CLINICAL REHABILITATION

Clinical Rehabilitation
2017, Vol. 31(12) 1653–1663

DOI: 10.1177/0269215517708328
journals.sagepub.com/home/cre

Psychometric evaluation of the Signs of Depression Scale with a revised scoring mechanism in stroke patients with communicative impairment

Mariska J van Dijk[1], Janneke M de Man-van Ginkel[2,3], Thóra B Hafsteinsdóttir[1,2,3] and Marieke J Schuurmans[1,2,3]

Abstract

Objectives: To investigate (1) the diagnostic value of the Signs of Depression Scale (SODS) in a Likert scale format and (2) whether the Likert scale improves the diagnostic value compared with the original dichotomous scale.

Design: Cross-sectional multicentre study.

Setting: One general and one university hospital in the Netherlands.

Subjects: A total of 116 consecutive hospitalized stroke patients, of whom 53 were patients with communicative impairment.

Main measures: Depression was diagnosed with the Composite International Diagnostic Interview (CIDI) administered to the patients' relatives. The Barthel Index (BI) was used as an external validator.

Results: The correlation between the CIDI and the SODS-Likert or the SODS was small ($r_b = 0.18$), and the correlation between the Barthel Index and the SODS-Likert ($r_s = -0.30$) or the SODS ($r_s = -0.33$) was moderate. For both instruments, the discriminatory power for diagnosing depression when compared with the CIDI was best at a cut-off score of $\geqslant 2$. The internal consistency of the SODS-Likert was acceptable ($\alpha = 0.69$) and slightly higher than that of the SODS ($\alpha = 0.57$). The inter-rater reliability of the SODS-Likert and the SODS was acceptable (intraclass correlation coefficient (ICC) 0.66 and ICC 0.80, respectively). The clinical utility was rated good.

Conclusion: The diagnostic value of the SODS did not improve using a Likert scale format. However, the diagnostic value of the original dichotomous SODS is reasonable for the initial mood assessment of stroke patients with communicative impairment.

[1]Faculty of Health Care, University of Applied Sciences Utrecht, Utrecht, The Netherlands
[2]Julius Center for Health Sciences and Primary Care, Department of Nursing Science, University Medical Center Utrecht, Utrecht, The Netherlands
[3]Nursing Science, Program in Clinical Health Sciences, University Medical Center Utrecht, Utrecht, The Netherlands

Corresponding author:
Janneke M de Man-van Ginkel, Julius Center for Health Sciences and Primary Care, Department of Nursing Science, University Medical Center Utrecht, PO Box 85500, 3508 GA Utrecht, The Netherlands.
Email: J.M.deMan@umcutrecht.nl

Keywords
Screening, communication disorder, depression, psychometrics, stroke

Received: 14 October 2016; accepted: 13 April 2017

Introduction

Depression affects about one-third of patients after stroke,[1] and it significantly impacts their well-being and recovery.[2,3] There is growing evidence that treatment of depression decreases the incidence of symptoms and improves physical recovery.[2] Although guidelines recommend screening for depression,[4,5] it remains undiagnosed and under-treated.[2]

Communicative impairment, caused by language and cognitive deficits, is associated with depression after stroke.[4,5] An assessment of depression in these patients is difficult because instruments for the clinical diagnosis depend on the patient's ability to communicate.[6]

In our systematic review,[7] we identified a number of non-language-based instruments that were developed for use in patients with impaired communication, which are both 'observer-rated instruments' and 'self-assessment instruments'. Self-assessment instruments seem less suitable for use in patients with impaired communication because these require the ability to understand and respond to spoken language; patients have to report their mood on a visual analogue scale which assumes that patients are able to interpret the pictures and understand the scoring of the scale. Of the observer-rated instruments, such as the Aphasia Depression Rating Scale,[8] the Clinical Global Impression Scale,[9] the Signs of Depression Scale (SODS),[10] and the Stroke Aphasic Depression Questionnaire,[11] the SODS was recommended based on its feasibility in daily care.[7] Studies on the psychometric properties of the SODS, however, provide limited evidence because of (1) the exclusion of patients with communicative impairment, (2) the use of a suboptimal reference test, and (3) small sample sizes.[7,12] Furthermore, the dichotomous format of the SODS was criticized in our previous study as it might limit the ability of clinicians to differentiate between the severity of the symptoms.[13] The literature reports that yes/no answers lead to a loss of information and negatively influence the diagnostic accuracy of an instrument,[14,15] whereas a Likert scale may improve the performance of an instrument.[14]

Considering these limitations, along with the potential for usefulness in daily practice, we concluded that further research is needed to provide conclusive evidence for the diagnostic accuracy of the SODS. We aimed to improve the earlier findings by (1) including patients with communicative impairment to strengthen the methodological quality and generalizability of the study findings; (2) using a psychiatric interview, considered to be the gold standard,[2,12] as the reference test; (3) including an adequate sample size; and (4) modifying the SODS into a four-point Likert scale.

Based on this background, the aims of our study were to investigate (1) the diagnostic value of the SODS in a Likert scale format administered during daily care of patients with communicative impairment after stroke and (2) whether a Likert scale improves the diagnostic value in comparison with the original dichotomous SODS.

Methods

Design

We conducted a cross-sectional study in stroke patients in two hospitals in the Netherlands. A reference test strategy was used to diagnose major depressive disorder in all stroke patients, with and without communicative impairment (Figure 1); in case of a communicative impaired patient, a relative (partner, child, sibling, close friend) rated the patient's mood based on the Composite International Diagnostic Interview (CIDI), a standardized

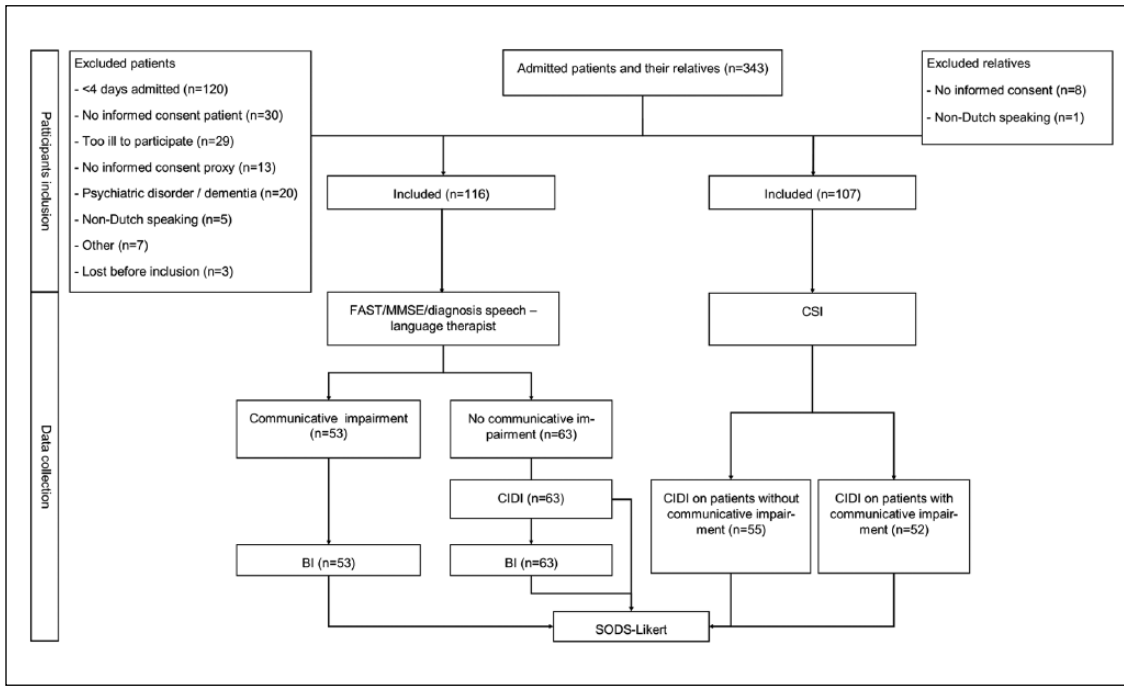

Figure 1. Flowchart of inclusion and data collection.
BI: Barthel Index; CIDI: Composite International Diagnostic Interview; CSI: Caregiver Strain Index; FAST: Frenchay Aphasia Screening Test; MMSE: Mini-Mental State Examination; SODS-Likert; Signs of Depression Scale-Likert.

psychiatric interview.[16] For patients who were able to communicate, this standardized psychiatric interview was conducted with them and their relatives. Additionally, the Barthel Index[17] was used as a non-language-based external validator. Physical disability, often assessed by the Barthel Index,[18] has been identified as a factor associated with depression after stroke.[2,3]

We aimed to achieve a sample size of at least 50 patients and/or their relatives in both groups, which is considered to be sufficient to achieve a good methodological quality.[19]

Ethical approval was obtained by the Medical Ethics Committee of the participating hospitals (11/550C).

Participants

We included patients with a diagnosis of stroke (cerebral haemorrhage or cerebral infarction) who stayed in the hospital for at least four days. This due to the fact that we focused on identifying depressive symptoms in the early stage in a sample reflecting the total spectrum of stroke patients. Exclusion criteria were (1) major psychiatric comorbidity, (2) use of antidepressant medication at stroke onset, or (3) too ill to participate, as determined by clinicians. Relatives were included if they spoke Dutch and were in contact with the patient at least once a week. All patients and relatives provided written informed consent; on behalf of the patients with communicative impairment, consent was obtained by a relative.

Outcome measures

The index test to identify depressive symptoms was the SODS in a Likert format. The original SODS is a six-item dichotomous scale with a yes/no response format and a score of 0–6; higher scores indicate more depressive symptoms. Prior studies have shown a good sensitivity and fair

specificity in studies with non-aphasic patients[13,20,21] and a fair sensitivity and specificity (cut-off score ⩾2) in a study with aphasic patients.[22]

We translated the SODS into Dutch using backward translation[23] and transformed it into a six-item Likert scale, with item scores ranging from 0 ('symptom not present'), 1 ('several days'), 2 ('more than half of the days') to 3 ('symptom present nearly every day'), and a maximum score of 18. This four-point Likert scale format and the item labels are according to the format of the Patient Health Questionnaire-9,[24] an instrument showing good psychometric properties within the stroke population.[25,26]

A major depressive disorder was diagnosed with the CIDI, a structured diagnostic interview for *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (DSM-IV) and International Classification of Diseases, Tenth Edition (ICD-10) psychiatric disorders.[16] The CIDI has shown good reliability and diagnostic agreement with the *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition, Revised (DSM-III-R) and the ICD-10 diagnosis of a major depressive disorder.[16]

To provide a diagnosis of depression in patients with communicative impairment, the CIDI was administered to the relatives on behalf of these patients. To be able to use the CIDI-relative with confidence, we compared ratings of patients without communicative impairment (CIDI-patient) and their relatives to check whether these rating sufficiently correlated, thereby indicating that relatives provide a reliable diagnosis of patients depression using the CIDI.

The Barthel Index measures the physical disability after stroke on a score of 0–20 (lower scores indicating more dependence)[17] which was used as an external validator. The Barthel Index shows good internal consistency, validity, and reliability.[18]

Both a positive correlation between the SODS-Likert and the CIDI-relative (concurrent validity) and a negative correlation between the SODS-Likert and the Barthel Index (divergent validity) are considered part of the diagnostic value and represent the basis for calculating the diagnostic accuracy of the SODS-Likert.

The clinical utility of the SODS-Likert was measured with a dichotomous questionnaire in which nurses indicated item clarity and administration time required.[27]

Communicative ability was assessed with the shortened Frenchay Aphasia Screening Test (FAST)[28] and the Mini-Mental State Examination (MMSE).[29] The shortened FAST consists of the subscales 'comprehension' and 'verbal expression' and excludes the sections 'reading' and 'writing'. The cut-off scores of the FAST depend on age: a score <17 for patients <60 years of age, a score <16 for patients ⩾61 years and ⩽70 years, and a score <15 for patients ⩾71 years indicate communicative impairment. The maximum score is 20.[28] The MMSE has a maximum score of 30.[29] A cut-off score of ⩽18 indicates cognitive impairment.[26] In case of communicative impairment based on the FAST and/or the MMSE, only the patient's relative administered the CIDI. If no communicative impairment was found, both patient and relative administered the CIDI.

Caregiver burden was measured with the Caregiver Strain Index (CSI)[30] because of a potential bias of a relative's rating of a patient's depression being influenced by his or her own perceived burden.[31] The CSI has 13 items with yes/no answers, a maximum total score of 13, and a cut-off score of ⩾7 for 'considerable strain'.[30]

Procedure

The researcher (M.J.v.D.) visited both hospitals daily to recruit patients and relatives between day one and four of the patient's admission, assessed whether the patient met the inclusion criteria, asked patients and relatives for informed consent, and collected patients' sociodemographic and clinical data and relatives' sociodemographic data.

Different pairs of nurses independently rated the patient's mood using the SODS-Likert after receiving verbal and written instruction by the researcher on how to administer the instrument. The instruction emphasized that nurses should rate the instrument based on their observations during daily care representing their perception of patient's mood and that they should not discuss the rating with colleagues

prior to completion. Ratings of nurses taking care of the patient for several days as well as ratings of nurses who took care of the patient for the first day were included because this reflected the usual daily care in hospital. The researcher administered the CIDI to patients who were able to communicate and to the relatives on behalf of all patients, regardless of their communicative abilities. The researcher was trained to administer the CIDI by the principal investigator (J.M.d.M.G), experienced in administering the CIDI in patients after stroke. Nurses were asked to complete the SODS-Likert at the same time the researcher assessed the CIDI. However, the administration of the SODS-Likert was not always completed immediately after the instrument was provided by the researcher, mainly due to other activities that had to be performed. To prevent bias, in all cases, the researcher and the nurses were blinded to each other's ratings.

Before administering the CIDI, patient's cognitive and communicative functioning was assessed by the researcher using the FAST and the MMSE and relative's perceived caregiver burden using the CSI. Patients and their relatives were blinded to each other's responses.

The researcher rated the patient's functional disability with the Barthel Index, based on information from the patient or the nurse, at the same time as the CIDI.

Nurses completed the clinical utility questionnaire after the first administration of the SODS-Likert.

All assessments were conducted from the fourth day of patient's admission to the hospital. This time period enabled nurses to observe patient's mood, which is important because the SODS is an observer-rated instrument. Data were collected between September 2013 and April 2014.

Statistical analysis

Correlation between the CIDI-patient and the CIDI-relative was measured with Cramer's V, resulting in a correlation of $V=0.62$ ($P<0.001$). A correlation of $\geqslant 0.50$ is considered large;[32] therefore, we concluded that the CIDI-relative could be used as a reference test.

The CIDIs rated positive by patients were also rated positive by their relatives, regardless of their perceived caregiver burden. Furthermore, the proportion of relatives of depressed patients who reported caregiver burden was smaller than the proportion of those without caregiver burden. This indicated that caregiver burden did not influence relatives' ratings of patients' depression, and adjustment of the outcome of the CIDI-relative for caregiver burden was not indicated.

Descriptive statistics were used to summarize the baseline characteristics of the participants.

The convergent validity was calculated using biserial correlations (r_b)[33] between the SODS-Likert and the SODS and the CIDI. The divergent validity was calculated using Spearman's correlation coefficient between the SODS-Likert and the SODS and the Barthel Index. A correlation of 0.50 was considered 'large', 0.30 'moderate', and 0.10 'small'.[32]

To determine the diagnostic accuracy, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated based on a 2×2 table using different cut-off scores.

The internal consistency of the SODS-Likert and the SODS was assessed using Cronbach's alpha. The inter-rater reliability was calculated using the intraclass correlation coefficient (ICC) in a one-way random-effects model. A value $\geqslant 0.70$ for both Cronbach's alpha and the ICC was considered 'positive', a rating considered acceptable for clinical practise.[34]

The clinical utility was evaluated using descriptive statistics.

To investigate whether the Likert scale of the SODS improved the diagnostic value compared with the SODS, statistical analyses were conducted on both versions.

Data were analysed using SPSS 23.0 (IBM Corp., Armonk, NY, USA).

Results

Participant characteristics

Of the 343 eligible patients, we included 116 patients, of which 53 (45.7%) with communicative impairment and 107 relatives, of which 53 (49.5%) were the relative of patients with communicative

Table 1. Patient characteristics ($n = 116$).

Characteristic	Value		
	Total study population	Patients with communicative impairment ($n = 53$)[a]	Patients without communicative impairment ($n = 63$)
Gender, *n* (%)			
Male	58 (50)	27 (50.1)	31 (49.9)
Age, mean (SD) (range)	70 (14.6) (27–90)	73.6 (12.2) (45–90)	67 (15.9) (27–88)
Type of stroke, *n* (%)			
Intracerebral haemorrhage	26 (22.4)	13 (24.5)	13 (20.6)
Infarction	88 (75.9)	39 (73.6)	49 (77.8)
Both	2 (1.7)	1 (1.9)	1 (1.6)
Side of lesion, *n* (%)			
Left	66 (56.9)	40 (75.5)	26 (41.3)
Right	39 (33.6)	12 (22.6)	27 (42.9)
Other[b]	7 (3.4)	1 (1.9)	6 (9.6)
Unknown	4 (6.0)	–	4 (6.3)
BI score, median (IQR) (range)	11 (14) (0–20)	5 (9) (0–20)	16 (11) (0–20)
FAST score, mean (SD) (range)[c]	15.8 (5) (0–20)	9.5 (4.3) (0–18)	18.5 (1.3) (15–20) $n = 28$[d]
MMSE score, mean (SD) (range)	24.1 (3.5) (8–30)	–	24.6 (2.5) (18–30)
Diagnosis of depression, *n* (%)			
CIDI-patient	8 (12.7)	–	8 (12.7)
CIDI-relative	25 (23.4)	19 (35.8)	6 (10.9)
Time between stroke onset and depression diagnosis in days, mean (SD) (range)	6.4 (2.7) (3–18)	6.9 (3.1) (3–18)	6 (2.1) (3–12)
SODS-Likert score, mean (SD) (range)	2.3 (3.2) (0–17)	3 (3) (3–18)	1.7 (3.2) (0–17)
SODS score, mean (SD) (range)	1.8 (1.8) (0–6)	2.4 (1.7) (0–5)	1.2 (1.7) (0–6)

BI: Barthel Index; FAST: Frenchay Aphasia Screening Test; IQR: interquartile range; MMSE: Mini-Mental State Examination; SODS: Signs of Depression Scale; SODS-Likert: Signs of Depression Scale-Likert.
[a]Based on FAST score or MMSE score or diagnosis by speech and language therapist.
[b]Brainstem or cerebellum.
[c]Cut-off value for communicative impairment: <17 for ages ⩾60 years, <16 for ages ⩾61 and ⩽70 years, and <15 for ages ⩾71 years, indicating communicative impairment.
[d]The FAST was not completed for 35 patients.

impairment (Figure 1). Patients were not included if they were admitted for less than four days ($n = 120$, 35%), had not provided informed consent ($n = 30$, 8.7%), or were too ill to participate ($n = 29$, 8.6%). Nine relatives (7.8%) of the 116 included patients did not give informed consent (Figure 1).

Patients with communicative impairment had moderate to severe disability (median Barthel Index score of 5 (interquartile range (IQR) 9, range: 0–20)), whereas patients who were able to communicate had mild to moderate disability (median Barthel Index score 16 (IQR 11, range: 0–20)), with a significant difference between the groups ($t = 6.45$, $P < 0.001$, 95% CI: 5.21–9.82) (Tables 1 and 2).

The prevalence of depression, measured with the CIDI, was 35.8% ($n = 19$) in patients with communicative impairment and 12.7% ($n = 8$) in patients who were able to communicate.

Validity

The correlation between the CIDI-relative and the SODS-Likert was small ($r_b = 0.18$, $P = 0.30$) and similar compared with the correlation between the

Table 2. Characteristics of relatives ($n = 107$).

Characteristic	Value
Gender, *n* (%)	
Male	40 (34.5)
Age in years, mean (SD) (range)	57.9 (16.0) (19–92)
Relation to patient, *n* (%)	
Spouse, living together	56 (48.3)
Spouse, not living together	2 (1.7)
Child	40 (34.5)
Parent	3 (2.6)
Sibling	1 (0.9)
Friend	1 (0.9)
Other	4 (3,4)
CSI score	
Relative of patient with communicative impairment, mean (SD) (range)	6.6 (2.4) (1–11)
Relative of patient without communicative impairment, mean (SD) (range)	4.7 (2.4) (0–10)

CSI: Caregiver Strain Index.

CIDI-relative and the SODS ($r_b = 0.18$, $P = 0.30$). A moderate correlation was found between the Barthel Index and the SODS-Likert ($r_s = -0.30$, $P = 0.03$), also indicating similarity compared with the correlation between the Barthel Index and the SODS ($r_s = -0.33$, $P = 0.02$).

Diagnostic accuracy

The discriminatory power of the SODS-Likert was best at a cut-off score of ⩾2 and showed a sensitivity of 0.74 (95% CI: 0.49–0.91), a specificity of 0.36 (95% CI: 0.20–0.55), a PPV of 0.40 (95% CI: 0.31–0.49), an NPV of 0.71 (95% CI: 0.50–0.85), and an area under the curve (AUC) of 0.59 (95% CI: 0.42–0.75). The diagnostic accuracy of the SODS-Likert was almost equal to that of the SODS, as shown in Table 3, indicating that the Likert scale did not improve the discriminatory power.

Reliability

The internal consistency of the SODS-Likert was $\alpha = 0.69$, showing a slight improvement compared with the SODS ($\alpha = 0.57$). The inter-rater reliability of the SODS-Likert was ICC = 0.66 (95% CI: 0.46–0.80), which was lower than the inter-rater reliability of the SODS (ICC = 0.80; 95% CI: 0.63–0.89).

Clinical utility

The clinical utility of the SODS-Likert was investigated by analysing 83 questionnaires (41%), achieving an agreement of >90% on all items, rating the clinical utility as good (Table 4).

Discussion

This study provides estimates of the diagnostic value of the SODS-Likert as well as the original dichotomous SODS in a sample of stroke patients in a hospital setting. We included patients with communicative impairment using a reference test strategy diagnosing depression with a standardized psychiatric interview based on relatives' ratings of depression and an external validator. The Likert scale format did not improve the diagnostic value of the SODS-Likert, as the findings showed equal performance of the SODS-Likert and the SODS. The performance of the SODS was satisfactory.

Our analysis was based on a small number of depressed patients; however, the prevalence of depression after stroke in patients with and without communicative impairment corresponds with other studies.[1,35] Furthermore, the range in CIDI outcomes and the scores on the SODS indicate that the sample reflected the entire range of severity of depression. Also, the sample sizes in our study are defined as good for both samples of patients with and without communicative impairment (53 and 63, respectively) and excellent for the total sample size.[19] These facts indicate that the sample size is sufficient to investigate the diagnostic value.

In this study, we found a correlation of 0.62 between patients' and relatives' ratings of the patient's mood. Due to the nature of the SODS focussing on behavioural symptoms, only partial criteria for depression according to the *Diagnostic and Statistical Manual of Mental Disorders* (DSM)[36] can be assessed. Meyer et al.[15] argue that in such cases a correlation of 0.30 defined as

Table 3. Diagnostic value of the SODS-Likert and SODS in patients with communicative impairment ($n = 52$).

	Cut-off score	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	Area under the curve (95% CI)
SODS-Likert	⩾1	0.79 (0.54–0.94)	0.24 (0.11–0.42)	0.38 (0.31–0.45)	0.67 (0.41–0.85)	0.59 (0.42–0.75)
	⩾2	0.74 (0.49–0.91)	0.36 (0.20–0.55)	0.40 (0.31–0.49)	0.71 (0.50–0.85)	
	⩾3	0.63 (0.38–0.84)	0.58 (0.39–0.75)	0.46 (0.34–0.59)	0.73 (0.58–0.84)	
SODS	⩾1	0.79 (0.54–0.94)	0.24 (0.11–0.42)	0.38 (0.23–0.54)	0.67 (0.35–0.90)	0.58 (0.42–0.75)
	⩾2	0.74 (0.49–0.91)	0.40 (0.23–0.58)	0.41 (0.25–0.59)	0.72 (0.47–0.90)	
	⩾3	0.58 (0.34–0.80)	0.64 (0.45–0.80)	0.48 (0.27–0.69)	0.72 (0.53–0.87)	

CI: confidence interval; SODS-Likert: Signs of Depression Scale-Likert; SODS: Signs of Depression Scale.

Table 4. Clinical utility of the SODS-Likert ($n = 83$).

Items	Agree *n* (%)
Includes outcomes important to the patient	79 (95.2)
Is easy to administer (low respondent burden)	78 (94.0)
Time investment to assess is acceptable	82 (98.8)
Questions are easy to understand	79 (95.2)
Instructions are clear	82 (98.8)
Is easily scored; response categories are clear	75 (90.4)
Scores are readily understandable (possible follow-up is clear)	76 (91.6)
Provides information that is clinically useful	78 (94)
Time to administer in minutes, mean (SD)	3 (3)

SODS-Likert: Signs of Depression Scale-Likert.

moderate and of 0.50 as large[32] are realistic. This is confirmed by more recent criteria qualifying a correlation of 0.50 as 'positive'.[37] These criteria support our decision to use the CIDI-relative as a reference test.

The correlations between the CIDI-relative and either the SODS-Likert or the SODS were small. We did not find meaningful differences between the diagnostic accuracy of the SODS-Likert and the SODS, which both achieved good sensitivity and NPV at a cut-off value of ⩾2. Earlier studies found moderate correlations between the original SODS and the Hospital Anxiety Depression Scale ($r_s = 0.34$)[21] and the Patient Health Questionnaire-9 ($r_b = 0.32$),[13] reporting a higher sensitivity of 0.86[21] and 0.80[13] in patients able to communicate. The use of a screening instrument instead of a diagnostic instrument could explain this discrepancy since this leads to some misclassification in the identification of depression after stroke.[35] Furthermore, these results may indicate that identifying depressive symptoms in patients with communicative impairment is more complex than in patients who can communicate adequately. The assessment of depressive symptoms using the SODS is solely based on observation of the patient's behavioural signs. As a result, symptoms as 'feeling of guilt' or 'thoughts on death or suicide' which are both DSM criteria for major depressive disorder cannot be included in the screening as these symptoms can only be assessed by patients verbal expression. A study on the symptom profile of depression after stroke supports this assumption; all symptoms of depression should be taken into account and evaluated as the symptoms profiles of depression after stroke and other patient populations do not differ.[38] To improve the assessment of depression after stroke, the use of relatives' ratings could be a valuable method to supplement the screening performed by clinicians. Considering the large correlation between patients' and relatives' ratings of patients' mood we found in our study, relatives are able to reliably rate the patient's mood status.[39]

The internal consistency of the SODS-Likert was acceptable, whereas in the SODS, it was low. Our findings are in line with the findings of Bennet et al.,[21] who found an internal consistency of $\alpha = 0.53$. Their study as well as ours had a heterogeneous sample of hospitalized patients, which could explain the low internal consistency. Another study,

which was conducted in a rehabilitation setting with a more homogeneous sample, showed an acceptable internal consistency ($\alpha = 0.71$).[13] The inter-rater reliability of the SODS was acceptable, and very similar to the reliability identified in a previous study.[13] The inter-rater reliability of the SODS-Likert was also acceptable.

The clinical utility of the SODS-Likert was found to be good, confirming earlier assumptions that the SODS is a clinically useful instrument.[10,22]

The results of our study suggest that a four-point Likert scale does not improve the diagnostic value of the SODS, thereby rejecting our hypothesis. Analyses of the SODS-Likert scores on item level revealed that the levels 'not at all' (score = 0) and 'several days' (score = 1) were the most used (85%-96%). The items 'more than half of the days' (score = 2) and 'almost every day' (score = 3) were used much less, which indicates that the nurses did not use the Likert scale optimally, and that the scores on the SODS-Likert in our study resemble a dichotomous scale. This might be caused by the fact that our study took place in the early stage after a stroke, and that nurses are cautious in labelling observed signs as being depressive symptoms. Furthermore, in our study, we aimed to reflect usual care in hospital, and therefore ratings of nurses who took care of the patient for several days as well as ratings of nurses who took care of the patient for the first day were included. In the latter case, it could be that the nurse is less prone to score the level 'more than half of the days' or 'several days', given the fact that her rating is based on her own observations combined with nursing reports of the previous days. It is known that many nurse reports lack information regarding depressive symptoms.[40]

With an optimal cut-off score of $\geqslant 2$ for both instruments, the SODS, in which scores range from 0 to 6, was more appropriate for screening depressive symptoms in patients unable to communicate than the SODS-Likert, which has a score range of 0–18. Therefore, the SODS-Likert cannot be recommended. However, the good sensitivity and NPV found in this study show that the SODS can be used as an initial screening instrument to identify patients who require further thorough assessment, thereby preventing the unnecessary use of expensive and more burdensome psychiatric diagnostic resources.[6] Also, the repeated assessment of depression is recommended because of the fluctuating nature of mood states after stroke, which may limit the effect of the poor specificity.

In summary, the psychometrics of the SODS are sufficient for the initial mood assessment of stroke patients with communicative impairment and the Likert scale did not improve its diagnostic value. Our findings suggest that relatives are able to screen patients for depression after stroke which may be important to further optimize stroke care.

Study limitations

Our study has some limitations that need to be emphasized.

Although a standardized psychiatric interview to diagnose depression is considered a reliable and valid method,[2] it has some limitations when used in a hospital setting. The criterion that symptoms have to be present for a minimum of two weeks to provide a diagnosis of depression could not be satisfied in our study because the mean length of hospital stay after stroke is <14 days in the Netherlands.[35] We accepted this limitation because we believed that a standardized psychiatric interview provided the best possible diagnosis.[2] Moreover, all patients fulfilled all other criteria for major depressive disorder. Nevertheless, when interpreting the results of our study, this limitation has to be considered.

Another possible limitation is the use of relatives' ratings to provide a diagnosis of depression. Based on the large correlation we found between the CIDI administered on patients without communicative impairment and their relatives, and supported by other studies who found similar results on agreement between patients' and relatives' ratings,[31,41] we made the assumption that these findings can be generalized to patients with communicative impairment. Due to the fact that patients with communicative impairment are not able to respond reliably to a psychiatric interview, evidence on patient's and relative's agreement in these patients using the CIDI, or any other language-based

diagnostic instrument, is lacking. However, the assumption on generalizability seems reasonable given the evidence and therewith enabled the inclusion of stroke patients with communicative impairment, which strengthens the methodological quality of our study.

Finally, the psychometric properties of the SODS-Likert were only investigated in a hospital setting. Our sample represented stroke patients within the total range of stroke diagnoses and handicaps, and therefore the results can be generalized to the total hospital population. However, the psychometric properties of the SODS-Likert could vary when studied in other health care facilities where patients with stroke stay for longer periods of time. One can assume that if clinicians become more acquainted with the patients in a rehabilitation or geriatric setting than in the hospital, due to the longer stay in these care facilities, this familiarity can result in a better identification of depressive symptoms and thereby a higher diagnostic value of the SODS-Likert.

Clinical Messages

- Clinicians can use the Signs of Depression Scale for the initial assessment of depressive symptoms in patients with communicative impairments.
- Relatives' ratings in the assessment of depression after stroke in patients with communicative impairments could be a valuable method to supplement the screening performed by clinicians.

Acknowledgements

The authors thank the patients and staff for their participation and support. Also, we thank Dr. C.L.J.J. Kruitwagen for guidance with statistical analysis strategy.

Author Contributions

M.J.v.D. contributed to designing the study, recruiting patients and their relatives, collecting the data, deciding on the analytic strategy, performing analysis, and writing the paper. J.M.d.M.G. contributed to designing the study, deciding on the analytic strategy, and revising the paper critically for important intellectual content. T.B.H. and M.J.S. contributed deciding on the analytic strategy and revising the paper critically for important intellectual content. M.J.S. is the guarantor.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Foundation of Innovation Alliance, Regional Attention and Action for Knowledge circulation (SIA RAAK) and by an internal grant of the University of Applied Sciences Utrecht.

References

1. Hackett ML and Pickles K. Part I: frequency of depression after stroke: an updated systematic review and meta-analysis of observational studies. *Int J Stroke* 2014; 9: 1017–1025.
2. Robinson RG and Jorge RE. Post-stroke depression: a review. *Am J Psychiatry* 2015; 173: 221–231.
3. Kouwenhoven SE, Kirkevold M, Engedal K, et al. Depression in acute stroke: prevalence, dominant symptoms and associated factors. A systematic literature review. *Disabil Rehabil* 2011; 33: 539–556.
4. Boddice G, Brauer S, Gustafsson L, et al. *Clinical guidelines for stroke management*. Melbourne, VIC, Australia: National Stroke Foundation, 2010.
5. Miller EL, Murray L, Richards L, et al. Comprehensive overview of nursing and interdisciplinary rehabilitation care of the stroke patient. A scientific statement from the American Heart Association. *Stroke* 2010; 41: 2402–2448.
6. Salter K, Bhogal SK, Foley N, et al. The assessment of poststroke depression. *Top Stroke Rehabil* 2007; 14: 1–24.
7. Van Dijk MJ, de Man-van Ginkel JM, Hafsteinsdottir TB, et al. Identifying depression post-stroke in patients with aphasia: a systematic review of the reliability, validity and feasibility of available instruments. *Clin Rehabil* 2016; 30: 795–810.
8. Benaim C, Cailly B, Perennou D, et al. Validation of the aphasic depression rating scale. *Stroke* 2004; 35: 1692–1696.
9. Guy W. The clinical global impression scale. In: Guy W (ed.) *ECDEU assessment manual for psychopharmacology-revised*. Rockville, MD: US Department of Health, Education & Welfare, ADAMHA, NIMH Psychopharmacology Research Branch, 1976, pp. 218–222.
10. Hammond MF, O'Keeffe ST and Barer DH. Development and validation of a brief observer-rated screening scale for

depression in elderly medical patients. *Age Ageing* 2000; 29: 511–515.
11. Sutcliffe L and Lincoln N. The assessment of depression in aphasic stroke patients: the development of the Stroke Aphasic Depression Questionnaire. *Clin Rehabil* 1998; 12: 506–513.
12. Burton L and Tyson S. Screening for mood disorders after stroke: a systematic review of psychometric properties and clinical utility. *Psychol Med* 2015; 45: 29–49.
13. Van Dijk MJ, de Man-van Ginkel JM, Hafsteinsdóttir TB, et al. Psychometric Properties of the Dutch Version of the Signs of Depression Scale. *Clin Nurs Res*. Epub ahead of print 5 July 2016. DOI: 10.1177/1054773816657798.
14. Streiner DL and Norman GR. Scaling responses. In: Streiner DL and Norman GR (eds) *Health measurement scales: a practical guide to their development and use*. 4th ed. New York: Oxford University Press, 2008, pp. 37–75.
15. Meyer GJ, Finn SE, Eyde LD, et al. Psychological testing and psychological assessment: a review of evidence and issues. *Am Psychol* 2001; 56: 128–165.
16. Robins LN, Wing J, Wittchen HU, et al. The Composite International Diagnostic Interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiat* 1988; 45: 1069–1077.
17. Mahoney FI. The Barthel Index. *Maryland State Med J* 1965; 14: 61–65.
18. Kasner SE. Clinical interpretation and use of stroke scales. *Lancet Neurol* 2006; 5: 603–612.
19. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010; 10: 22.
20. Watkins C, Leathley M, Daniels L, et al. The signs of depression scale in stroke: how useful are nurses observations. *Clin Rehabil* 2001; 15: 447–457.
21. Bennett H, Thomas S, Austen R, et al. Validation of screening measures for assessing mood in stroke patients. *Br J Clin Psychol* 2006; 45: 367–376.
22. Lightbody CE, Auton M, Baldwin R, et al. The use of nurses' and carers' observations in the identification of poststroke depression. *J Adv Nurs* 2007; 60: 595–604.
23. Maneesriwongul W and Dixon JK. Instrument translation process: a methods review. *J Adv Nurs* 2004; 48: 175–186.
24. Kroenke K and Spitzer RL. The PHQ-9: a new depression diagnostic and severity measure. *Psychiatr Ann* 2002; 32: 509–515.
25. Williams LS, Brizendine EJ, Plue L, et al. Performance of the PHQ-9 as a screening tool for depression after stroke. *Stroke* 2005; 36: 635–638.
26. De Man-van Ginkel JM, Hafsteinsdottir T, Lindeman E, et al. An efficient way to detect poststroke depression by subsequent administration of a 9-item and a 2-item Patient Health Questionnaire. *Stroke* 2012; 43: 854–856.
27. De Man-van Ginkel JM, Gooskens F, Schepers VP, et al. Screening for poststroke depression using the patient health questionnaire. *Nurs Res* 2012; 61: 333–341.
28. Enderby PM, Wood VA, Wade DT, et al. The Frenchay Aphasia Screening Test: a short, simple test for aphasia appropriate for non-specialists. *Int Rehabil Med* 1986; 8: 166–170.
29. Folstein MF, Folstein SE and McHugh PR. 'Mini-mental state': a practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 189–198.
30. Robinson BC. Validation of a caregiver strain index. *J Gerontol* 1983; 38: 344–348.
31. Williams LS, Bakas T, Brizendine E, et al. How valid are family proxy assessments of stroke patients' health-related quality of life? *Stroke* 2006; 37: 2081–2085.
32. Cohen J. A power primer. *Psychol Bull* 1992; 112: 155–159.
33. Field A. Correlation. In: Field A (ed.) *Discovering statistics using SPSS*. 3rd ed. London: SAGE, 2009, pp.166–196.
34. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007; 60: 34–42.
35. De Man-van Ginkel JM, Hafsteinsdottir TB, Lindeman E, et al. In-hospital risk prediction for post-stroke depression: development and validation of the Post-stroke Depression Prediction Scale. *Stroke* 2013; 44: 2441–2445.
36. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*, text revision. 4th ed. Washington, DC: American Psychiatric Association, 2004.
37. Schellingerhout JM, Verhagen AP, Heymans MW, et al. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. *Qual Life Res* 2012; 21: 659–670.
38. De Man-van Ginkel JM, Hafsteinsdóttir TB, Lindeman E, et al. Clinical manifestation of depression after stroke: is it different from depression in other patient populations? *PLoS ONE* 2015; 10: e0144450.
39. Sneeuw K, Aaronson N, De Haan R, et al. Assessing quality of life after stroke: the value and limitations of proxy ratings. *Stroke* 1997; 28: 1541–1549.
40. Edwards DF, Hahn MG, Baum CM, et al. Screening patients with stroke for rehabilitation needs: validation of the post-stroke rehabilitation guidelines. *Neurorehabil Neural Repair* 2006; 20: 42–48.
41. Pickard AS, Johnson JA, Feeny DH, et al. Agreement between patient and proxy assessments of health-related quality of life after stroke using the EQ-5D and Health Utilities Index. *Stroke* 2004; 35: 607–612.