses.

The DIA-S4, GDS4, and GDS5 were constructed on the basis of the longer forms as has been described by Weeks et al. (2003).

Table 3. Sample Characteristics.

	Diagnoses of depression		
	Yes (*n* = 180) *n* (%)	No (*n* = 151) *n* (%)	Total (*N* = 331) *n* (%)
Gender			
Male	45 (25%)	43 (29%)	88 (27%)
Female	135 (75%)	108 (72%)	243 (73%)
Status of depression			
No depression	133 (74%)		133 (74%)
Subthreshold	47 (26%)		47 (26%)
Mild	—	127 (84%)	127 (84%)
Moderate	—	19 (13%)	19 (6%)
Severe	—	5 (0.3%)	5 (2%)
	M (SD)	*M (SD)*	*M (SD)*
Age	81.0 (7.8)	80.3 (7.4)	80.7 (7.6)
Assessment MMSE	25.5 (3.5)	24.2 (4.0)	25.0 (3.7)
Barthel index	56 .3(26.8)	47.7 (22.3)	51.8 (24.7)

Measures

DIA-S4. The short version of the DIA-S consists of four items from the original scale (Heidenblut & Zank, 2010, 2014) that were chosen based on item statistics and content considerations.

GDS15. The 15-item short form of the Geriatric Depression Scale (Yesavage et al., 1986) is the most commonly used self-rating scale for geriatric depression. The maximum score is 15 points, while a score above five points serves as the cut-off for clinically relevant depression.

GDS4. The four-item short form of the GDS that is considered in this article was published by D'Ath et al. (1994). It represents the most commonly used very short depression scale in geriatric wards in Germany. The items for this short version were selected based on logistic regression analyses that used single items to predict a dichotomous "caseness" outcome of the GDS15 as the dependent variable.

GDS5. The five-item version of the GDS that is cited most often in literature was published by Hoyl et al. (1999). For this scale, correlations of the single items with clinical diagnoses as the gold standard criterion were considered when picking the five items with the strongest predictive power.

MADRS. The Montgomery and Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979; Neumann & Schulte, 1989) is an interview-based depression rating scale that has been used successfully in various populations of geriatric patients (Leentjens et al., 2000; Mottram et al., 2000). The entire score ranges from 0 to 60 points. In clinical practice, a score above 13 indicates a mild depressive episode, a score above 21 a moderate depressive episode, and a score above 28 a severe depressive episode. In the current study, the outcome of the scale serves as the gold standard criterion for depression, with participants being divided into depressed or nondepressed categories via the first cut-off at 13 points. The internal consistency of the scale was very good for the present sample (alpha = .86).

Cognitive impairment. The participants' cognitive impairment was measured with the Mini-Mental State Examination (AGAST, 1997; Folstein et al., 1975). The MMSE has a good internal consistency in geriatric inpatients (alpha = .79; Beyermann et al., 2013). It is recommended as a basic assessment instrument for in-patient geriatric healthcare units in Germany (DGG, 2019).

Functional impairment. The patients' functional impairment in basic activities of daily life was measured via the Barthel index (AGAST, 1997; Mahoney & Barthel, 1965). In a study with 2,634 geriatric inpatients (Lübke et al., 2004), the internal consistency of the scale was excellent (alpha = .90). It is recommended as a basic assessment instrument by the German Geriatrics Society (DGG, 2019).

Results

Sample

Sample characteristics are presented in Table 3.

The age and gender distribution of the sample is comparable to the population of geriatric inpatients in Germany as presented in epidemiological studies, whereas cognitive and physical functioning are slightly better due to the criteria for inclusion in the study (Renteln-Kruse & Ebert, 2003). Of the depressed subsample, 84% showed symptoms comparable to a mild depressive episode and 16% showed symptoms comparable to a moderate or severe

Table 4. Item Characteristics.

Item	Missing data N (%)	Difficulty (M)	Discriminatory power	Diagnostic effectiveness (%)
DIA-S4				
Feeling down	8 (2.4)	.48	.50	.71
Hard to motivate	7 (2.1)	.48	.44	.78
Can enjoy my life	18 (5.4)	.32	.47	.71
Tend to worry a lot	3 (0.9)	.55	50	.71
GDS4				
Basically satisfied	9 (2.7)	.29	.31	.68
Life is empty	13 (3.9)	.29	.45	.67
Afraid that something bad is going to happen	7 (2.1)	.35	.11	.59
Happy most of the time	31 (9.4)	.39	.47	.69
GDS5				
Basically satisfied	9 (2.7)	.28	.31	.68
Often get bored	2 (0.6)	.35	.32	.66
Often helpless	10 (3.0)	.54	.40	.63
Rather at home	10 (3.0)	.75	.09	.55
Feel pretty worthless	11 (3.3)	.47	.44	.69

Note. The prognostic effectiveness of the item was estimated as the percentage of cases that were correctly identified as depressed or not depressed by the respective item.

Table 5. Scale Efficacy.

	AUC	Cut-off score	Sensitivity	Specificity	Positive LR	Negative LR	Diagnostic Log Odd
DIA-S4	.857	**1.5**	**87%**	**68%**	**2.72**	**0.19**	**14.32**
		2.5	64%	88%	2.92	0.41	7.12
GDS4	.742	0.5	82%	49%	1.61	0.57	2.82
		1.5	**58%**	**81%**	**3.05**	**0.52**	**5.87**
GDS5	.778	**1.5**	**88%**	**49%**	**1.73**	**0.24**	**7.21**
		2.5	69%	76%	2.87	0.41	7.00

Note. The results of each scale's best cut-off score considering the Diagnostic Log Odd are printed in bold letters.

depressive episode. Among those with no current depressive disorder, 26% of the participants suffered from subclinical symptoms of depression (with an MADR score between 10 and 12 points).

Scale Characteristics

Item characteristics and scale reliability. Item characteristics are presented in Table 4. The percentage of missing data ranged from 0.6 to 9.4 for single GDS items and from 0.9 to 5.4 for items of the DIA-S. Whereas item difficulties for the DIA-S items were moderate (range: .32–.55), the level of difficulty was higher among the GDS4 items (.29–.39). For the GDS5 items, the difficulty varied considerably (.28–.75). The discriminatory power of the DIA-S4 items ranged between .44 and .55. Both versions of the GDS included one item with very low discriminatory power respectively (GDS4, Item3: .11, GDS5, Item4: .09). Prognostic efficiency, which was estimated as the rate of agreement between the dichotomous outcome of the MADR scale and the single item, ranged from .71 to .78 for the DIA-S items and from .55 to .69 for the GDS items.

The DIA-S4 had an internal consistency of .70, the GDS5 of .55, and the GDS4 of .58.

Effectiveness. For a comparison of the scales, we conducted ROC analyses to estimate scale sensitivities, scale specificities, and AUC scores. As a measure of the tests' usefulness for clinical decision making, we also estimated the positive likelihood ratio and the negative likelihood ratio. We used the Diagnostic Odd Ratio (DOR) to identify the best cut-off score for the scales. Table 5 gives an overview of the effectiveness of the scales in predicting the depressive status of the sample participants.

The best cut-off score of the DIA-S4 was between 1 and 2 with scores for sensitivity of 87% and a specificity of 68% (DOR = 14.32). For the GDS4, it was not easy to pick an ideal cut-off point, as a change of one point reduced either the sensitivity or the specificity of the scale considerably (see Figure 1 for the graph of the curve). The best cut-off score of the GDS4 was between 1 and 2 with a sensitivity of 58% and a specificity of 81% (DOR = 5.82).

The predictive efficiency of the GDS5 was best between 1 and 2 points with 88% sensitivity and 49% specificity for
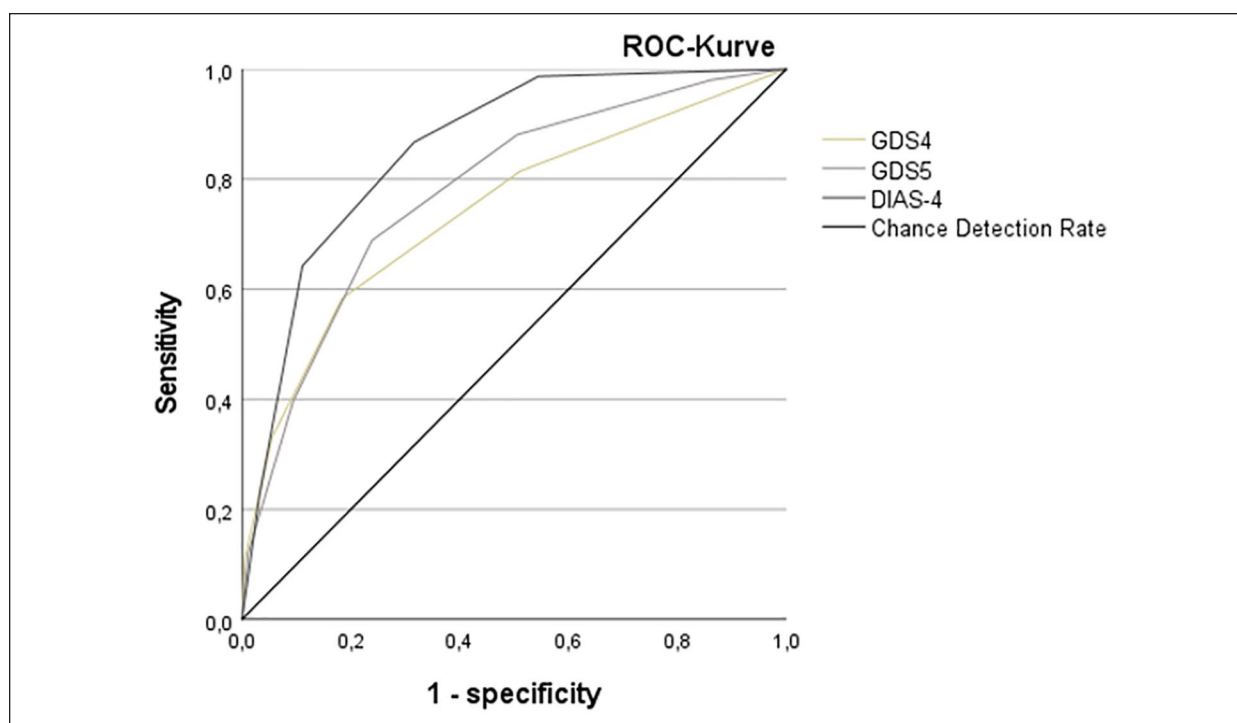
Figure 1. ROC curves of the DIA-S4, the GDS5, and the GDS4.

the total sample (DOR = 7.52). Comparing the scales based on their best cut-off scores, the DIA-S4 proved to be significantly more efficient than the GDS4 ($\chi^2_{(.049; 1, n=331)}$ = 2.867) and the GDS5 ($\chi^2_{(.002; 1, N=331)}$ = 9.006).

Discussion

Discussion of the Results

In this article, the diagnostic performance of three very short depression scales, namely the GDS5, the GDS4, and the DIA-S4, are discussed. The sample of the study included 331 geriatric inpatients. Among the depressed subsample ($n = 151$), 84% showed symptoms comparable to a mild depressive episode, 19% showed symptoms of a moderate depressive episode and 0.3% showed symptoms of a severe depressive episode. Among the sample group without clinical depression, 26% showed subclinical symptoms of depression. The proportion of mild depressive symptoms we found in our study is higher compared to prevalence studies with older medical inpatients (McCusker et al., 2005; Koenig, 1997). The spectrum of diseases in a sample has a considerable impact on the effectiveness of a screening scale within a sample. Screening instruments usually perform better in populations whose depressive status is easy to classify. Thus, if a large proportion of the sample either suffers from severe depressive symptoms or shows no symptoms at all, the results of the study can overestimate the diagnostic performance of a scale. This effect is known as spectrum bias (Dautzenberg et al., 2020). As both the DIA-S and the GDS were designed for use in a geriatric

setting rather than in a psychiatric setting, it is important to test the ability of the scales to correctly identify subjects that are more difficult to classify. We therefore chose a sample with a sufficient number of subjects with subclinical symptoms and mild depression.

Item and scale characteristics show better results for the DIA-S4 compared to the GDS4 and GDS5. The results for the GDS5 and GDS4 are very similar to those of former studies in scale reliability and are, especially in the case of the GDS4, lower in scale validity. One reason for this could be the large proportion of patients in this study who are difficult to classify. Another possible explanation is that in our study, we validated the GDS4 in an inpatient setting. This means that our sample suffered from more severe medical conditions than the outpatient participants or day care patients of former studies (D'Ath et al., 1994; Pomeroy, 2001), which probably had an effect on how the items were interpreted. For example, in our study, one item of the GDS4 ("Are you afraid that something bad is going to happen to you?") had a very low discriminatory power (.11). The reason for this may be that many patients in our sample had upcoming medical procedures or other health-related problems and interpreted the item in that way. As the DIA-S4 was designed to avoid this kind of bias, this did not have an effect on the discriminatory power of the DIA-S4 items (.44–.50).

For both GDS scales, it was difficult to find the best cut-off score, as an increase in sensitivity led to a significant decrease in specificity and vice versa. In a screening tool, sensitivity is usually more important than specificity. However, in the GDS5 and GDS4, cut-off

scores with a high sensitivity had very low specificity rates. For example, in the GDS5, the cut-off score of 1.5 had a high sensitivity (88%), but the specificity was only 49%. This indicates that a negative test result would be meaningless in clinical practice, as the scale would not perform reliably when it comes to correctly identifying non-depressed subjects. Furthermore, the internal consistency of both GDS scales were significantly lower than the internal consistency of the DIA-S4. This might be due to the fact that, whereas the short GDS scales were constructed based on item and scale statistics only, the DIA-S4 also considered theoretical aspects. The good internal consistency of the scale indicates that this approach was successful.

Implications for Clinical Practice

Based on the data of this study, the DIA-S4 proves to be a reliable screening tool for general practice. The instrument is easy to apply, and the interviewers reported that the subjects' acceptance of the scale was high. However, talking about one's depressive symptoms is a sensitive subject that requires mutual trust between the client and the interviewer. We therefore recommend the use of the DIA-S 4 within a broader diagnostic context. Furthermore, patients' compliance should be ensured by informing them about the assessment. The DIA-S4 can be applied by different healthcare practitioners. We recommend training interviewers in basic facts about test admission and depression in older subjects. In the case of a positive test result, further diagnostic investigation is required in order to assess the severity of the symptoms, suicidal tendencies, and relevant comorbidities. Patients should be informed about their diagnosis and, if applicable, of their treatment options.

Limitations and Future Prospects

In this validation study, the DIA-S4 was applied as part of a more extensive instrument. Applying the instrument by itself might lead to different results. In most validation studies, shortened versions of screening scales are applied within longer versions of the same scale. However, in clinical practice, patients might respond differently to very short depression screenings. Thus, it may take time for some patients to open up in regard to very personal questions. Therefore, it would be important to test the reliability and validity of shortened scales individually and under clinical conditions.

The results of the current study are based on the validation of the DIA-S in a German sample. This means that the English version of the scale as it is presented here has not yet been tested empirically. Thus, further data is required in order to gain an impression of the usefulness of the scale among English-speaking populations.

The items of the DIA-S4 were designed to be context-free so that the instrument can be useful in different populations (geriatric inpatients, geriatric outpatients,

community-dwelling older people). However, the scale has not yet been validated among these populations. Further research is required in order to assess how accurate the scale is in these different settings.

Conclusion

Very short depression scales can be useful tools for research and clinical practice. However, the shortening of a scale often reduces its reliability and validity considerably. In order to create consistent and reliable shorter scales, not only the psychometrical quality of the items should be considered but also the content of the items and their meaning within the scale. In this study, the DIA-S4 showed good psychometrical qualities and can be recommended as a screening tool for clinical practice and research.

ORCID iD

Sonja Heidenblut https://orcid.org/0000-0002-0280-7807

References

AGAST. (1997). *Geriatrisches Basisassessment: Handlungsanleitungen für die Praxis* [Geriatric Assessment: Guidelines for clinical practise]. Medizin Verlag.

Beyermann, S., Trippe, R. H., Bähr, A. A., & Püllen, R. (2013). Mini-Mental-Status-Test im stationären geriatrischen Bereich: Eine evaluation der diagnostischen qualität [Mini-Mental State Examination in geriatrics. An evaluation of diagnostic quality]. *Zeitschrift für Gerontologie und Geriatrie*, *46*(8), 740–747.

Chattat, R., Ellena, L., Cucinotta, D., Savorani, G., & Mucciarelli, G. (2001). A study on the validity of different short versions of the geriatric depression scale. *Archives of Gerontology and Geriatrics*, *7*, 81–86.

D'Ath, P., Katona, P., Mullan, E., Evans, S., & Katona, C. (1994). Screening, detection and management of depression in elderly primary care attenders. I: The acceptability

and performance of the 15 item Geriatric Depression Scale (GDS15) and the development of short versions. *Family Practice*, *11*(3), 260–266.

Dautzenberg, G., Lijmer, J., & Beekman, A. (2020). Diagnostic accuracy of the Montreal Cognitive Assessment (MoCA) for cognitive screening in old age psychiatry: Determining cutoff scores in clinical practice. Avoiding spectrum bias caused by healthy controls. *International Journal of Geriatric Psychiatry*, *35*(3), 261–269.

Deutsche Gesellschaft für Geriatrie. (2019). *S1-Leitlinie "Geriatrisches Assessment der Stufe 2" [S1 Guideline "Geriatric assessment step 2"]*. https://www.awmf.org/uploads/tx_szleitlinien/084-002l_Geriatrisches-Assessment-Stufe-2_2019-08_01.pdf

DGPPN. (2015). *Nationale Versorgungsrichtlinie Unipolare Depression [National Care Guideline for Unipolar Depression]*. https://www.awmf.org/leitlinien/detail/ll/nvl-005.html

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Gantner, a., Schubert, D., Wolf, S., & Creps, P. (2003). Screening for depression in a geriatric rehabilitation sample. *International Journal of Psychiatry in Medicine*, *33*(4), S.333–341.

Hautzinger, M., Hemmeter, U., Hirsbrunner, T., Holsboer-Trachsler, E., Leyhe, T. Mall, J.F., Mosimann, U., Rach, N., Trächsel, N. & Savaskan, E. (2018). Empfehlungen für diagnostik und therapie der depression im alter [Therapy recommendations for diagnosis and treatment of depression in old age]. *Praxis*, *107*(3), S.127–144.

Hegeman, J. M., Kok, R. M., van der Mast, R. C., & Giltay, E. J. (2012). Phenomenology of depression in older compared with younger adults: Meta-analysis. *The British Journal of Psychiatry*, *200*(4), 275–281.

Heidenblut, S, & Zank, S. (2010). Development of a new screening instrument for geriatric depression. The depression in old age scale (DIA-S). *Zeitschrift Für Gerontologie und Geriatrie*, *43*(3), 170–176.

Heidenblut, S., & Zank, S. (2014). Screening for depression with the Depression in Old Age Scale (DIA-S) and the Geriatric Depression Scale (GDS15): Diagnostic accuracy in a geriatric inpatient setting. *GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry*, *27*(1), 41–49.

Helvik, A-S., Engedal, K., Krokstad, S., Stordal, E., & Selbaek, G. (2012). A comparison of depressive symptoms in elderly medical inpatients and the elderly in a population-based health study (the Nord-Trondelag Health Study 3). *Nordic Journal of Psychiatry*, *66*, 189–197.

Horackova, K., Kopecek, M., Machu, V., Kagstrom, A., Aarsland, D., Bankovska Motlova, L., & Cermakova, P. (2019). Prevalence of late-life depression and gap in mental health service use across European regions. *European Psychiatry*, *57*, 19–25.

Hoyl, M. T., Alessi, C. A., Harker, J. O., Josephson, K. R., Pietruszka, F. M., Koelfgen, M., Mervis, J. R., Fitten, L. J., & Rubenstein, L. Z. (1999). Development and testing of a five-item version of the Geriatric Depression Scale. *Journal of the American Geriatrics Society*, *47*(7), 873–878.

Jung, Y. E., Kim, M. D., Bahk, W. M., Woo, Y. S., Nam, B., Seo, J. S., Jang, S-H., Sung, H-M., Shim, I-H., Yoon, B-H., Kim, J. S., & Kwon, Y. J. (2019). Validation of the Korean version of the depression in old age scale and comparison with other depression screening questionnaires used in elderly patients in medical settings. *Clinical Psychopharmacology and Neuroscience*, *17*(3), 369–376.

Koenig, H. (1997). Differences in psychosocial and health correlates of major and minordepression in medically ill older adults. *Journal of the American Geriatrics Society*, *45*(12), 1487–1495.

Leentjens, A. F., Verhey, F. R., Lousberg, R., Spitsbergen, H., & Wilmink, F. W. (2000). The validity of the Hamilton and Montgomery-Asberg depression rating scales as screening and diagnostic tools for depression in Parkinson's disease. *International Journal of Geriatric Psychiatry*, *15*(7), 644–649.

Lübke, N., Meinck, M., & von Renteln-Kruse, N. (2004). Der Barthel–Index in der Geriatrie. Eine Kontextanalyse zum Hamburger Einstufungsmanual. [The Barthel Index in geriatrics. A context analysis for the Hamburg Classification Manual]. *Zeitschrift für Gerontologie und Geriatrie, 37*, 316– 326.

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel Index. *Md State Med J, 14*(61–65), 61–65.

Mätzold, S., Geritz, J., Zeuner, K. E., Berg, D., Paschen, S., Hieke, J., . . . Maetzler, W. (2019). Functional movement disorders in neurogeriatric inpatients: Underdiagnosed, often comorbid to neurodegenerative disorders and treatable. *Zeitschrift für Gerontologie und Geriatrie, 52*(4), 324–329.

McCusker, J., Cole, M., Dufouil, C., Dendukuri, N., Latimer, E., Windholz, S., & Elie, M.(2005). The prevalence and correlates of major and minor depression in older medical inpatients. *Journal of the American Geriatrics Society*, *53*, 1344–1353.

Montgomery, S. a., & Asberg, M. (1979). A new depression scale designed to be sensitive to change. *The British Journal of Psychiatry*, *134*(4), 382–389.

Mottram, P., Wilson, K., & Copeland, J. (2000). Validation of the Hamilton depression rating scale and montgomery and asberg rating scales in terms of AGECAT Depression cases. *International Journal of Geriatric Psychiatry*, *60*, 1113–1119.

NCCFMA. (2010). *Depression in adults with a chronic physical health problem*. British Psychological Society.

Neumann, N.-U., & Schulte, R.-M. (1989). *Montgomery and Asberg Depression Rating Scale. Deutsche Fassung [Montgomery and Asberg Depression Ratin Scale. German Version]*. Perimed Fachbuch Verlagsgesellschaft.

Otto, F., & de Wall, S. (2019). Inpatient treatment for informal caregivers of relatives suffering from dementia. *Zeitschrift Fur Allgemeinmedizin*, *2019*(11), 462–467.

Pomeroy, I., Clark, C., & Philp, I.(2001). The effectiveness of very short scales for depressionscreening in elderly medical patients. *International Journal of Geriatric Psychiatry*,*16*, 321–326.

Rashedi, V., Rezaei, M., Foroughan, M., & Delbari, A. (2016). Validity and reliability of the depression in old age scale (DIA-S) in Iranian older adults. *Archives of Gerontology and Geriatrics*, *66*, 193–197.

Renteln-Kruse, W., & Ebert, D. (2003) Merkmale hospi-talisierter geriatrischer Patienten –Zwei Kohorten im Vergleich unter Verwendung des Screenings derArbe-itsgemeinschaft Geriatrisches Basisassessment (AGAST) [Characteristics of hospitalized geriatric patients – a comparison of two cohorts using the screening of the Arbeitsgemeinschaft Geriatrisches Basisassessment (AGAST)]. *Zeitschrift fürGerontologie und Geriatrie*, *36*, 223–232.

Sieske, L., Janssen, G., Babel, N., Westhoff, T. H., Wirth, R., & Pourhassan, M. (2019). Inflammation, appetite and food intake in older hospitalized patients. *Nutrients*, *11*(9), 1–10.

van Marwijk, H. W., Wallace, P., de Bock, G. H., Hermans, J., Kaptein, a a, & Mulder, J. D. (1995). Evaluation of the feasibility, reliability and diagnostic value of shortened versions of the geriatric depression scale. *The British Journal of General Practice : The Journal of the Royal College of General Practitioners*, *45*(393), 195–199.

Weeks, S. K., McGann, P. E., Michaels, T. K., & Penninx, B. W. J. H. (2003). Comparing various short-form Geriatric Depression Scales leads to the GDS-5/15. *Journal of Nursing Scholarship an Official Publication of Sigma Theta Tau International Honor Society of Nursing Sigma Theta Tau*, *35*(2), 133–137.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders : Clinical descriptions and diagnostic guidelines*. World Health Organization. https://apps.who.int/iris/handle/10665/37958

Xiang, X., Danilovich, M.K., Noth Tomasino, K., & Jordan, N. (2018). Depression prevalence and treatment among older home health services users in the United States. *Archives of Gerontology and Geriatrics*, *75*, 151–157.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1986). Development and validation of a geriatric depression screening scale: A pre-liminary report. *Journal of Psychiatric Research*, *17*(1), 37–49.