

Research paper

Predictive accuracy of the Post-Stroke Depression Prediction Scale: A prospective binational observational study[☆]



Julian Hirt^{a,b,1}, Lianne C.J. van Meijeren^{c,1}, Susanne Saal^a, Thóra B. Hafsteinsdóttir^{c,d},
Jeannette Hofmeijer^e, Andrea Kraft^f, Gabriele Meyer^a, Janneke M. de Man-van Ginkel^{c,d,*}

^a Institute for Health and Nursing Sciences, Medical Faculty, Martin Luther University Halle-Wittenberg, Magdeburger Straße 8, 06112 Halle (Saale), Germany

^b Institute of Applied Nursing Sciences, Department of Health, FHS St.Gallen, Rosenbergstrasse 59, 9000 St.Gallen, Switzerland

^c Nursing Sciences, Program in Clinical Health Sciences, University Medical Center Utrecht, University Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands

^d Julius Center for Health Science and Primary Care, Department of Nursing Science, University Medical Center Utrecht, University Utrecht, Heidelberglaan 100, 3584 CX Utrecht, the Netherlands

^e Department of Neurology, Rijnstate, Wagnerlaan 55, 6815 CE Arnhem, the Netherlands

^f Department of Neurology of the Hospital Martha Maria Halle-Dölau, Röntgenstraße 1, 06120 Halle (Saale), Germany

ARTICLE INFO

Keywords:

Stroke

Depression

Clinical prediction rule

Predictive value of tests

ABSTRACT

Background: Depression after a stroke is a common complication that negatively influences stroke rehabilitation. Early identification, followed by adequate treatment of depression, improves recovery from stroke. To support early identification, the Post-stroke Depression Prediction Scale (DePreS) was developed to predict in the first week after stroke, the risk of depression in the second month. In this study we investigate the predictive accuracy of the DePreS in stroke patients.

Methods: In this prospective multicenter observational study, hospitalized stroke patients were included from three stroke units in the Netherlands and Germany using consecutive sampling. In the first week after stroke, the predicted risk for depression was estimated with the DePreS. Two months after stroke, major depressive disorder was determined with the Composite International Diagnostic Interview.

Results: Of the 93 included patients, 17 (18.3%) showed symptoms of major depressive disorder. With a cut-off value of ≥ 0 , DePreS performed optimally with a sensitivity of 0.65 (95% CI 0.42–0.87), specificity of 0.74 (95% CI 0.64–0.84), positive predictive value of 0.35 (95% CI 0.19–0.52), and negative predictive value of 0.90 (95% CI 0.80–1.00). The AUC was 0.71 (95% CI 0.56–0.86).

Limitations: The generalizability of the study findings is limited to patients able to communicate adequately. **Conclusions:** This study demonstrates that the DePreS is an adequate instrument for early and reliable identification of stroke patients who are not at risk of MDD in the second months after stroke. This limits the need for structural diagnostic follow-up to patients with a high risk.

1. Background

Depression is a frequent complication after stroke. In the first two months after stroke, an average of 36% of patients develop symptoms of depression and 14% are diagnosed with major depressive disorder (MDD) (Hackett and Pickles, 2014). The symptoms of depression are a depressed mood, loss of interest or pleasure, significant change in appetite or weight, feelings of guilt or low self-worth, disturbed sleep,

change in activity level, feelings of tiredness, poor concentration, and suicidal ideation (American Psychiatric Association, 2000). According to the Diagnostic and Statistical Manual of Mental Disorders version V (DSM-V TR™), the diagnosis of major depressive disorder is defined as the consistent presence of five or more of the nine symptoms of depression over a two-week period, which represents a change from the previous criteria of experiencing at least one symptom, such as a depressed mood or loss of interest or pleasure (American Psychiatric

[☆]Clinical Trial Registration

URL: www.drks.de; Unique identifier: DRKS00011546; Date of registration: 30 January 2017.

* Corresponding author: Janneke M. de Man-van Ginkel, University Medical Center Utrecht, Julius Center for Health Sciences and Primary Care, Str.6.131, PO Box 85500, 3508 GA Utrecht, the Netherlands.

E-mail address: J.M.deMan@utrecht.nl (J.M. de Man-van Ginkel).

¹ These authors contributed equally.

<https://doi.org/10.1016/j.jad.2020.01.019>

Received 18 February 2019; Received in revised form 16 November 2019; Accepted 5 January 2020

Available online 08 January 2020

0165-0327/ © 2020 Elsevier B.V. All rights reserved.

Association, 2000).

Depression after stroke is associated with an increased disability, whereas recovery from depression after stroke within one year improves functional outcomes and quality of life (Shi et al., 2016), suggesting that depression hampers functional recovery after stroke (Robinson and Jorge, 2016). Early identification of patients at risk for depression may facilitate adequate treatment and follow up.

Over the last few decades, there have been major developments in the treatment of stroke patients combined with the use of care pathways, shortening the patient's length of stay in hospital to a mean of ≤ 7 days (Van Raak et al., 2010). This reduction in length of stay hinders optimal screening for depression during hospital admission since the two week time frame of the DSM criteria can no longer be used to diagnose depression. To enable early identification given the limited length of hospital stay, a prediction instrument, the Post-stroke Depression Prediction Scale (DePreS), was developed to support clinicians in predicting in the first week the risk of depression in the second month after stroke using routine care data (De Man-van Ginkel et al., 2013). The following clusters of risk factors were used in the development of the DePreS: socio-demographic and stroke-related factors such as age, gender, marital status, premorbid living situation, educational level, work, the type of stroke, the stroke location, and the consequences of stroke. The following medical history data were collected: 1) vascular risk factors such as diabetes mellitus, hypertension, dyslipidemia, impaired renal function, smoking and alcohol consumption; 2) vascular diseases such as prior stroke, transient ischaemic attack, angina pectoris, myocardial infarction, atrial fibrillation and peripheral arterial disease; and 3) other diseases such as other brain disorders, depression or other psychiatric disorders; functional status post stroke measured using the Barthel Index (BI) (Mahoney and Barthel, 1965) and the modified Rankin Scale (Van Swieten et al., 1988). The patients perceived lack or excess of social support was measured with the Social Support List (SSL-6) (Bridges et al., 2002). The final model, resulting in the DePreS was selected from this set of variables and resulted in following items: history of depression, hypertension, angina pectoris, and help with dressing in the first week after stroke (Table 1).

To strengthen the future applicability of the DePreS in clinical practice, only risk factors that are easily available to clinicians and that is generally collected in the hospital were used as predictors. Previously, the performance of the DePreS was evaluated in the same sample used for development, resulting in an area under the receiver operating characteristic curve (AUC) of 0.78 (0.72–0.85) and a negative predictive value and specificity of 0.94 (0.91–0.97) and 0.75 (0.70–0.80), respectively, with a cut-off value of ≥ 2 (De Man-van Ginkel et al., 2013). However, since validating an instrument in the development sample may result in overoptimistic predictive performance, evaluation of the predictive accuracy in a new sample of stroke patients is needed (Altman et al., 2009). The aim of this study is therefore to investigate the predictive accuracy of the DePreS administered in the first week after stroke for estimating the risk of depression in the second month after stroke in hospitalized stroke

patients.

2. Methods

2.1. Study design and participant selection

A multicenter clinimetric study with a longitudinal, prospective and observational design was conducted in three hospitals not involved in the development study of the DePreS, one in the Netherlands and two in Germany. Ethical approval was obtained by the Medical Ethical Committee of the University Medical centre Utrecht (16–809/C) and the Ethical Committee of the Medical Faculty of the Martin Luther University Halle-Wittenberg (2016–162). From February 2017 to June 2017, 204 patients admitted to the hospital with a diagnosis of ischemic stroke or intracerebral hemorrhage were approached consecutively for participation. If the patient met the inclusion criteria, the researcher informed the patient about the study, checked whether the patients was able to participate, obtained written informed consent and collected the baseline data.

Of the 204 patients, 111 were excluded for the following reasons: too ill to participate ($n = 35$, 32%); a communicative disorder based on the Frenchay Aphasia Screening Test in the Netherlands or a shortened version of the Aphasie-Schnelltest in Germany, with scores of ≥ 17 points in patients < 60 years of age, ≥ 16 in patients ≥ 60 and ≤ 70 years of age and ≥ 15 in patients > 70 years of age (Enderby et al., 1987; Kroker, 2006); cognitive disorder based on scores of ≥ 18 on the Mini-Mental State Examination ($n = 24$, 25.8%) (De Man-van Ginkel, 2013; Folstein et al., 1975); diagnosed depression; antidepressant prescription or other psychiatric comorbidity at stroke onset based on anamnesis or a medical diagnosis ($n = 10$, 10.8%); or no Dutch or German language skills ($n = 3$, 3.2%). Additional reasons for exclusion were: declined to participate ($n = 23$, 24.7%) and other reasons ($n = 16$, 17.2%).

This resulted in a sample size of 93 patients (48 in the Netherlands and 45 in Germany) (Fig. 1), which met the calculated number of at least 90 patients who had to be included to be able to reach adequate sensitivity and specificity (Buderer, 1996). The sample size was calculated according to the method of Burdener with a sensitivity of 0.73, a specificity of 0.75, and confidence intervals around the sensitivity and specificity of ± 0.10 , in combination with a prevalence of 0.23. These values were taken from the development study of the DePreS (De Man-van Ginkel et al., 2013). The estimated prevalence of 0.23 is the average of the prevalence presented in the literature (Hackett and Pickles, 2014) and the prevalence found in the development study (De Man-van Ginkel et al., 2013). Furthermore, a dropout rate of 10% was assumed. This resulted in a sample size of 90 patients.

2.2. Data collection

The risk of depression in the first week after stroke was measured with the DePreS and the German version of the index test, the DePreS-G. The DePreS consists of the following items: a medical history of depression or other psychiatric disorders, history of hypertension, history of angina pectoris, and the Barthel Index item 'Dressing'. The DePreS score ranges from -19 to 23 . A higher score indicates a higher risk of depression after stroke (De Man-van Ginkel et al., 2013). For the German part of the study, the DePreS-G was translated from Dutch to German following the ISPOR criteria (Wild et al., 2005) with two independent forward and backward translations and a cognitive debriefing with 18 potential users of the instrument.

As the reference test, the Composite International Diagnostic Interview (CIDI) was applied, which is a structured psychiatric interview used to establish mental disorders, such as MDD, according to the DSM-IV and the ICD-10. The CIDI has been proven to be a reliable and valid instrument, with an inter-observer agreement (Cohen's kappa) of 0.71 and sensitivity and specificity of 80% (Wittchen, 1994). In the

Table 1
Score table Post-stroke Depression Prediction Scale (DePreS).

Does the patient have a medical history of depression or other psychiatric disorders?	No	☐ 0
	Yes	☐ 13
Does the patient have a medical history of hypertension?	No	☐ 0
	Yes	☐ -5
Does the patient have a medical history of angina pectoris?	No	☐ 0
	Yes	☐ 7
To what extent does the patient need help with dressing in the first week after stroke (Barthel Index item dressing)	Completely independent	☐ 0
	Needs help but can do about half unaided	☐ -9
	Dependent	☐ 3
Sum score	Add the item scores to obtain the sum score	—

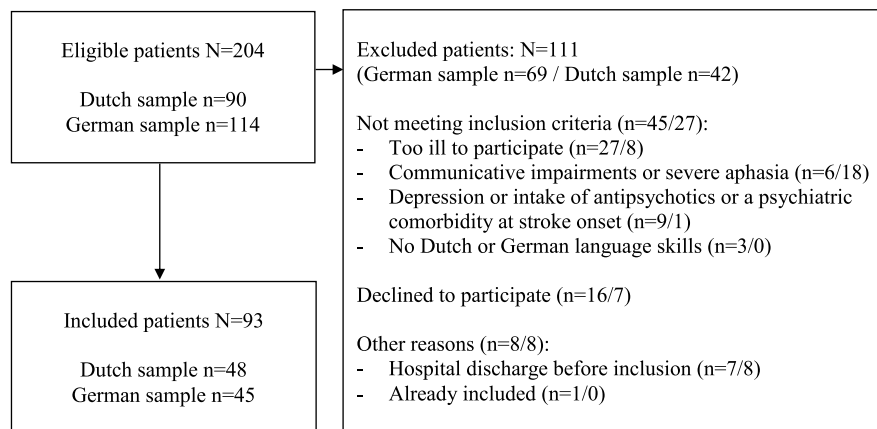


Fig. 1. Recruitment process.

Table 2
Baseline characteristics per subsample.

		Netherlands <i>n</i> = 48	Germany <i>n</i> = 45	p-value
Sex, <i>n</i> (%)	Male	30 (62.5)	27 (60)	0.805 ^a
Age, mean (SD; min-max)	74 (11.8; 44–93)	66.7 (13; 44–91)	0.006 ^b	
Marital status, <i>n</i> (%)	Single	9 (18.8)	4 (8.9)	0.390 ^a
	Married/living together	24 (50)	25 (55.6)	
	Widowed/divorced	15 (31.3)	16 (35.6)	
Home situation, <i>n</i> (%)	Independent-living alone	21 (43.8)	9 (20)	0.060 ^a
	Independent-living together with someone	25 (52.1)	34 (75.6)	
	Nursing home	2 (4.2)	1 (2.2)	
	Assisted Living	0	1 (2.2)	
Discharge direction, <i>n</i> (%)	Home	29 (60.4)	16 (35.6)	0.029 ^a
	Rehabilitation center	19 (39.6)	27 (60)	
	Other	0	2 (4.4)	
FAST/AST, mean (SD; min-max)	19 (1; 15–20)	19 (1; 16–20)	0.004 ^b	
MMSE, mean (SD; min-max)	26 (3; 18–30)	27 (3; 18–30)	0.114 ^b	
Type of stroke, <i>n</i> (%)	Intracerebral hemorrhage	0	2 (4.4)	0.140 ^a
	Ischemic stroke	48 (100)	43 (95.6)	
Localization, <i>n</i> (%)	Left	14 (29.2)	18 (40)	0.001 ^b
	Right	13 (27.1)	12 (26.7)	
	Cerebellum	17 (35.4)	2 (4.4)	
	Other	4 (8.3)	13 (24.5)	
Symptoms, <i>n</i> (%)	Motor	36 (75)	19 (42.2)	0.001 ^a
	Sensibility disturbances	12 (25)	16 (35.6)	0.267 ^a
	Speech	19 (39.6)	14 (31.1)	0.393 ^a
	Language	3 (6.3)	4 (8.9)	0.630 ^a
	Visual disturbances	13 (27.1)	12 (26.7)	0.964 ^a
Vascular risk factors, <i>n</i> (%)	Diabetes mellitus	10 (20.8)	18 (40)	0.044 ^a
	Hypertension	31 (64.6)	33 (73.3)	0.363 ^a
	Dyslipidemia	19 (39.6)	17 (37.8)	0.858 ^a
	Impaired renal function	4 (8.3)	9 (20)	0.105 ^a
	Smoking	12 (25)	13 (28.9)	0.672 ^a
	Alcohol consumption	29 (60.4)	26 (57.8)	0.796 ^a
Vascular diseases, <i>n</i> (%)	Prior stroke	13 (27.1)	10 (22.2)	0.587 ^a
	Transient ischemic attack	6 (12.5)	0	0.014 ^a
	Angina pectoris	8 (16.7)	8 (17.8)	0.887 ^a
	Myocardial infarction	8 (16.7)	6 (13.3)	0.653 ^a
	Atrial fibrillation	7 (14.6)	8 (17.8)	0.676 ^a
	Peripheral arterial disease	7 (14.6)	3 (6.7)	0.279 ^a
Other diseases, <i>n</i> (%)	Other brain disorders	3 (6.3)	5 (11.1)	0.403 ^a
	Depression	2 (4.2)	1 (2.2)	0.596 ^a
	Other psychiatric disorders	2 (4.2)	4 (8.9)	0.354 ^a
Barthel Index T0, Mean (SD; min-max)	14 (5; 3–20)	16 (5; 2–20)	0.005 ^b	
Barthel Index T1, Mean (SD; min-max)	18 (4; 3–20)	18 (3; 7–20)	0.879 ^b	

Abbreviations: AST = Aphasia-Schnelltest; FAST = Frenchay Aphasia Screening Test; MMSE = Mini Mental State Examination; *n* = number; SD = Standard Deviation; T0 = within the first week after stroke; T1 = second month after stroke.

^a = Result based on chi-squared test

^b = Result based on t-test.

current study, the CIDI-auto 2.1 version was used, and both interviewers (JH, LM) were trained in administering the CIDI-auto (Ter Smitten et al., 2008).

Disability was measured with the Barthel Index (BI). The BI is a 10-

item scale, including activities of daily living (ADL) and mobility. The score ranges from 0 to 20. A higher score indicates greater independence (Collin et al., 2009). Inter-rater reliability showed a Cohen's kappa of 0.88 and diagnostic accuracy of 0.82 (AUC) (Banks and

Marotta, 2007; Balu, 2009). Furthermore, the baseline characteristics were collected regarding age, sex, type of stroke, lesion location, consequences of stroke and discharge directions.

To maintain the blinding of the researchers towards the index test in the participating Dutch hospital, ward nurses collected the baseline data included in the DePreS and the researcher (LM) collected data for the other baseline characteristics. In the participating German hospitals, the research assistant collected all baseline and DePreS data to enable blinding of the researcher (JH). Six to eight weeks after stroke onset, the researcher (JH, LM) conducted the CIDI by telephone. In case of a positive result on the first few questions of the CIDI, which gives an indication of potential MDD, an interview was continued at a later time point at the home or residential health care facility of the patient. To prevent observer bias, the DePreS and CIDI were measured by different persons, independently. The researchers who conducted the CIDI were not aware of the results of the DePreS.

2.3. Data analysis

The baseline characteristics and prevalence of a major depressive disorder were analyzed separately in the Dutch and German subsamples to identify potential differences. Chi-squared test was used for categorical data and *t*-test for metric data. For each score of the DePreS, all participants were classified as true positive, true negative, false positive and false negative, based on a 2×2 table. The sensitivity, specificity, positive predictive value, and negative predictive value, including corresponding 95% confidence intervals, were calculated for each cut-off score. These analyses were conducted separately in both subsamples to establish whether potential differences in the baseline characteristics influenced the predictive accuracy of the DePreS. Finally, the AUC and Youden's index were calculated (Hughes, 2015). Since the subsamples were not different, only the results of the total sample are presented. All analyses were conducted using IBM SPSS Statistics Version 23 (IBM, Armonk, NY, USA). The study data are available from the corresponding author upon request.

3. Results

The mean age of the 93 included patients was 70.4 (SD = 12.8), and 57 (61.3%) were male. Most patients (97.8%; $n = 91$) suffered from an ischemic stroke. The baseline characteristics are presented in Table 2. The baseline characteristics of the subsamples were predominantly comparable. The Dutch and German subsamples differed in terms of vascular risk factors and stroke severity. Patients in the German sample were younger, on average, and suffered from diabetes mellitus or hypertension more often. The Dutch sample had stroke-related motor symptoms more often and, consequently, a lower mean score on the Barthel Index. By contrast, in the German sample, the proportion of patients who were discharged to inpatient rehabilitation was higher.

The prevalence of a major depressive disorder differed between samples: 20.8% ($n = 10$) in the Netherlands sample versus 15.6% ($n = 7$) in the German samples, resulting in an overall prevalence of major depressive disorder of 18.3% ($n = 17$).

The observed sum scores of the DePreS ranged from -14 to 18 , and 62.4% of patients ($n = 58$) had a sum score ≥ 0 . The predictive accuracy of the DePreS was found to be best for a cut-off score of 0, with a specificity of 0.74 (0.64–0.84), sensitivity of 0.64 (0.42–0.87), negative predictive value of 0.90 (0.80–1.00), and positive predicted value of 0.35 (0.19–0.52). The discriminative power of the DePreS was good, with an area under the ROC-curve of 0.71 (95% CI: 0.56–0.86) (Fig. 2). An overview of these findings for the different cut-off values is presented in Table 3.

4. Discussion

The aim of this study was to investigate the predictive accuracy of

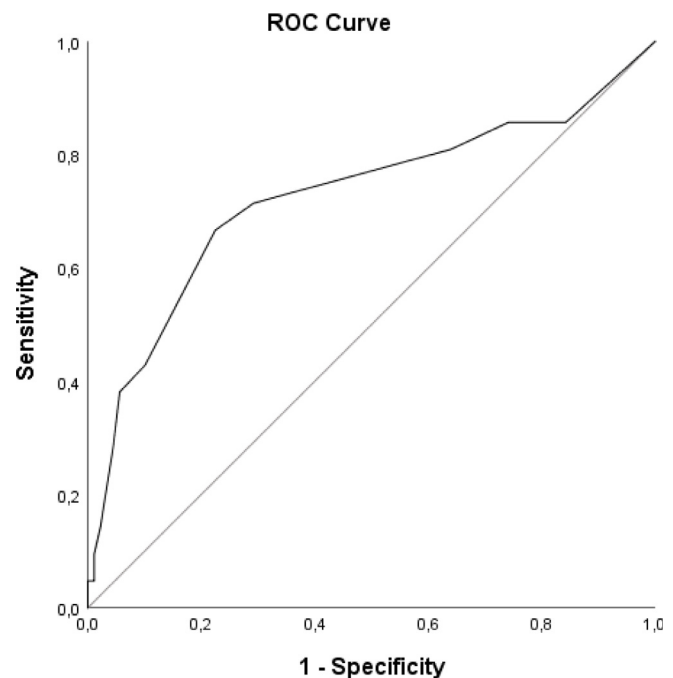


Fig. 2. Receiver operating characteristic curve.

the DePreS administered in the first week after ischemic or hemorrhagic stroke for estimating the risk of MDD in the second month after stroke. Overall, the DePreS showed an adequate predictive accuracy and discriminative power, with a high negative predictive value and specificity and a relatively high AUC. These results indicate that the DePreS can be used to reliably identify in the first week after stroke patients in whom the risk of depression in the second month after stroke is low.

The results in this study endorse that the DePreS can serve as the basis for a simple decision tree to guide a more selective screening process. Since the motive for the development of the DePreS was to enable clinicians to estimate the risk for depression for an individual patient in the first week after stroke considering the limited window of in-hospital opportunity (De Man-van Ginkel et al., 2013), all predictors included in the DePreS are available to clinicians in the first week after stroke from routine care data. As a result, the DePreS can be used without need for structural additional assessment in those patients for whom the DePreS indicate a low risk for MDD after stroke. This is an important advantage over other early identification strategies that require active involvement of patients, apart from the fact that these strategies are sub-optimal because of the short length of stay in hospital (Van Raak, 2010).

In the current study the main focus was to investigate the predictive accuracy of the DePreS, for which the optimal strategy is to conduct a study in a new sample of patients admitted to other institutions than those in which the patients of the development study were recruited (Altman et al., 2009). Therefore, we included patients admitted to a hospital in the Netherlands who did not participate in the previous study in which the DePreS was developed and admitted to two hospitals in Germany. The inclusion of these patients increased the power of the findings for the entire population since involving new institutes from different countries optimally reflects the normal variability of patients in the population. This bi-national nature, however, can easily introduce differences in the baseline characteristics of the participating patients due to potential differences in the stroke population between the Netherlands and Germany. When comparing the baseline characteristics between the Dutch and German samples we made some interesting observations. The vascular risk factors, such as hypertension and diabetes mellitus, were more prevalent and the mean age was lower in the German sample than that in the Dutch sample. By contrast, the

Table 3
Performance of the DePreS.

Cut-off	n	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)	True-positive n (%)	False-positive n (%)	False-negative n (%)	True-negative n (%)	Youden's index
–5	65	0.82 (0.64–1.00)	0.33 (0.22–0.43)	0.22 (0.12–0.32)	0.89 (0.82–0.97)	14 (21.5)	51 (78.5)	3 (10.7)	25 (89.3)	0.15
–2	38	0.71 (0.49–0.92)	0.66 (0.55–0.76)	0.32 (0.17–0.46)	0.91 (0.82–1)	12 (31.6)	26 (68.4)	5 (9.1)	50 (90.9)	0.37
0	31	0.65 (0.42–0.87)	0.74 (0.64–0.84)	0.35 (0.19–0.52)	0.90 (0.80–1)	11 (35.5)	20 (64.5)	6 (9.7)	56 (90.3)	0.39
2	16	0.41 (0.18–0.65)	0.88 (0.81–0.95)	0.44 (0.19–0.68)	0.87 (0.71–1.03)	7 (43.7)	9 (56.3)	10 (13.0)	67 (87.0)	0.29
3	11	0.35 (0.13–0.58)	0.93 (0.88–0.99)	0.55 (0.25–0.84)	0.87 (0.66–1.07)	6 (54.5)	5 (45.5)	11 (13.4)	71 (86.6)	0.28

Abbreviations: CI = Confidence interval; n = Number.

Notes: Youden's Index = Sensitivity + Specificity–1.

prevalence of motor symptoms was higher and the mean BI sum score was lower in the Dutch sample than those in the German sample. Interestingly however, the proportion of patients discharged to inpatient rehabilitation was considerably lower in the Dutch sample than that in the German sample. These differences are found in the literature as well and seem to reflect the differences in socioeconomic status and healthcare systems that still exist between the two populations represented by the study samples (Grau et al., 2001; Haerting et al., 2012; Arboix, 2015).

Another remarkable finding is that in the Dutch sample no patients, and in the German sample only two patients, with intracerebral hemorrhage were recruited. Since patients with intracerebral hemorrhage often demonstrate more severe stroke consequences in the first few days after stroke, it is reasonable that these patients were excluded because they were too ill to participate (Ojaghihaghghi et al., 2017; Chiu et al., 2010). This finding is comparable to that of an earlier study in the same population in Halle (Saale) that also included a low proportion of patients with intracerebral hemorrhage (Saal et al., 2015). In addition, the absence of patients with an intracerebral hemorrhage in the Dutch sample can be explained by the fact that the Dutch sample was recruited in a non-academic hospital and patients admitted with an intracerebral hemorrhage are often referred to an academic hospital for specialized treatment directly after assessment in the emergency department.

The prevalence of MDD after stroke in this study was higher than that in the development study (De Man-van Ginkel et al., 2013). However, the prevalence was comparable to the internationally known prevalence of MDD after stroke, indicating a 13% (95% CI 8–19%) pooled estimate of the prevalence of MDD after stroke according to the DSM criteria within six weeks after a stroke (Hackett and Pickles, 2014).

The AUC of the DePreS of 0.71 (95% CI 0.56–0.86) was slightly lower than that of the development study (AUC 0.78, 95% CI 0.72–0.85) (De Man-van Ginkel et al., 2013). Although it was expected that the results of a diagnostic validation study in a new sample of patients would show lower discriminative findings (Bleeker et al., 2003), the difference was fairly small (>0.1), indicating that the AUC of the DePreS, as it was between 0.7 and 0.8, is acceptable and makes the DePreS a stable instrument (Mandrekar, 2010). This finding also confirms that the DePreS discriminates well between stroke patients not at risk for depression from those who are at risk for depression after a stroke (Steyerberg, 2009).

In the current study, the predictive performance was found to be best at a cut-off score of ≥ 0 , as was reflected by the highest negative predictive value, acceptable specificity, and highest specificity and Youden's index. In the development study, the best cut-off score was found and the Youden's index was highest at a cut-off score of ≥ 2 (De Man-van Ginkel et al., 2013).

However, unpublished findings showed a higher negative predictive value at a cut-off score of ≥ 0 , which also emphasizes the stability of the DePreS (Mandrekar, 2010).

The consequence of the focus on specificity and negative predictive value, however, is that the performance of the DePreS resulted in a low

sensitivity and positive predictive value. This can be explained by a slightly lower prevalence than assumed based on the literature and the findings of the development study when calculating the needed sample size. Although the calculated sample size was reached, we found a lower prevalence. Since the prevalence of disease is directly related to the prevalence of the disease this lower prevalence might explain the predictive values found for the DePreS (Parikh et al., 2008). Furthermore, the resulted sensitivity was lower in comparison with the development study of the DePreS throughout the calculated cut-off scores which could be explained by the assumable lower predictive accuracy parameters in validation studies (Bleeker et al., 2003; De Man-van Ginkel et al., 2013). The resulted sensitivity and positive predictive value indicate a relatively large proportion of patients that are incorrectly identified as at risk for depression. However, since additional screening and diagnosing of MDD are indicated in these patients, they will be spared unnecessary treatment. This is preferable to being incorrectly identified as at no risk for depression, as this results in these patients, who need adequate monitoring and, if necessary, treatment, in fact will not get any follow up. This is expected to be more burdensome than the additional screening, not only because of the burden of depression but also because of its negative impact on rehabilitation outcomes (Robinson and Jorge, 2016). When determining the optimal cutoff value, we focused on an optimal balance between the NPV and the PPV as well as the sensitivity and the specificity (Trevethan, 2017). This is indicated by the Youden's index (Hughes, 2015) which is highest at a cut-off score of ≥ 0 . Another argument for choosing a cut-off score of ≥ 0 is that it is easier to interpret this score intuitively, which contributes to clinical usability and acceptance.

The strength of the design of the study is the multicenter approach that used stroke patients from two countries to validate the predictive performance of the DePreS. Considering the differences between the samples, the instrument can be used with confidence in other patient populations.

This study has certain limitations. First, the focus on the investigation of the predictive accuracy of the DePreS is not equal to validating the modeling process of an instrument, and therefore does not answer the question whether in a new sample of stroke patients the same set of predictors for depression would result in the same set of predictors in the final model, and whether the performance of the model, in terms of calibration and discrimination, is comparable with the models' performance in the data of the original sample (Steyerberg, 2009; Altman et al., 2009). A second limitation is the high number of patients who were too ill to participate, predominately in the German setting in this study. Since this criterion was based on the clinical judgement of the researchers and clinicians, it might have led to some differences in the sample with regard to the patient's condition at inclusion and might explain the differences found in the severity of stroke symptoms between the two samples. The generalizability of the study findings is limited to patient able to communicate adequately because we exclude patients with cognitive and communicative impairments.

5. Conclusions

This study demonstrates that the DePreS is an adequate instrument for early and reliable identification of stroke patients who are not at risk of MDD in the second months after stroke, which enables clinicians to provide further assessments only in stroke patients who are identified to be at risk of depression. Provided a timely diagnosis followed by adequate therapy, using the DePreS routinely in daily hospital care may reduce the burden of living with untreated consequences of depression entails. Further research for early, reliable risk prediction for depression after a stroke is needed in stroke patients with communicative and/or cognitive impairments (Van Dijk, 2016). Moreover, the development of a care program that starts with an adequate risk prediction during hospital admission and that is embedded in the stroke rehabilitation pathway is essential for improving recovery and psychosocial well-being.

CRedit authorship contribution statement

Julian Hirt: Conceptualization, Data curation, Formal analysis, Writing - review & editing. **Lianne C.J. van Meijeren:** Conceptualization, Data curation, Formal analysis, Writing - review & editing. **Susanne Saal:** Conceptualization, Data curation, Formal analysis, Writing - review & editing. **Thóra B. Hafsteinsdóttir:** Conceptualization, Writing - review & editing. **Jeannette Hofmeijer:** Data curation, Writing - review & editing. **Andrea Kraft:** Data curation, Writing - review & editing. **Gabriele Meyer:** Writing - review & editing. **Janneke M. de Man-van Ginkel:** Conceptualization, Data curation, Formal analysis, Writing - review & editing.

Declaration of Competing Interest

None.

Funding sources

None.

Acknowledgments

We would like to thank Beatrix Kötteritzsch (nurse in charge), Tobias Müller (neurologist), Kathleen Rechlin (student assistant), and Josephine Reiche (case manager) for supporting the recruitment and data collection process in Halle (Saale), Germany.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:[10.1016/j.jad.2020.01.019](https://doi.org/10.1016/j.jad.2020.01.019).

References

Altman, D.G., Vergouwe, Y., Royston, P., Moons, K.G.M., 2009. Prognosis and prognostic research: validating a prognostic model. *BMJ* 338, b605.

American Psychiatric Association, 2000. Diagnostic and statistical manual of mental disorders. Diagnostic and Statistical Manual of Mental Disorders. Text revision (DSM-IV-TR TM), 4th Ed. American Psychiatric Association, Washington DC.

Arboix, A., 2015. Cardiovascular risk factors for acute stroke: risk profiles in the different subtypes of ischemic stroke. *World J. Clin. Case* 3 (5), 418–429.

Balu, S., 2009. Differences in psychometric properties, cut-off scores, and outcomes between the Barthel Index and modified Rankin scale in pharmacotherapy-based stroke trials: systematic literature review. *Curr. Med. Res. Opin.* 25 (6), 1329–1341.

Banks, J.L., Marotta, C.A., 2007. Outcomes validity and reliability of the modified Rankin scale: implications for stroke clinical trials: a literature review and synthesis. *Stroke* 38 (3), 1091–1096.

Bleeker, S.E., Moll, H.A., Steyerberg, E.W., Donders, A.R.T., Derksen-Lubsen, G., Grobbee, D.E., et al., 2003. External validation is necessary in prediction research. *J. Clin. Epidemiol.* 56 (9), 826–832.

Bridges, K.R., Sanderman, R., van Sonderen, E., 2002. An English language version of the social support list: preliminary reliability. *Psychol. Rep.* 90, 1055–1058.

Buderer, N.M.F., 1996. Statistical methodology. I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad. Emerg. Med.* 3 (9), 895–900.

Chiu, D., Peterson, L., Elkind, M.S.V., Rosand, J., Gerber, L.M., Silverstein, M.D., 2010. Comparison of outcomes after intracerebral hemorrhage and ischemic stroke. *J. Stroke Cerebrovasc. Dis.* 19 (3), 225–229.

Collin, C., Wade, D.T., Davies, S., Horne, V., 2009. The Barthel ADL Index. A reliability study. *Int. Disabil. Stud.* 10 (2), 61–63.

de Man-van Ginkel, J.M., Hafsteinsdóttir, T.B., Lindeman, E., Ettema, R.G.A., Grobbee, D.E., Schuurmans, M.J., 2013. In-hospital risk prediction for post-stroke depression: development and validation of the Post-Stroke Depression Prediction Scale. *Stroke* 44 (9), 2441–2445.

Enderby, P.M., Wood, V.A., Wade, D.T., Hewer, R.L., 1987. The Frenchay Aphasia screening test. A short, simple test for aphasia appropriate for non-specialists. *Int. Rehabil. Med.* 8 (4), 166–170.

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. *J. Psychiatr. Res.* 12 (3), 189–198.

Grau, A.J., Weimar, C., Buggle, F., Heinrich, A., Goertler, M., Neumaier, S., et al., 2001. Risk factors, outcome, and treatment in subtypes of ischemic stroke. The German Stroke Data Bank. *Stroke* 32 (11), 2559–2566.

Hackett, M.L., Pickles, K., 2014. Part I: frequency of depression after stroke: an updated systematic review and meta-analysis of observational studies. *Int. J. Stroke* 9 (8), 1017–1025.

Haerting, J., Kluttig, A., Greiser, K.H., Nuding, S., Werdan, K., 2012. Kohortenstudie zu Risikofaktoren für Herz-Kreislauf-Krankheiten in einer urbanen älteren ostdeutschen Allgemeinbevölkerung (CARLA-Studie). *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* 55 (6–7), 795–800.

Hughes, G., 2015. Youden's index and the weight of evidence. *Method. Inf. Med.* 54 (2), 198–199.

Krocker, C., 2006. Aphasie-Schnell-Test (AST). Ein standardisierter Test für die Differenzialdiagnose Aphasie – keine Aphasie – Dysarthrie in Der Akutphase, 3rd ed. Schulz-Kirchner Verlag, Idstein.

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

Mandrekar, J.N., 2010. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* 5 (9), 1315–1316.

Ojaghhihaghghi, S., Vahdati, S.S., Mikaeilpour, A., Ramouz, A., 2017. Comparison of neurological clinical manifestation in patients with hemorrhagic and ischemic stroke. *World J. Emerg. Med.* 8 (1), 34–38.

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., Thomas, R., 2008. Understanding and using sensitivity, specificity and predictive values. *Ind. J. Ophthalmol.* 56, 45–50.

Robinson, R.G., Jorge, R.E., 2016. Post-Stroke depression: a review. *Am. J. Psychiatry* 173 (3), 221–231.

Saal, S., Becker, C., Lorenz, S., Schubert, M., Kuss, O., Stang, A., et al., 2015. Effect of a stroke support service in Germany: a randomized trial. *Top. Stroke Rehabil.* 22 (6), 432–439.

Shi, Y.Z., Xiang, Y.T., Yang, Y., Zhang, N., Wang, S., Ungvari, G.S., et al., 2016. Depression after minor stroke: the association with disability and quality of life—a 1-year follow-up study. *Int. J. Geriatr. Psychiatry* 31 (4), 421–427.

Steyerberg, E.W., 2009. Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating (Statistics for Biology and Health). Springer-Verlag, New York.

Ter Smitten, R., Smeets, R., Van den Brink, W., 2008. Composite International Diagnostic Interview Computerized Version 2.1: Dutch Translation and Adaptation: WHO-CIDI Training en Referentie Centrum. Psychiatrisch Centrum AMC, Amsterdam, The Netherlands.

Trevethan, R., 2017. Sensitivity, specificity, and predictive values: foundations, plabilities, and pitfalls in research and practice. *Front Public Health* 5, 307.

Van Dijk, M.J., De Man-van Ginkel, J.M., Hafsteinsdóttir, T.B., Schuurmans, M.J., 2016. Identifying depression post-stroke in patients with aphasia: a systematic review of the reliability, validity and feasibility of available instruments. *Clin. Rehabil.* 30 (8), 795–810.

Van Raak, A., Groothuis, S., van der Aa, R., Limburg, M., Vos, L., 2010. Shifting stroke care from the hospital to the nursing home: explaining the outcomes of a Dutch case. *J. Eval. Clin. Pract.* 16 (6), 1203–1208.

Van Swieten, J.C., Koudstaal, P.J., Visser, M.C., Schouten, H.J., van Gijn, J., 1988. Interobserver agreement for the assessment of handicap in stroke patients. *Stroke* 19, 604–607.

Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., et al., 2005. Principles of good practice for the translation and cultural adaptation process for patient-reported outcomes (PRO) measures: report of the ISPOR task force for translation and cultural adaptation. *Value in Health* 8 (2), 94–104.

Wittchen, H.-U., 1994. Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. *J. Psychiatr. Res.* 28 (1), 57–84.