Psychological Medicine (2015), **45**, 29–49. © Cambridge University Press 2014
doi:10.1017/S0033291714000336

REVIEW ARTICLE

Screening for mood disorders after stroke: a systematic review of psychometric properties and clinical utility

L.-J. Burton and S. Tyson*

Stroke and Vascular Research Centre, School of Nursing, Midwifery and Social Work, University of Manchester, UK

Background. Routine mood screening is recommended after stroke. However, clinicians report difficulty selecting appropriate tools from the wide range available. We aimed to systematically review the psychometric properties and clinical utility of mood screening tools for stroke survivors.

Method. Electronic databases (AMED, EMBASE, CINAHL, Medline and PsycINFO) were searched to identify studies assessing the sensitivity and specificity of mood screening tools. Tools that demonstrated at least 80% sensitivity and 60% specificity with stroke survivors with identifiable cut-off scores indicating major and/or any mood disorder in at least one study were selected and clinical utility was assessed. Those with high clinical utility (against predefined criteria) were selected for recommendation.

Results. Thirty papers examining 27 screening tools were identified and 16 tools met the psychometric and clinical utility criteria: 10 were verbal self-report tools, four were observational and two incorporated visual prompts for those with communication problems. Only the Stroke Aphasic Depression Questionnaire –Hospital version (SADQ-H) met all the psychometric and utility criteria. The nine-item Patient Health Questionnaire (PHQ-9) can detect major depression and the 15-item Geriatric Depression Scale (GDS-15) can identify milder symptoms; both are feasible to use in clinical practice. The Hospital Anxiety and Depression Scale (HADS) was the only tool able to identify anxiety accurately, but clinical utility was mixed.

Conclusions. Valid and clinically feasible mood screening tools for stroke have been identified but methodological inconsistency prevented recommendations about the optimal cut-off scores.

Received 27 August 2013; Revised 24 January 2014; Accepted 24 January 2014; First published online 27 February 2014

Key words: Mood, review, screening, stroke.

Introduction

Around a third of stroke survivors suffer from low mood, with about 10% receiving a diagnosis of major depression and many experiencing co-morbid anxiety (House *et al.* 1991; Ayerbe *et al.* 2013). Mood disturbance has been linked to greater dependence in activities of daily living, institutionalization and mortality, and poorer quality of life (Kotila *et al.* 1999; House *et al.* 2001; Pohjasvaara *et al.* 2001; Williams *et al.* 2004; Donnellan *et al.* 2010). However, the management of mood disorders post-stroke is often suboptimal; problems are frequently undiagnosed and inadequately treated (Hackett *et al.* 2005) and most survivors report insufficient help to deal with their emotional needs (National Audit Office, 2010). This may result from difficulties in professionals' recognition of the symptoms of low mood in patients with stroke because of the overlap between stroke-related impairments, hospitalization and somatic symptoms of mood disorder (Hart & Morris, 2007).

In acknowledgement of this shortcoming, improving management and access to psychological services has become a priority (NHS Improvement, 2011). Timely diagnosis is an essential element, facilitating early access to treatment and improving prognosis (Jorge *et al.* 2003; Mitchell *et al.* 2009); thus, clinical guidelines recommend that stroke survivors should be routinely screened for the presence of mood disorders, using a validated tool (National Stroke Foundation, 2010; NICE, 2010). Screening rates have improved substantially over the past decade (Bowen *et al.* 2005; National Audit Office, 2010) but 20% of appropriate patients are still not screened (NSSA, 2011). One

* Address for correspondence: S. Tyson, F.C.S.P., M.Sc., Ph.D., Stroke and Vascular Research Centre, School of Nursing, Midwifery and Social Work, University of Manchester, Jean McFarlane Building, Oxford Road, Manchester M13 9PL, UK.
(Email: Sarah.Tyson@Manchester.ac.uk)

issue that affects implementation of the guidelines is the choice of screening tool; clinicians report that lack of knowledge and consensus about the best measures to use are barriers (Burton *et al.* 2013). As part of an ongoing programme of work to select and implement effective measurement tools in stroke rehabilitation, we aimed to systematically review the psychometric properties of tools to screen for mood disorders post-stroke to identify the most suitable for clinical practice. To facilitate uptake in clinical practice, we also aimed to identify optimal cut-off scores for major and any degree of depression and anxiety, and assess clinical utility (or feasibility).

Method

Study identification and selection

Electronic databases (AMED, PsycINFO, CINAHL, Medline and EMBASE) were searched from their inception to May 2013, using the following keywords: stroke OR cerebrovascular accident OR CVA and screen* OR tool OR measure* OR questionnaire OR scale AND mood OR depression OR anxiety OR emotion OR distress. All searches were limited to English language and human studies.

We also searched the reference lists of selected articles, previously published reviews and the Stroke Group of the Cochrane Library. The titles, abstracts and then full texts were screened by two independent reviewers to identify articles that reported validation of tools to screen for low mood, distress and/or anxiety in people with stroke. Articles that assessed the properties of mood screening tools, reported both sensitivity and specificity compared with a gold standard measure and aimed to identify people who needed further evaluation or treatment were selected. We excluded studies of tools that were not designed as a screening tool and intended to make a full assessment of mood or diagnosis, including the Montgomery–Asberg Depression Rating Scale (MADRS; Montgomery & Asberg, 1979), the Hamilton Rating Scale for Depression (HAMD; Hamilton, 1960), the Cornell Scale (Alexopoulos *et al.* 1988), the Geriatric Mental State Examination (GMS; Copeland *et al.* 1987), the Post-Stroke Depression Rating Scale (PSDS; Gainotti *et al.* 1997) and the Symptom Checklist 90 (SCL-90; Derogatis *et al.* 1973). We also excluded studies of tools that assessed generic related constructs such as quality of life; merely involved the validation of a language translation of a tool; were conference papers or abstracts where the data could not be extracted; and where less than 50% of the participants had suffered a stroke or data from people with stroke could not be extracted.

Data extraction

We extracted independently from the selected articles data regarding the participant samples and settings (where available), selection criteria, tools evaluated, type of disorder assessed, and sensitivity and specificity. Cut-off scores for major depression and any degree of depression and anxiety were identified. Positive (PPV) and negative predictive values (NPV) at each cut-off score were calculated from available data where possible. Final data were agreed by consensus with a third party to arbitrate if necessary. Sensitivity $\geqslant 0.8$ and specificity $\geqslant 0.6$ were considered sufficiently accurate. There is often a trade-off between sensitivity and specificity and these (widely used) criteria are considered appropriate for clinical practice, where the costs of failing to identify an individual with difficulties are greater than the costs of further evaluation of those who may not require treatment (Bennett & Lincoln, 2006).

Any cut-off score that did not yield data meeting these criteria in at least one study was excluded and any tools that did not report any cut-off scores with sufficient sensitivity and specificity were rejected.

Tools that met these criteria were then assessed for clinical utility (the feasibility of using a tool in clinical practice) from the original articles (where possible), marketing material (including costs), the tools' authors and instruction manuals. A previously published tool to assess clinical utility of outcome measures (Connell & Tyson, 2012) was reviewed and adapted by a consultation group of occupational therapists and clinical psychologists working in stroke services in a large UK conurbation to reflect their priorities for selecting screening tools. These are summarized as follows: as access to clinical psychology is limited for most stroke survivors, it is important that screening tools can be completed by any member of the multidisciplinary team to identify moderate to severe difficulties for onward referral (NSSA, 2008). This is often undertaken in addition to their traditional workload so screening tools need to be quick and easy to administer, with minimal training requirements. Finally, they need to be inexpensive or, preferably, freely available, particularly because, in the current financial climate, a cheaper tool would be chosen over one incurring costs if it performed equally well in terms of psychometrics.

The final utility criteria and scores were:

(*a*) Time to administer and score the measure: $\leqslant 5$ min (score 2); 6–10 min (score 1); $\geqslant 11$ min (score 0).
(*b*) Initial costs for purchase of the measure (e.g. starter kit including manual): 2=freely available; 1=cost<£100; 0=cost $\geqslant$£100 or unavailable.
(*c*) Additional cost per record form: 1=no additional costs; 0=additional cost or unavailable.

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

(*d*) Need for specialist training to administer and score the measure: 1=no specialist training required; 0=specialist training required.

Summing these scores gave a maximum of six points, with higher scores indicating greater clinical utility. Tools that scored<6 were rejected at this stage.

Results

The searches revealed 30 papers that met the selection criteria, involving 3751 stroke survivors and 27 screening tools. All 27 tools were tested to detect depression and eight were also used to identify anxiety. The tools' progress through the review is summarized in Fig. 1 and detailed below. The selected tools are described in Table 1 and details of the population tested are presented in Table 2. Most of the selected papers recruited their participants through acute admissions to hospital (Parikh *et al.* 1988; Williams *et al.* 2005; Bennett *et al.* 2006; Lightbody *et al.* 2007; Healey *et al.* 2008; Lee *et al.* 2008; Hacker *et al.* 2010; de Man-van Ginkel *et al.* 2012*a*; Kang *et al.* 2012), often consecutively (Watkins *et al.* 2001*a*,*b*, 2007; Aben *et al.* 2002; Benaim *et al.* 2004; Tang *et al.* 2004*a*; Berg *et al.* 2009; Sagen *et al.* 2009; de Man-van Ginkel *et al.* 2012*b*; Tham *et al.* 2012). Others recruited from in-patient rehabilitation facilities (Tang *et al.* 2004*b*,*c*; Turner-Stokes *et al.* 2005; Roger & Johnson-Greene, 2009; de Man-van Ginkel *et al.* 2012*a*), or a mixed in-patient and community-dwelling population (Shinar *et al.* 1986; Lincoln *et al.* 2003; Sivrioglu *et al.* 2009; Turner *et al.* 2012); only one study recruited solely in the community (Agrell & Dehlin, 1989). Three studies involved participants who were also taking part in a clinical trial (O'Rourke *et al.* 1998; Lincoln *et al.* 2003; Williams *et al.* 2005). Most assessments were made in the acute (within 1 month) (Shinar *et al.* 1986; Watkins *et al.* 2001*a*,*b*, 2007; Aben *et al.* 2002; Tang *et al.* 2004*b*,*c*; Bennett *et al.* 2006; Lightbody *et al.* 2007; Lee *et al.* 2008; Berg *et al.* 2009; Roger & Johnson-Greene, 2009; Hacker *et al.* 2010; Kang *et al.* 2012) or subacute (from 1 to 6 months) stages post-stroke (Johnson *et al.* 1995; O'Rourke *et al.* 1998; Tang *et al.* 2004*a*; Turner-Stokes *et al.* 2005; Williams *et al.* 2005; Bennett *et al.* 2006; Watkins *et al.* 2007; Healey *et al.* 2008; Berg *et al.* 2009; Sagen *et al.* 2009; de Man-van Ginkel *et al.* 2012*a*,*b*). Five papers considered mood disorders in the long term (more than 6 months) after stroke (Agrell & Dehlin, 1989; Berg *et al.* 2009; Sivrioglu *et al.* 2009; Kang *et al.* 2012; Turner *et al.* 2012) and one study combined assessment scores of participants between 1 week and 2 years post-stroke (Parikh *et al.* 1988).

The identified studies used a range of criterion measures as the reference gold standard; most used a psychiatrist's opinion (Agrell & Dehlin, 1989; Benaim *et al.* 2004), usually based on a semi-structured interview or assessment tool (Shinar *et al.* 1986; Parikh *et al.* 1988; Johnson *et al.* 1995; O'Rourke *et al.* 1998; Aben *et al.* 2002; Lincoln *et al.* 2003; Tang *et al.* 2004*a*,*b*,*c*; Williams *et al.* 2005; Lightbody *et al.* 2007; Healey *et al.* 2008; Roger & Johnson-Greene, 2009; Sagen *et al.* 2009; Kang *et al.* 2012; de Man-van Ginkel *et al.* 2012*b*; Tham *et al.* 2012; Turner *et al.* 2012). Most used DSM criteria (Shinar *et al.* 1986; Parikh *et al.* 1988; Johnson *et al.* 1995; O'Rourke *et al.* 1998; Aben *et al.* 2002; Lincoln *et al.* 2003; Tang *et al.* 2004*a*,*b*,*c*; Turner-Stokes *et al.* 2005; Williams *et al.* 2005; Lightbody *et al.* 2007; Healey *et al.* 2008; Lee *et al.* 2008; Berg *et al.* 2009; Roger & Johnson-Greene, 2009; Sagen *et al.* 2009; Sivrioglu *et al.* 2009; de Man-van Ginkel *et al.* 2012*b*; Kang *et al.* 2012; Tham *et al.* 2012; Turner *et al.* 2012) for classification of psychiatric disorders, although one also used ICD criteria (Lincoln *et al.* 2003). Others used another a screening or assessment tool as the gold standard (Watkins *et al.* 2001*a*,*b*, 2007; Bennett *et al.* 2006; Hacker *et al.* 2010; de Man-van Ginkel *et al.* 2012*b*).

Screening for depression

Measures identified

Twenty-seven tools met the inclusion criteria and fell into three categories:

(1) Verbal tools for those who could self-report their mood (n=15): the Beck Depression Inventory (BDI; Beck *et al.* 1961), BDI Fast Screen (BDI-S; Beck *et al.* 2000) and BDI – Second Edition (BDI-II; Beck *et al.* 1996), the Center for Epidemologic Studies Depression Scale (CES-D; Radloff, 1977), the 28-item General Health Questionnaire (GHQ-28; Goldberg & Williams, 1988), the Geriatric Depression Scale (GDS, 30 items; Yesavage *et al.* 1983) and its 15-item version (GDS-15; Sheikh & Yesavage, 1986), the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983), the Kessler-10 (K10; Kessler *et al.* 2003), the two-item Patient Health Questionnaire (PHQ-2; Kroenke *et al.* 2003) and the nine-item version (PHQ-9; Spitzer *et al.* 1999), the Stroke Inpatient Depression Inventory (SIDI; Rybarczyk *et al.* 1996), the Wakefield Depression Inventory (WDI; Snaith *et al.* 1971), the Yale question (Lachs *et al.* 1990; Mahoney *et al.* 1994) and the Zung Self-Rating Depression Scale (Zung SDS; Zung, 1965).
(2) Tools involving visual aids (visual analogue scales or pictures) to aid self-report for those with

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

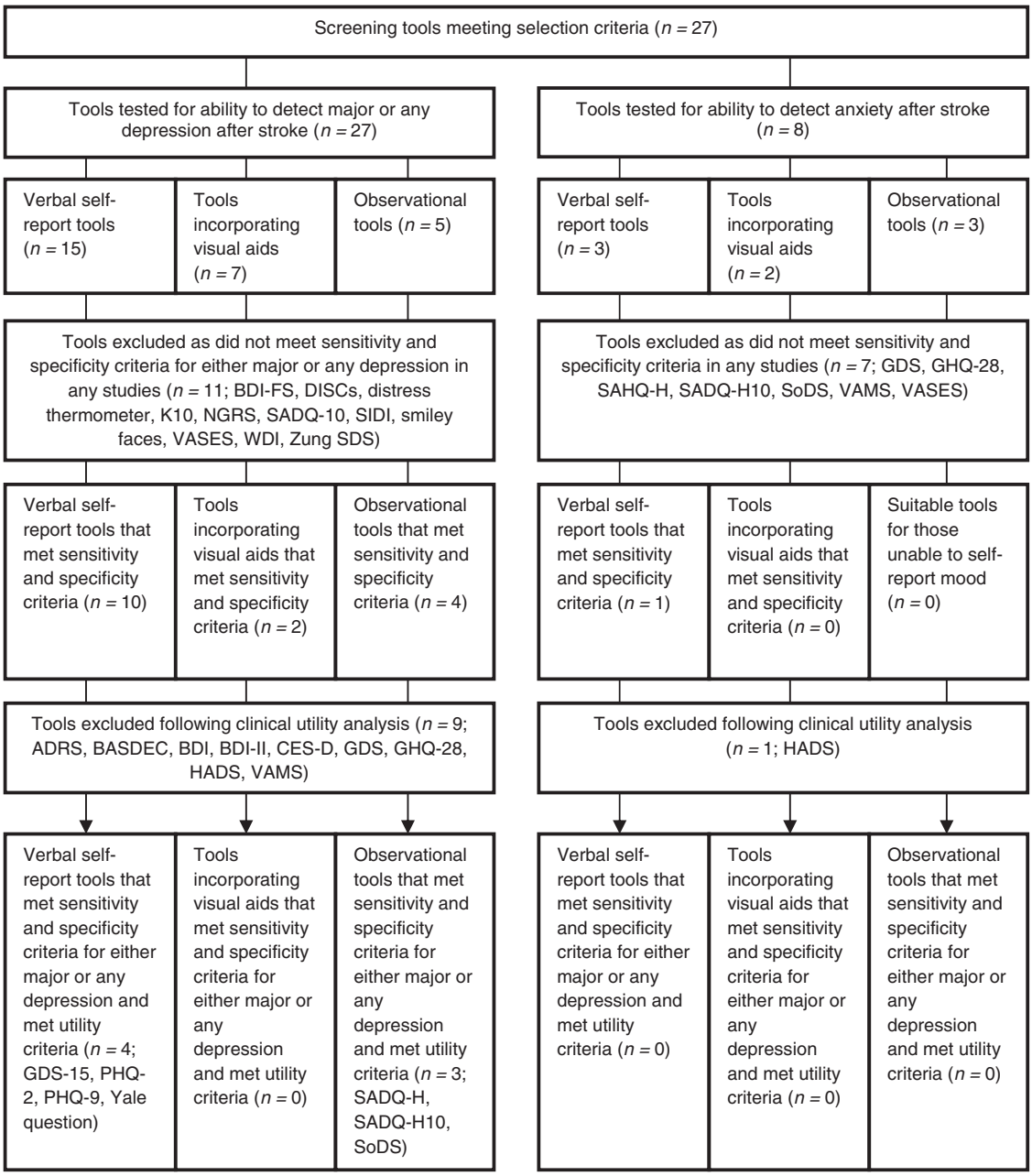

Fig. 1. Study flowchart. ADRS, Aphasia Depression Rating Scale; BASDEC, Brief Assessment Schedule Depression Cards; BDI, Beck Depression Inventory (FS, Fast Screen); CES-D, Center for Epidemologic Studies Depression Scale; DISCs, Depression Intensity Scale Circles; GDS, Geriatric Depression Scale (30-item); GDS-15, 15-item Geriatric Depression Scale; GHQ-28, 28-item General Health Questionnaire; HADS, Hospital Anxiety and Depression Scale; K10, Kessler-10; NGRS, Numeric Graphic Rating Scale; PHQ-2, two-item Patient Health Questionnaire; PHQ-9, nine-item version Patient Health Questionnaire; SADQ-H, Stroke Aphasic Depression Questionnaire – Hospital version; SADQ-10, 10-item SADQ; SADQ-H10, 10-item SADQ-H; SIDI, Stroke Inpatient Depression Inventory; SoDS, Signs of Depression Scale; VAMS, Visual Analogue Mood Scales; VASES, Visual Analogue Self-Esteem Scale; WDI, Wakefield Depression Inventory; Zung SDS, Zung Self-Rating Depression Scale.

communication problems (n=7): the Brief Assessment Schedule Depression Cards (BASDEC; Adshead *et al.* 1992), the Depression Intensity Scale Circles (DISCs; Turner-Stokes *et al.* 2005), the distress thermometer (Holland *et al.* 2011), the Numbered Graphic Rating Scale (NGRS; Turner-Stokes *et al.* 2005), 'smiley faces' (Lee *et al.* 2008), Visual Analogue Mood Scales (VAMS) 'sad item' (Stern, 1997) and Visual Analogue Self-Esteem Scales (VASES; Brumfitt & Sheeran, 1999).

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

Table 1. *Brief description of each identified measure meeting psychometric criteria in at least one validation study*

Measure	Brief description of measure	Type of tool	Time to administer (min)	Training required?	Initial costs	Recurring costs	Clinical utility total score/6
ADRS (Benaim *et al.* 2004)	Observer-rated. A health professional rates the patient's behaviour based on observation and/or interview on nine items. Items are scored between 0 and 6 points, and totalled to give a score of up to 32 points. Includes somatic symptoms	Observational	Not reported	Yes	Freely available	N.A.	3
BDI (Beck *et al.* 1961)	Self-report. Patients choose one of four statements (graded according to severity) for 21 items regarding symptoms and attitudes over the previous week. Item scores are totalled with a maximum of 63. Includes somatic symptoms	Self-report	<10	No	Unavailable; superseded by BDI-II	Unavailable	2
BDI-II (Beck *et al.* 1996)	Self-report. Patients choose one of four statements for 21 items (graded according to severity) regarding symptoms and attitudes over the previous 2 weeks. Item scores are totalled to provide a maximum of 63. Includes somatic symptoms	Self-report	<10	No	Must be purchased: <£100 for complete kit	Ongoing costs for record forms/ interpretative reports	3
BASDEC (Adshead *et al.* 1992)	Self-report with visual prompts. Patients indicate their present feelings by sorting cards with pictorial representations of 19 items into piles labelled 'true' or 'false'. Includes somatic items	Visually aided self-report	3–4	Minimal	Unavailable	Unavailable	3
CES-D (Radloff, 1977)	Self-report. Patients rate the frequency of 20 mood-related statements over the past week on a four-point scale from 'rarely or none of the time' to 'most or all of the time'. Item scores are summed to a maximum of 60. Includes somatic symptoms	Self-report	<15	No	Freely available	N.A.	4
GHQ-28 (Goldberg & Williams, 1988)	Self-report. Patients choose one of four statements for 28 items regarding symptoms of distress over the past 2 weeks, scored on a three-point scale, which is totalled to give a maximum of 84 points. Alternatively, a binary method ('not at all'=0, 'more than usual'=1) with a maximum of 28 points. Includes somatic symptoms	Self-report	<5	No	Must be purchased: >£100 for user guide plus record forms	Ongoing costs for record forms	3

Table 1 (*cont.*)

Measure	Brief description of measure	Type of tool	Time to administer (min)	Training required?	Initial costs	Recurring costs	Clinical utility total score/6
GDS (Yesavage *et al.* 1983)	Self-report. Patients give 'yes/no' responses to 30 items regarding their experience over the previous week. Scores are totalled with a maximum of 30. Minimal somatic items	Self-report	8–10	No	Freely available	N.A.	5
GDS-15 (Sheikh & Yesavage, 1986)	Self-report. Patients give 'yes/no' responses to 15 items regarding their experience over the previous week. Scores are totalled with a maximum of 15. Minimal somatic symptoms	Self-report	5	No	Freely available	N.A.	6
HADS (Zigmond & Snaith, 1983)	Self-report. Patients rate their agreement with seven anxiety and seven depression items over the previous week on a four-point scale to give a maximum of 42 (or 21 for each subscale). Does not include somatic symptoms	Self-report	2–6	No	Must be purchased: >£100 for complete kit	Ongoing costs for record forms/digital administrations	3
PHQ-2 (Kroenke *et al.* 2003)	Self-report. Patients use a four-point scale for two items regarding the presence of depressive symptoms over the previous 2 weeks. Scores are summed to give a maximum of 6. Does not include somatic symptoms	Self-report	<2	No	Freely available	N.A.	6
PHQ-9 (Spitzer *et al.* 1999)	Self-report. Patients use a four-point scale for nine items regarding the presence of depressive symptoms over the previous 2 weeks. Scores are summed to give a maximum score of 27. Includes somatic symptoms	Self-report	3–5	No	Freely available	N.A.	6
SoDS (Hammond *et al.* 2000)	Observer rated. An observer rates the occurrence of six behaviours associated with low mood with a 'yes/no' response. Scores are summed to provide a score from 0 to 6. Minimal somatic symptoms	Observational	<5	No	Freely available	N.A.	6
SADQ-H10 (Lincoln *et al.* 2000)	Observer rated. An observer rates the frequency of hospitalized patients' behaviour on 10 items associated with low mood on a scale of 0–3 to give a total of 30. Includes somatic symptoms	Observational	<5	No	Freely available	N.A.	6
SADQ-H (Lincoln *et al.* 2000)	Observer rated. An observer rates the frequency of hospitalized patients' behaviour on 21 items associated with low mood on a scale of 0–3 to give a total of 63. Includes somatic symptoms	Observational	<5	No	Freely available	N.A.	6

VAMS 'sad item' (Stern, 1997)	Self-report. Patients are presented with a vertical 10-cm line with a cartoon sad face with a verbal descriptor and a neutral face at opposite ends. The patient indicates the severity of their feeling by the position on the line, which is measured. Does not include somatic items	Visually aided self-report	<5	No	Must be purchased: >£100 for kit (includes eight mood states)	Ongoing costs for response booklets	3
Yale question (Lachs *et al.* 1990; Mahoney *et al.* 1994)	Self-report. The patient is asked about current feelings of sadness or depression, with a 'yes/no' response format. Does not include somatic symptoms	Self-report	<5	No	Freely available	N.A.	6

ADRS, Aphasia Depression Rating Scale; BASDEC, Brief Assessment Schedule Depression Cards; BDI, Beck Depression Inventory; BDI-II, Beck Depression Inventory Second Edition; CES-D, Center for Epidemiologic Studies Depression Scale; GDS, Geriatric Depression Scale (30-item); GDS-15, 15-item Geriatric Depression Scale; GHQ-28, 28-item General Health Questionnaire; HADS, Hospital Anxiety and Depression Scale; PHQ-2, two-item Patient Health Questionnaire; PHQ-9, nine-item Patient Health Questionnaire; SADQ-H, Stroke Aphasic Depression Questionnaire – Hospital version; SADQ-H10, 10-item Stroke Aphasic Depression Questionnaire – Hospital version; SoDS, Signs of Depression Scale; VAMS, Visual Analogue Mood Scale; N.A., not available.

(3) Observer-rated measures for people unable to self-report their mood due to communication or cognitive impairments (n=5): the Aphasia Depression Rating Scale (ADRS; Benaim *et al.* 2004), the Stroke Aphasic Depression Questionnaire (SADQ; Sutcliffe & Lincoln, 1998), SADQ – Hospital version (SADQ-H; Lincoln *et al.* 2000) and its 10-item version (SADQ-H10; Lincoln *et al.* 2000), and the Signs of Depression Scale (SoDS; Hammond *et al.* 2000).

The sensitivity and specificity for each cut-off score for the depression screening tools are detailed in Table 2. Eleven tools did not meet sensitivity and specificity criteria at any cut-off scores and were rejected (see Fig. 1). This left 16 tools that could accurately detect depression in people with stroke. Ten were verbal self-report tools, two used visual aids and four were observational measures; these are described briefly in the following sections.

Verbal self-report tools

The verbal self-report tools were all questionnaires completed by the person with stroke or verbally by interview with a health-care professional, apart from the single-question 'Yale' tool (Lachs *et al.* 1990; Mahoney *et al.* 1994), which was reported verbally. All were developed to identify problems in a general population and subsequently applied to stroke.

The questionnaires ranged from two (PHQ-2) to 30 questions (GDS). Five used four- or five-point Likert scales for people to rate the severity or frequency of their symptoms (GHQ-28, HADS, PHQ-2, PHQ-9 and CES-D). Three used a 'yes/no' response format (GDS, GDS-15 and Yale question) whereas the BDI and its second edition BDI-II included four multiple choice statements that graded symptom severity. Most scales were designed to detect depression alone; exceptions were the GHQ-28, which is a measure of general distress, and the HADS, which includes separate scales for depression and anxiety. The time scale over which people rated their mood varied: present mood state (Yale question), the previous week (BDI, CES-D, GDS, GDS-15 and HADS) and the past 2 weeks (BDI-II, GHQ-28, PHQ-2 and PHQ-9) in line with DSM (APA, 1987, 1994, 2000). Most tools include somatic items (BDI, BDI-II, CES-D, GHQ-28 and PHQ-9) whereas others seek to limit their inclusion (GDS and GDS-15) or omit them completely (HADS, PHQ-2 and Yale).

Tools incorporating visual aids

Two tools incorporated visual prompts to aid self-report. The BASDEC consists of 19 cards, each with a single printed statement, that the person sorts into

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

Table 2. *Descriptions of the selected papers*

Study	Stroke participants	Exclusion criteria	Tool	Criterion measure	Time post-stroke assessment was made	Cut-off score	Sensitivity	Specificity	PPV	NPV
Aben *et al.* 2002	202 consecutive admissions with first infarct, mean age=69 years	Co-morbid intracerebral disease, severe communication or cognitive problems, neuropsychiatric diagnosis	BDI	SCID DSM-IV major depression	1 month (n=166)	9/10	80	61	22	96
			HADS-T	SCID DSM-IV major depression	1 month (n=171)	10/11	92	65	30	98
				SCID DSM-IV major and minor depression		10/11	87	70	45	95
			HADS-A	SCID DSM-IV major and minor depression	1 month (n=171)	4/5	89	72	64	92
			HADS-D	SCID DSM-IV major depression	1 month (n=171)	7/8	73	82	41	95
				SCID DSM-IV major and minor depression		6/7	73	79	51	91
Agrell & Dehlin, 1989	40 elderly patients, mean age=80 (range 61–90) years	Dysphasia, severe cognitive impairment	GDS	Psychiatric interview; any depression	14 months (range 4 months to 2.5 years)	9/10	88	64	58	88
Benaim *et al.* 2004	50 rehabilitation in-patients, mean age=60 (S.D.=13) years	None	ADRS	Psychiatrist diagnosis of depression	60±45 days	8/9	83	71		
Bennett *et al.* 2006	100 acute in-patients, median age=72 (IQR=65–76) years	Dementia, non-English speaking, sensory defect, unable to complete the HADS	SoDS	HADS-D 7/8	2–4 weeks (n=79)	1/2	86	62		
			SADQ-H10			5/6	100	78		
			SADQ-H			17/18	100	81		
			VAMS 'sad item'			22/23	88	62		

Berg *et al.* 2009	100 consecutive admissions with first stroke, mean age=55 years	>70 years, history of alcohol abuse, dementia or psychosis, current antidepressant treatment, severe concomitant disease	BDI	DSM-III-R 2 major depression	2 weeks	6/7	100	46		
						9/10	80	76		
						13/14	80	88		
					2 months	6/7	100	60		
						9/10	100	76		
						13/14	43	89		
					6 months (n=96)	6/7	100	66		
						9/10	71	83		
						13/14	29	93		
					12 months (n=93)	6/7	100	68		
						9/10	100	86		
						13/14	63	91		
					18 months (n=92)	6/7	100	64		
						9/10	83	84		
						13/14	50	93		
Crabtree *et al.* 2012	32 patients with stroke	Aphasia	HADS-D	MINI major depression	6 weeks	5/6	100	100		
de Man-van Ginkel *et al.* 2012*a*	55 hospital in-patients, mean age=65 (S.D.=15) years	Severe physical illness, cognitive or communication impairment, psychiatric diagnosis, discharged<2 weeks post-stroke	PHQ-9	GDS-15; 5/6 (major depression)	60 days (n=53)	9/10	100	86	50	100
			PHQ-2			1/2	100	77	38	100
de Man-van Ginkel *et al.* 2012*b*	164 consecutive hospital admissions, mean age=71 (S.D.=14) years	Severe cognitive or communication impairment, co-morbid or pre-morbid physical or psychiatric illness, non-English or Dutch speaking	PHQ-9	CIDI major depression	7 weeks (range 5–9 weeks)	9/10	80	78	34	97
			PHQ-2			1/2	75	76	30	96
Hacker *et al.* 2010	125 acute in-patients, mean age=73 (S.D.=13) years	<1 week post-stroke, non-English speaking, dysphasia	SADQ-H10	BASDEC 6/7	Mean=28 days	5/6	70	69		
Healey *et al.* 2008	49 rehabilitation in-patients, mean age=79 (S.D.=7) years	<65 years old, severe physical illness, cognitive or communication impairment	HADS-D	SCID DSM-IV major depression	Median=41 days (range 16–113)	7/8	86	69	32	97
				SCID DSM-IV major and minor depression		7/8	62	69	42	83
			BASDEC	SCID DSM-IV major depression		6/7	100	95	78	100

Table 2 (*cont.*)

Study	Stroke participants	Exclusion criteria	Tool	Criterion measure	Time post-stroke assessment was made	Cut-off score	Sensitivity	Specificity	PPV	NPV
Johnson *et al.* 1995	204 patients, mean age=71 (range=23–95) years	No details of exclusion criteria	GHQ-28	PAS DSM-III any depressive disorder (dysthymia/major depression)	4 months (*n*=66)	4/5	89	75	47	96
			GDS		4 months (*n*=120)	9/10	84	50	44	87
						10/11	84	66	53	90
			HADS-D		4 months (*n*=93)	3/4	94	32	25	96
						4/5	83	44	26	92
			HADS-A	PAS DSM-III anxiety	4 months (*n*=93)	6/7	57	56	28	82
Kang *et al.* 2012	423 patients with acute stroke, mean age=65 (S.D.=10) years	Severe physical or psychiatric conditions, communication or severe cognitive difficulties	BDI	MINI DSM-IV major depression	2 weeks	10/11	92	84		
					12 months (*n*=288)	10/11	96	84		
				MINI DSM-IV any depressive disorder	2 weeks	7/8	84	76		
					12 months (*n*=288)	7/8	84	76		
			HADS-D	MINI DSM-IV major depression	2 weeks	6/7	83	76		
					12 months (*n*=288)	6/7	93	81		
				MINI DSM-IV any depressive disorder	2 weeks	4/5	85	72		
					12 months (*n*=288)	4/5	87	74		
Lee *et al.* 2008	253 patients with ischaemic stroke	Age<50 years, non-Chinese speaking, previous stroke, aphasia	GDS-15	DSM-IV any depressive disorder	1 month	4/5	84	77	85	75
Lightbody *et al.* 2007	71 acute in-patients, median age=70 (range=59–76) years	Severe physical illness, no carer/relative available	SoDS; nurse-rated	SCID DSM-IV depression	In-patients	1/2	64	61		
Lincoln *et al.* 2003	143 in-patient or community-dwelling patients, 123 were participating in a trial of CBT. Mean age=66 (S.D.=14) years	Dementia, non-English speaking, in residential care pre-stroke, severe communication problems, score>10 on the BDI or>18 on the WDI (trial participants only)	BDI-II	SCAN DSM-III-R major depression	Unknown	11/12	91	30		
						13/14	91	48		
			GHQ-28	SCAN DSM-III-R major depression		10/11	81	63		
						11/12	81	68		
				SCAN ICD-10 major or mild depression		4/5	98	35		
						7/8	85	61		

O'Rourke *et al.* 1998	105 community dwelling patients, median age=68 (range=18–90) years	Severe cognitive or communication impairment	HADS-D	SADS DSM-IV any depressive disorder	6 months post-stroke	6/7	80	79		
			HADS-A	SADS DSM-IV any anxiety disorder		6/7	83	68		
Parikh *et al.* 1998	80 patients evaluated in hospital, at 3 and 6 months, 1 and 2 years; mean age=58 (S.D.=13.5) years	Severe communication problems	CES-D	PSE DSM-III major or minor depression	1 week to 2 years post-stroke	15/16	86	90	80	93
Roger & Johnson-Greene, 2009	67 rehabilitation in-patients. Mean age=71 (S.D.=9) years	Severe sensory, communication or cognitive impairment, non-English speaking, pre-morbid neurological illness	CES-D	SCID DSM-IV major or minor depression	Mean=8 (S.D. =4.5) days	15/16	60	76	28	38
			GDS-15			4/5	46	90	23	49
Sagen *et al.* 2009	104 consecutive admissions, mean age=65 (S.D.=14) years	Non-Norwegian speaking, severe communication or cognitive problems, psychosis, terminal illness	HADS-T	SCID DSM-IV any depressive disorder	4 months (n=101)	9/10	90	83	55	97
						10/11	90	83	55	97
			HADS-D			2/3	84	66	36	95
						3/4	84	73	42	95
						4/5	79	82	50	94
						6/7	58	92	61	90
						7/8	58	94	69	91
			HADS-T	SCID DSM-IV anxiety		5/6	83	60	38	92
			HADS-A	SCID DSM-IV anxiety		3/4	83	65	41	93
						6/7	70	83	55	90
Shinar *et al.* 1986	27 consecutive patients, median age=56 (range=28–73) years	Aphasia	CES-D	PSE DSM-III any depressive disorder	7–10 days	15/16	73	100	100	84
Sivrioglu *et al.* 2009	85 in or out-patients, mean age=60 (range=25–87) years	Illiteracy, aphasia, history of depression, co-morbid neurological or psychiatric disorder (including major depression)	GDS	DSM-IV-TR minor depression	237 days (17–704)	7/8	80	61	60	81
						8/9	80	61	60	81
						9/10	75	67	63	78
						10/11	69	75	67	77
Tang *et al.* 2004*a*	127 consecutive admissions, mean age=76 (S.D.=6) years	Stroke>7 days before admission,<65 years old, non-Cantonese speaking, other neurological disease, communication or severe cognitive problems	GDS-15	SCID DSM-IV any depressive disorder	14 weeks	6/7	89	73	37	98

Table 2 (*cont.*)

Study	Stroke participants	Exclusion criteria	Tool	Criterion measure	Time post-stroke assessment was made	Cut-off score	Sensitivity	Specificity	PPV	NPV
Tang *et al.* 2004*b*	60 rehabilitation in-patients, mean age=71 (S.D.=9) years	As above but any age>18 years	HADS-D	SCID DSM-III-R any depressive disorder	23 (S.D.=6) days	3/4	86	78	55	93
Tang *et al.* 2004*c*	100 rehabilitation in-patients, mean age=74 (S.D.=7) years	As above except<60 years old and any time since stroke	HADS-D	SCID DSM-III-R any depressive disorder	3–4 weeks post-stroke	6/7 7/8	88 82	53 58	28 29	96 95
Tham *et al.* 2012	114 consecutive admissions to hospital rehabilitation unit	Exclusion criteria unknown	PHQ-2	SCID major depressive episode and adjustment disorder	Unknown	2/3	62	95		
Turner *et al.* 2012	72 in- and out-patients, mean age=67 (S.D.=13) years	<3 weeks post-stroke, non-English speaking, severe cognitive or physical impairment	BDI-II	SCID DSM-IV major depressive episode	Median=14 months (3 weeks to 540 months)	11/12	92	71	41	98
						13/14	85	75	42	96
			HADS-T			10/11	92	63	35	97
						14/15	85	75	42	96
			HADS-D			5/6	92	68	39	98
						7/8	62	83	44	91
			PHQ-9			6/7	85	63	33	95
						9/10	69	78	41	92
			PHQ-2			1/2	77	63	31	93
						2/3	69	83	47	92
Turner-Stokes *et al.* 2005	114 in-patients with acquired brain injury (76 with stroke), mean age=43 (S.D.=15) years	None	Yale question	DSM-IV major or minor depression	Median=12 weeks	0/1	68	73	62	78
Watkins *et al.* 2001*a*	110 acute in-patients, median age=75, (range=70–79) years	Severe cognition or communication problems, discharged<2 weeks of stroke	Yale question	MADRS 6/7 any depressive disorder	2 weeks	0/1	86	78	82	82
Watkins *et al.* 2001*b*	137 consecutive admissions, median age=74 (IQR=68–79) years	Severe cognitive or communication problems	SoDS	MADRS depression 6/7	2 weeks	1/2	70	56	65	62

Watkins *et al.* 2007	122 consecutive admissions, median age=74 years	Severe cognition or communication problems, discharged<2 weeks of stroke	Yale question	MADRS 6/7 any depressive disorder	2 weeks	0/1	86	84		86	84
					3 months	0/1	95	89		93	91
Williams *et al.* 2005	316 hospital patients enrolled in clinical trial	Severe cognitive or communication impairment, not depressed	PHQ-9	SCID DSM-IV major depression	1–2 months	9/10	91	89		80	95
			PHQ-2			2/3	83	84		72	91

ADRS, Aphasia Depression Rating Scale; BASDEC, Brief Assessment Schedule Depression Cards; BDI, Beck Depression Inventory; BDI-II, Beck Depression Inventory, Second Edition; CES-D, Center for Epidemiologic Studies Depression Scale; CIDI, Composite International Diagnostic Interview (Robins *et al.* 1988); GDS, Geriatric Depression Scale (30-item); GDS-15, 15-item Geriatric Depression Scale; GHQ-28, 28-item General Health Questionnaire; HADS-A, Hospital Anxiety and Depression Scale – Anxiety subscale; HADS-D, Hospital Anxiety and Depression Scale – Depression subscale; HADS-T, Hospital Anxiety and Depression Scale – Total score; IQR, interquartile range; MADRS, Montgomery–Asberg Depression Rating Scale; MINI, Mini International Neuropsychiatric Interview (Sheehan *et al.* 1998); NPV, negative predictive value; PAS, Psychiatric Assessment System; PHQ-2, two-item Patient Health Questionnaire; PHQ-9, nine-item Patient Health Questionnaire; PPV, positive predictive value; PSE, Present State Examination (Wing *et al.* 1974); SADQ-H, Stroke Aphasic Depression Questionnaire – Hospital version; SADQ-H10, 10-item Stroke Aphasic Depression Questionnaire – Hospital version; SADS, Schedule for Affective Disorders and Schizophrenia (Endicott & Spitzer, 1978); SCAN, Schedules for Clinical Assessment in Neuropsychiatry (WHO, 1992); SCID, Structured Clinical Interview for DSM-IV (First *et al.* 1995); S.D., standard deviation; SoDS, Signs of Depression Scale; VAMS, Visual Analogue Mood Scale.

'true' and 'false' piles to describe their present mood state. The VAMS 'sad item' involves a vertical line with a cartoon sad face at one end with a verbal descriptor and a neutral face at the opposite end. The person points to where they see themselves on the scale at the present time. Neither measure incorporated somatic symptoms of low mood.

Observational measures

Four observational measures were identified for use with people who could not self-report. The SADQ-H and the SADQ-H10 require an observer to rate the frequency of behaviours related to low mood over the previous week, using a four-point Likert scale, with elevated scores on two consecutive weeks suggesting low mood. The SoDS requires the observer to rate the presence of behaviours related to low mood over the previous week in a 'yes/no' response format. The ADRS uses an interview with, or observation of, the person to indicate severity. All of the measures except for the SoDS incorporated somatic mood-related symptoms.

Optimal cut-off scores

Having selected screening tools that could accurately identify depression, the optimal cut-off score to detect either major depression or any depressive disorder was explored (Tables 3 and 4). Sufficient data were available for only four verbal screening tools to identify effective cut-off scores for both major depression and any depressive disorder (BDI, GHQ-28, HADS total score and HADS depression subscale). Of the observational tools and those incorporating visual prompts, only the ADRS, BASDEC, SADQ-H and VAMS 'sad item' demonstrated effective cut-off scores although no distinction was made between severity of depression. For all tools, data were reported for multiple cut-off scores, but the results were so highly varied that optimal cut-off scores could not be identified for any of them.

Clinical utility

The 16 selected screening tools were then assessed for clinical utility (Table 1). All of the tools took less than 15 min to administer; however, some of the reported administration times were not specific for people with stroke, who often have communication and cognitive problems and may therefore take longer to complete the mood screen. All except the ADRS could be administered without specialist training, but the time needed to complete this tool could not be ascertained. Ten tools were freely available (ADRS, CES-D, GDS, GDS-15, PHQ-2, PHQ-9, SADQ-H, SADQ-H10, SoDS

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

Table 3. *Sensitivity and specificity of verbal self-report tools to detect major depression or any depressive disorder*

Screening instrument	Major depression				Any depressive disorder			
	Study	Cut-off	Sensitivity (%)	Specificity (%)	Study	Cut-off	Sensitivity (%)	Specificity (%)
BDI	Berg *et al.* 2009	6/7	100	46–68	Kang *et al.* 2012	7/8	**84**	**76**
	Aben *et al.* 2002; Berg *et al.* 2009	9/10	71–100	18–86				
	Kang *et al.* 2012	10/11	**92–96**	**84**				
	Berg *et al.* 2009	13/14	29–80	88–93				
BDI–II	Lincoln *et al.* 2003; Turner *et al.* 2012	11/12	91–92	30–71				
	Johnson *et al.* 1995; O'Rourke *et al.* 1998; Sagen *et al.* 2009	13/14	85–91	48–75				
CES-D					Shinar *et al.* 1986; Parikh *et al.* 1988; Roger & Johnson-Greene, 2009	15/16	60–86	76–100
GHQ-28	Lincoln *et al.* 2003	10/11	**81**	**63**	Johnson *et al.* 1995; Lincoln *et al.* 2003	4/5	89–98	35–75
	Lincoln *et al.* 2003	11/12	**81**	**68**	Lincoln *et al.* 2003	7/8	**85**	**61**
GDS					Sivrioglu *et al.* 2009	7/8	**80**	**61**
					Sivrioglu *et al.* 2009	8/9	**80**	**61**
					Agrell & Dehlin, 1989; Johnson *et al.* 1995; Sivrioglu *et al.* 2009	9/10	75–88	44–88
					Johnson *et al.* 1995; Sivrioglu *et al.* 2009	10/11	69–84	66–75
GDS-15					Lee *et al.* 2008; Roger & Johnson-Greene, 2009	4/5	46–84	77–90
					Tang *et al.* 2004*a*	6/7	**89**	**73**
HADS-Total	Aben *et al.* 2002; Turner *et al.* 2012	10/11	**92**	**63–65**	Sagen *et al.* 2009	9/10	**90**	**83**
	Turner *et al.* 2012	14/15	**85**	**75**	Aben *et al.* 2002; Sagen *et al.* 2009	10/11	**87–90**	**70–83**
HADS-Anxiety	Aben *et al.* 2002	4/5	92	56	Aben *et al.* 2002	4/5	**89**	**72**
HADS-Depression	Crabtree *et al.* 2012; Turner *et al.* 2012	5/6	**92–100**	**68–100**	Sagen *et al.* 2009	2/3	**84**	**66**
	Healey *et al.* 2008; Kang *et al.* 2012	6/7	**83–93**	**76–81**	Johnson *et al.* 1995; Tang *et al.* 2004*b*; Sagen *et al.* 2009	3/4	84–94	32–78
	Aben *et al.* 2002; Healey *et al.* 2008; Turner *et al.* 2012	7/8	62–86	69–83	Johnson *et al.* 1995; Sagen *et al.* 2009; Kang *et al.* 2012	4/5	79–87	44–82
					O'Rourke *et al.* 1998; Aben *et al.* 2002; Tang *et al.* 2004*c*; Sagen *et al.* 2009	6/7	58–88	53–92
					Tang *et al.* 2004*c*; Healey *et al.* 2008; Sagen *et al.* 2009	7/8	58–82	58–94

PHQ-2	de Man-van Ginkel *et al.* 2012*a*,*b*; Turner *et al.* 2012	1/2	75–100	76–95			
	Williams *et al.* 2005; de Man-van Ginkel *et al.* 2012*a*; Turner *et al.* 2012	2/3	67–83	83–91			
PHQ-9	Turner *et al.* 2012	6/7	**85**	**63**			
	Williams *et al.* 2005; de Man-van Ginkel *et al.* 2012*a*,*b*; Turner *et al.* 2012	9/10	69–100	78–89			
Yale question					Watkins *et al.* 2001*a*, 2007; Turner-Stokes *et al.* 2005	0/1	68–95	73–89

BDI, Beck Depression Inventory; BDI-II, Beck Depression Inventory, Second Edition; CES-D, Center for Epidemiologic Studies Depression Scale; GDS, Geriatric Depression Scale (30-item); 15-item GDS-15, Geriatric Depression Scale; GHQ-28, 28-item General Health Questionnaire; HADS, Hospital Anxiety and Depression Scale; PHQ-2, two-item Patient Health Questionnaire; PHQ-9, nine-item Patient Health Questionnaire.

Cut-off scores with at least one study reporting 80% sensitivity and 60% specificity criteria have been included. Where no studies reached the sensitivity and specificity criteria at a cut-off score, the cut-off score has been removed. Sensitivity and specificity levels that reach the selection criteria are highlighted in bold.

and the Yale question), one (BDI-II) required the purchase of an initial starter kit, which costs under £100 (approximately US$150), whereas the three other tools cost over £100 (GHQ-28, HADS and VAMS). All tools that incurred costs for initial purchase also generated additional costs per person screened, either for paper record forms or for electronic reports. Two tools (BASDEC and BDI) were not available for use.

Seven tools met all clinical utility criteria (see Fig. 1): four verbal measures (PHQ-2, PHQ-9, GDS-15 and Yale) and three observational measures (SADQ-H, SADQ-H10 and SoDS). None of the tools that incorporated visual prompts to aid self-report met the clinical utility criteria (Table 5).

Screening for anxiety

Eight screening tools tested to detect anxiety post-stroke were identified: three were verbal tools (GDS, GHQ-28 and HADS), three were observational (SADQ-H, SADQ-H10 and SoDS) and two used visual analogue scales (VAMS and VASES; Fig. 1). Only the HADS anxiety subscale was designed specifically to measure anxiety, although the total HADS has also been used and these were the only measures to yield adequate sensitivity and specificity data at any cut-point (Table 4). None of the visual analogue scales or observational measures met selection criteria.

Clinical utility

Clinical utility for the HADS is mixed (Table 1): it can be administered in under 5 min with minimal training for staff, but it incurs both initial and recurrent costs and therefore does not meet clinical utility criteria.

Discussion

Our extensive search strategies identified a wide range of tools to screen for mood disorders after stroke, but only the SADQ-H met both the psychometric and clinical utility criteria for both major depression and any depressive disorder; none of the tools incorporating verbal self-report or visual aids met all these criteria. The HADS, BDI, ADRS, BASDEC, GHQ-28 and VAMS 'sad item' all yielded good psychometric data for both major depression and any depressive disorder, indicating that they could accurately identify stroke survivors who needed further assessment and possibly treatment for depression. However, the BASDEC and the BDI are currently unavailable, the former could not be located and the latter was superseded by BDI-II; the other tools incur financial costs (although 'pirate' copies can be downloaded from the internet), require specialist training and/or are time-consuming to administer. Two tools met the clinical

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

Table 4. *Sensitivity and specificity of screening tools for post-stroke depression for those who are unable to self-report or require visual prompts*

Screening instrument	Cut-off score for depression	Study	Sensitivity (%)	Specificity (%)
ADRS	8/9	Benaim *et al.* 2004	**83**	**71**
BASDEC	6/7	Healey *et al.* 2008	**100**	**95**
SADQ-H	17/18	Bennett *et al.* 2006	**100**	**81**
SADQ-H10	5/6	Bennett *et al.* 2006	70–100	69–78
SoDS	1/2	Watkins *et al.* 2001*b*; Bennett *et al.* 2006; Lightbody *et al.* 2007	64–86	56–62
VAMS 'sad item'	22/23	Bennett *et al.* 2006	**88**	**62**

ADRS, Aphasia Depression Rating Scale; BASDEC, Brief Assessment Schedule Depression Cards; SADQ-H, Stroke Aphasic Depression Questionnaire – Hospital version; SADQ-H10, 10-item Stroke Aphasic Depression Questionnaire – Hospital version; SoDS, Signs of Depression Scale; VAMS, Visual Analogue Mood Scale.

Cut-off scores with at least one study reporting 80% sensitivity and 60% specificity criteria have been included. Where no studies reached the sensitivity and specificity criteria at a cut-off score, the cut-off score has been removed. Sensitivity and specificity levels that reach the selection criteria are highlighted in bold.

Table 5. *Sensitivity and specificity of verbal self-report screening tools for post-stroke anxiety*

Screening instrument	Cut-off score for anxiety	Study	Sensitivity (%)	Specificity (%)
HADS-Total	5/6	Sagen *et al.* 2009	**83**	**60**
HADS-Anxiety	3/4	Sagen *et al.* 2009	**83**	**65**
subscale	6/7	Johnson *et al.* 1995; O'Rourke *et al.* 1998; Sagen *et al.* 2009	57–83	56–83

HADS, Hospital Anxiety and Depression Scale.

Cut-off scores with at least one study meeting 80% sensitivity and 60% specificity criteria have been included. Where no studies reached the sensitivity and specificity criteria at a cut-off score, the cut-off score has been removed. Sensitivity and specificity that reach selection criteria are highlighted in bold.

utility criteria and yielded acceptable psychometrics at a specific cut-off score for either major depression or any depressive disorder; the GDS-15 can detect any depressive disorder (but not specifically major depression) and so may be best used as an initial screen to identify people with stroke who require further evaluation. The PHQ-9 can detect major depression although sensitivity drops to 78% in identifying milder symptoms (Williams *et al.* 2005). The HADS (both total score and anxiety subscale) was the only effective tool to identify anxiety, but it does incur a financial cost.

There has only been one previous systematic review of screening tools for depression after stroke and this focused on the detection of major depression only (Meader *et al.* 2014). One of our objectives in this study was to enable implementation of screening tools by identifying optimal cut-off scores for the tools we could recommend. However, this proved impossible because of the heterogeneity of participant characteristics, study designs and observed sensitivity and specificity estimates. Some of the heterogeneity was due to the varied choice of criterion measures against which the accuracy of the screening tools was judged. A semi-structured interview by a psychiatrist is widely accepted as the 'gold standard' reference criterion but use of different interview tools and diagnostic criteria can result in variances in diagnostic accuracy. The variability and accuracy of these tools will affect the reported accuracy of the screening tools against which they are measured. More consistent use of a criterion measure would enhance meta-analysis and comparison between tools.

A limitation of the psychiatric interview and diagnostic tools is that they are difficult for people with communication and cognitive difficulties to complete. Subsequently, many studies have used other screening tools as the criterion measure (Watkins *et al.* 2001*a*,*b*, 2007; Bennett *et al.* 2006; Lee *et al.* 2008; Hacker *et al.* 2010; de Man-van Ginkel *et al.* 2012*a*) or excluded

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

people with these problems (even those that tested observational tools for people unable to self-report), making the assumption that the results from more able populations would generalize to people with stroke. There is limited evidence to support or refute this assumption. Communication and cognitive problems are common after stroke and their exclusion limits the generalizability and relevance of the findings to clinical practice. Further research is needed involving pragmatic samples of people with stroke at all stages of recovery and survivorship, using predefined cut-off scores to establish the optimal thresholds so that people with both major and mild disorders can be identified.

Most of the selected tools were originally developed for a psychiatric population and then applied to stroke survivors, assuming that the experience of mood disorder after stroke has the same construct as other populations. Although stroke survivors display similar distributions of symptoms scores as people with other physical illnesses (House *et al.* 1991), mood disorders may be experienced differently in someone with additional physical or cognitive impairments (Gainotti *et al.* 1999). It is a matter of modern health policy that measurement tools should reflect the issues that are important and relevant to service users (so-called patient- reported outcome measures) (Darzi, 2008) and service users' views and expertise should be involved in all levels of research. Thus, it is notable that none of the selected screening tools included stroke survivors' perspectives in their construction and, although they have apparent face validity, the content validity for stroke survivors is unknown.

The inclusion of somatic items in the mood screening tools is particularly controversial because these can overlap with symptoms of stroke and/or effects of being in hospital. For example, many included items assess fatigue, concentration and memory problems or altered activity or sleeping patterns, which are common impairments after stroke independent of any emotional difficulties or hospitalization. Inclusion of such items would conflate the scores and might lead to ineffective clinical decision making or inaccurate research conclusions. To avoid the confounding effects of somatic items, some screening tools exclude them; however, the impact is unclear as there is some evidence that somatic symptoms are among the best differentiators between stroke survivors with and without depression (de Coster *et al.* 2005). It might be better to adjust the cut-off scores to reflect the increase in prevalence of these symptoms in the stroke population. Further research is required to investigate the construct of post-stroke depression and anxiety and to establish the content validity of the screening tools for stroke survivors.

A related issue is the construct validity of the screening tools for stroke survivors. Most selected tools were developed using classic test theory and are scored by summing the scores from the different items to produce a 'total' score. This is a controversial approach; many proponents of item response theory would consider this an inappropriate use of categorical data that could produce misleading scores (Tennant & Conaghan, 2007). Three of the selected tools have been subjected to Rasch analysis in relatively small samples: the BDI-II, GDS and HADS depression subscale. They report inconsistent findings regarding the unidimensionality of the tools and all identified redundant or disordered items, or ineffective scoring methods that did not fit the model (Pickard *et al.* 2006; Tang *et al.* 2007; Siegert *et al.* 2010). Only one of the studies provided data to enable the ordinal scoring data to be transformed into interval level data to allow use of parametric statistics and change scores to be calculated (Siegert *et al.* 2010), and it remains a moot point whether this makes an important difference to the data and how they are reported. Further development of the tools should include Rasch analysis to ensure an effective and efficient scale structure.

Most of the selected tools rely on the person's ability to self-report their mood, which is often compromised after stroke. In an attempt to overcome this, several tools use visual aids to facilitate non-verbal responses. However, these are based on the assumption that stroke survivors are able to interpret the visual aids. Work applying visual aids in tools to measure pain suggests that many stroke survivors, particularly those with right-hemisphere damage, find this difficult (Benaim *et al.* 2007). Acceptability and clinical utility of these tools need to be examined with people with stroke with different types and severity of impairment.

Finally, we identified a wide range of tools to detect depression in stroke survivors, but there are few standardized tools to detect anxiety and emotionalism. Further work is needed to develop person-centred tools for these purposes.

Limitations

A limitation of this study is that the quality is dependent on the articles identified. There is evidence of selection bias in the two studies that excluded participants who did not report low mood (Lincoln *et al.* 2003; Williams *et al.* 2005), thereby artificially increasing the prevalence of depression in the sample and probably affecting the reported PPV and NPV. We also included studies from around the world and note that the construct of depression could vary in different cultures, such as collectivist societies, which may have contributed to the heterogeneity of cut-off scores.

https://doi.org/10.1017/S0033291714000336 Published online by Cambridge University Press

Furthermore, we only included published English language studies and so may have missed relevant publications in other languages or unpublished data. To produce a generalizable result, we included studies that assessed stroke survivors at all stages of recovery, from the acute hospital setting to several years post-stroke. This may have contributed to our difficulties in ascertaining the optimum cut-off score for each tool, as sensitivity and specificity values at different cut-offs have been demonstrated to vary over time in the period following stroke (Berg *et al.* 2009). Finally, although we used recommended sensitivity and specificity criteria to reflect clinical priorities, alternative criteria may be warranted in different situations; for example, higher specificity may be required where resources for further assessment are limited. Further research should examine and compare the factor structure of depression globally and at different stages post-stroke to identify the most effective tools for each situation.

Clinical utility is rarely considered in tool development although it is key to uptake in clinical practice and research. We worked with clinicians in a range of stroke services in one of the largest conurbations in the UK to develop our measure of clinical utility, so we are confident that it is representative of the issues that limit implementation in the UK, at least. However, although the barriers to implementing tools in practice are fairly universal, the cut-off points may be context specific. Different models of health care may have different funding limits or time available for assessment. A further limitation is our reliance on reported administration times within the general population, as there were few reports within a stroke population. It is likely that screening tools would take longer to administer with patients with stroke-related impairments and activity limitations so the reported administration times may be underestimates. To address this, we contacted the authors or publishers of the selected tools for further information regarding clinical utility, but the information was not always available and should be considered a limitation.

Conclusions

The following tools can accurately screen for depression in stroke survivors in clinical practice: the GDS-15 can detect any depressive disorder and the PHQ-9 can detect severe depression whereas the SADQ-H can be used with stroke survivors who are unable to self-report. The HADS (both the total scale and the anxiety subscale) can effectively identify anxiety post-stroke but clinical utility is limited by the costs involved. We were unable to establish the optimal cut-off scores for these or the other selected tools.

Acknowledgements

This project was funded through a Knowledge Transfer Partnership (grant no. 0007812) by the Technology Strategy Board and the Greater Manchester and Cheshire Cardiovascular Network.

Declaration of Interest

None.

References

Aben I, Verhey F, Lousberg R, Lodder J, Honig A (2002). Validity of the Beck Depression Inventory, Hospital Anxiety and Depression Scale, SCL-90, and Hamilton Depression Rating Scale as screening instruments for depression in stroke patients. *Psychosomatics* **43**, 386–393.

Adshead F, Cody D, Pitt B (1992). BASDEC: a novel screening instrument for depression in elderly medical patients. *British Medical Journal* **305**, 397.

Agrell B, Dehlin O (1989). Comparison of six depression rating scales in geriatric stroke patients. *Stroke* **20**, 1190–1194.

Alexopoulos G, Abrams R, Young R, Shamoian C (1988). Cornell scale for depression in dementia. *Biological Psychiatry* **23**, 271–284.

APA (1987). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised. DSM-III-R*. American Psychiatric Association: Washington, DC.

APA (1994). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. DSM-IV*. American Psychiatric Association: Washington, DC.

APA (2000). *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision. DSM-IV-TR*. American Psychiatric Association: Washington, DC.

Ayerbe L, Ayis S, Crichton S, Wolfe C, Rudd A (2013). The natural history of depression up to 15 years after stroke: the South London Stroke Register. *Stroke* **44**, 1105–1110.

Beck A, Brown G, Steer R (1996). *Beck Depression Inventory-II (BDI-II)*. The Psychological Corporation: San Antonio, TX.

Beck A, Steer R, Brown G (2000). *BDI-FS Screen for Medical Inpatients: Manual*. The Psychological Corporation: San Antonio, TX.

Beck A, Ward C, Mendelson M, Mock J, Erbaugh J (1961). An inventory for measuring depression. *Archives of Genernal Psychiatry* **4**, 561–571.

Benaim C, Cailly B, Perennou D, Pelissier J (2004). Validation of the Aphasic Depression Rating Scale. *Stroke* **35**, 1692–1696.

Benaim C, Froger J, Cazottes C, Gueben D, Porte M, Desnuelle C, Pelissier J (2007). Use of the Faces Pain Scale by left and right hemispheric stroke patients. *Pain* **128**, 52–58.

Bennett HE, Lincoln NB (2006). Potential screening measures for depression and anxiety after stroke. *International Journal of Therapy and Rehabilitation* **13**, 401–406.

Bennett HE, Thomas SA, Austen R, Morris AM, Lincoln NB (2006). Validation of screening measures for assessing

mood in stroke patients. *British Journal of Clinical Psychology* **45**, 367–376.

Berg A, Lonnqvist J, Palomaki H, Kaste M (2009). Assessment of depression after stroke: a comparison of different screening instruments. *Stroke* **40**, 523–529.

Bowen A, Knapp P, Hoffmann A, Lowe D (2005). Psychological services for people with stroke: compliance with the UK National Clinical Guidelines. *Clinical Rehabilitation* **19**, 323–330.

Brumfitt S, Sheeran P (1999). *VASES: Visual Analogue Self-Esteem Scale*. Winslow Press: Oxon, UK.

Burton L, Tyson S, McGovern A (2013). Staff perceptions of using outcome measures in stroke rehabilitation. *Disability and Rehabilitation* **35**, 828–834.

Connell L, Tyson S (2012). Clinical reality of measuring upper-limb ability in neurologic conditions: a systematic review. *Archives of Physical Medicine and Rehabilitation* **93**, 221–228.

Copeland JR, Dewey ME, Wood N, Searle R, Davidson IA, McWilliam C (1987). The range of mental illness amongst the elderly in the community. Prevalence in Liverpool using the GMS-AGECAT package. *British Journal of Psychiatry* **150**, 815–823.

Crabtree A, Brookes R, Moynihan B (2012). Evaluating the use of a self-administered screening tool for depression after stroke. *International Journal of Stroke* **7**, 1747–4930.

Darzi A (2008). *High Quality Care For All*. Department of Health: London.

de Coster L, Leentjens AF, Lodder J, Verhey FR (2005). The sensitivity of somatic symptoms in post-stroke depression: a discriminant analytic approach. *International Journal of Geriatric Psychiatry* **20**, 358–362.

de Man-van Ginkel J, Gooskens F, Schepers V, Schuurmans M, Lindeman E, Hafsteinsdóttir T (2012*a*). Screening for post-stroke depression using the Patient Health Questionnaire. *Nursing Research* **61**, 333–341.

de Man-van Ginkel J, Hafsteinsdóttir T, Lindeman E, Burger H, Grobbee D, Schuurmans M (2012*b*). An efficient way to detect poststroke depression by subsequent administration of a 9-item and 2-item Patient Health Questionnaire. *Stroke* **43**, 854–856.

Derogatis L, Lipman R, Covi L (1973). SCL-90: an outpatient psychiatric rating scale – preliminary report. *Psychopharmacology Bulletin* **9**, 13–28.

Donnellan C, Hickey A, Hevey D, O'Neill D (2010). Effect of mood symptoms on recovery one year post-stroke. *International Journal of Geriatric Psychiatry* **25**, 1288–1295.

Endicott J, Spitzer RL (1978). A diagnostic interview: the schedule for affective disorders and schizophrenia. *Archives of General Psychiatry* **35**, 837–844.

First MB, Spitzer RL, Gibbon M, Williams JBW (1995). *Structured Clinical Interview for DSM-IV Axis I Disorders – Patient Edition (SCID-I/P), Version 2.0*. Biometrics Research Department, New York State Psychiatric Institute: New York.

Gainotti G, Azzoni A, Marra C (1999). Frequency, phenomenology and anatomical-clinical correlates of major post-stroke depression. *British Journal of Psychiatry* **175**, 163–167.

Gainotti G, Azzoni A, Razzano C, Lanzillotta M, Marra C, Gasparini F (1997). The Post-Stroke Depression Rating Scale: a test specifically devised to investigate affective disorders of stroke patients. *Journal of Clinical and Experimental Neuropsychology* **19**, 340–356.

Goldberg D, Williams P (1988). *A User's Guide to the General Health Questionnaire*. NFER-Nelson: Basingstoke.

Hacker VL, Stark D, Thomas S (2010). Validation of the Stroke Aphasic Depression Questionnaire using the brief assessment schedule depression cards in an acute stroke sample. *British Journal of Clinical Psychology* **49**, 123–127.

Hackett M, Yapa C, Parag V, Anderson C (2005). Frequency of depression after stroke: a systematic review of observational studies. *Stroke* **36**, 1330–1340.

Hamilton M (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry* **23**, 56–62.

Hammond M, O'Keefe S, Barer D (2000). Development and validation of a brief observer-rated scale screening scale for depression in elderly medical patients. *Age and Ageing* **29**, 511–551.

Hart S, Morris R (2007). Screening for depression after stroke: an exploration of professionals' compliance with guidelines. *Clinical Rehabilitation* **22**, 60–70.

Healey AK, Kneebone II, Carroll M, Anderson SJ (2008). A preliminary investigation of the reliability and validity of the Brief Assessment Schedule Depression Cards and the Beck Depression Inventory-Fast Screen to screen for depression in older stroke survivors. *International Journal of Geriatric Psychiatry* **23**, 531–537.

Holland J, Andersen B, Breitbart W, Compas B, Dudley M, Fleishman S (2011). *NCCN Clinical Practice Guidelines in Oncology: Distress Management*. Version 1.2011.

House A, Dennis M, Mogridge L, Warlow C, Hawton K, Jones L (1991). Mood disorders in the year after first stroke. *British Journal of Psychiatry* **158**, 83–92.

House A, Knapp P, Bamford J, Vail A (2001). Mortality at 12 and 24 months after stroke may be associated with depressive symptoms at 1 month. *Stroke* **32**, 696–701.

Johnson G, Burvill P, Anderson C, Jamrozik K, Stewart-Wynne E, Chakera T (1995). Screening instruments for depression and anxiety following stroke: experience in the Perth community stroke study. *Acta Psychiatrica Scandinavica* **91**, 252–257.

Jorge R, Robinson R, Arndt S, Starkstein S (2003). Mortality and poststroke depression: a placebo-controlled trial of antidepressants. *American Journal of Psychiatry* **160**, 1823–1829.

Kang H-J, Stewart R, Kim J-M, Jang J-E, Kim S-Y, Bae K-Y, Kim S-W, Shin I-S, Park M-S, Cho K-H, Yoon J-S (2012). Comparative validity of depression assessment scales for screening poststroke depression. *Journal of Affective Disorders* **147**, 186–191.

Kessler R, Barker P, Colpe L, Epstein J, Gfroerer J, Hiripi E, Howes M, Normand S, Manderscheid R, Walters E, Zaslavsky A (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry* **60**, 184–189.

Kotila M, Numminen H, Waltimo O, Kaste M (1999). Post-stroke depression and functional recovery in a

population-based stroke register. The Finnstroke study. *European Journal of Neurology* **6**, 309–312.

Kroenke K, Spitzer R, Williams J (2003). The Patient Health Questionnaire-2: validity of a two-item depression screener. *Medical Care* **41**, 1284–1292.

Lachs M, Feinstein A, Cooney L, Drickamer M, Marottoli R, Pannill F, Tinetti M (1990). A simple procedure for general screening for functional disability in elderly patients. *Annals of Internal Medicine* **112**, 699–706.

Lee A, Tang S, Yu G, Cheung R (2008). The smiley as a simple screening tool for depression after stroke: a preliminary study. *International Journal of Nursing Studies* **45**, 1081–1089.

Lightbody C, Auton M, Baldwin R (2007). The use of nurses' and carers' observations in the identification of poststroke depression. *Journal of Advanced Nursing* **60**, 595–604.

Lincoln NB, Nicholl CR, Flannaghan T, Leonard M, Van der Gucht E (2003). The validity of questionnaire measures for assessing depression after stroke. *Clinical Rehabilitation* **17**, 840–847.

Lincoln NB, Sutcliffe LM, Unsworth G (2000). Validation of the Stroke Aphasic Depression Questionnaire (SADQ) for use with patients in hospital. *Clinical Neuropsychological Assessment* **1**, 88–96.

Mahoney J, Drinka T, Abler R, Gunter-Hunt G, Matthews C, Gravenstein S, Carnes M (1994). Screening for depression: single question versus GDS. *Journal of the American Geriatrics Society* **42**, 1006–1008.

Meader N, Moe-Byrne T, Llewellyn A, Mitchell A (2014). Screening for poststroke major depression: a meta-analysis of diagnostic validity studies. *Journal of Neurology, Neurosurgery and Psychiatry* **85**, 198–206.

Mitchell P, Veith R, Becker K, Cain K, Fruin M, Tirschwell D, Teri L (2009). Brief psychosocial-behavioral intervention with antidepressant reduces poststroke depression significantly more than usual care with antidepressant: living well with stroke: randomized, controlled trial. *Stroke* **40**, 3073–3078.

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

National Audit Office (2010). *Progress in Improving Stroke Care*. The Stationery Office: London.

National Stroke Foundation (2010). *Clinical Guidelines for Stroke Management*. National Stroke Foundation: Melbourne, Australia.

NHS Improvement (2011). *Psychological Care After Stroke: Improving Stroke Services for People with Cognitive and Mood Disorders*. NHS Improvement: London.

NICE (2010). *Stroke Quality Standard* (*QS2*). National Institute for Health and Clinical Excellence: London.

NSSA (2008). *National Sentinel Stroke Audit, Phase 1 Organisational Audit 2008: Report for England, Wales and Northern Ireland*. Intercollegiate Stroke Working Party/Royal College of Physicians: London.

NSSA (2011). *National Sentinel Stroke Audit 2010 Round 7*. Intercollegiate Stroke Working Party/Royal College of Physicians: London.

O'Rourke S, MacHale S, Signorini D (1998). Detecting psychiatric morbidity after stroke: comparison of the GHQ and the HAD scale. *Stroke* **29**, 980–985.

Parikh R, Eden D, Price TR, Robinson RG (1988). The sensitivity and specificity of the Center for Epidemiologic Studies Depression Scale in screening for post-stroke depression. *International Journal of Psychiatry in Medicine* **18**, 169–181.

Pickard A, Dalal M, Bushnell D (2006). A comparison of depressive symptoms in stroke and primary care: applying Rasch models to evaluate the Center for Epidemiologic Studies-Depression Scale. *Value in Health* **9**, 59–64.

Pohjasvaara T, Vataja A, Leppävuori A, Kaste M, Erkinjuntti T (2001). Depression is an independent predictor of poor long-term functional outcome post-stroke. *European Journal of Neurology* **8**, 315–319.

Radloff LS (1977). The CES-D scale: a self-report depression scale for research in the general population. *Applied Psychological Measurement* **1**, 385–401.

Robins L, Wing J, Wittchen H, Helzer J, Babor T, Burke J, Farmer A, Jablenski A, Pickens R, Regier D, Sartorius N, Towle L (1988). The Composite International Diagnostic Interview. An epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Archives of General Psychiatry* **45**, 1069–1077.

Roger P, Johnson-Greene D (2009). Comparison of assessment measures for post-stroke depression. *Clinical Neuropsychologist* **23**, 780–793.

Rybarczyk B, Winemiller D, Lazarus L, Haut A, Hartman C (1996). Validation of a depression screening measure for stroke inpatients. *American Journal of Geriatric Psychiatry* **4**, 131–139.

Sagen U, Gunnar Vik T, Moum T, Morland T, Finset A, Dammen T (2009). Screening for anxiety and depression after stroke: comparison of the Hospital Anxiety and Depression Scale and the Montgomery and Asberg Depression Rating Scale. *Journal of Psychosomatic Research* **67**, 325–332.

Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, Hergueta T, Baker R, Dunbar GC (1998). The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *Journal of Clinical Psychiatry* **59** (Suppl. 20), 22–33.

Sheikh J, Yesavage J (1986). Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clinical Gerontologist* **5**, 165–173.

Shinar D, Gross CR, Price TR, Banko M, Bolduc PL, Robinson RG (1986). Screening for depression in stroke patients: the reliability and validity of the Center for Epidemiologic Studies Depression Scale. *Stroke* **17**, 241–246.

Siegert R, Tennant A, Turner-Stokes L (2010). Rasch analysis of the Beck Depression Inventory-II in a neurological rehabilitation sample. *Disability and Rehabilitation* **32**, 8–17.

Sivrioglu E, Sivrioglu K, Ertan T, Ertan F, Cankurtaran E, Aki O, Uluduz D, Ince B, Kirli S (2009). Reliability and validity of the Geriatric Depression Scale in detection of

poststroke minor depression. *Journal of Clinical and Experimental Neuropsychology* **31**, 999–1006.

Snaith R, Ahmed S, Mehta S, Hamilton M (1971). Assessment of severity of primary depressive illness: Wakefield self-assessment depression inventory. *Psychological Medicine* **1**, 143–149.

Spitzer R, Kroenke K, Williams J (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Journal of the American Medical Association* **282**, 1737–1744.

Stern R (1997). *Visual Analog Mood Scales Professional Manual*. Psychological Assessment Resources: Odessa, FL.

Sutcliffe L, Lincoln N (1998). The assessment of depression in aphasic stroke patients: the development of the Stroke Aphasic Depression Questionnaire. *Clinical Rehabilitation* **12**, 506–513.

Tang W, Chan S, Chiu H, Wong K, Kwok T, Mok V, Ungvari G (2004*a*). Can the Geriatric Depression Scale detect poststroke depression in the Chinese elderly? *Journal of Affective Disorders* **81**, 153–156.

Tang W, Ungvari G, Chiu H, Sze K (2004*b*). Detecting depression in Chinese stroke patients: a pilot study comparing four screening instruments. *International Journal of Psychiatry in Medicine* **34**, 155–163.

Tang WK, Ungvari GS, Chiu HF, Sze KH, Yu AC, Leung TL (2004*c*). Screening post-stroke depression in Chinese older adults using the Hospital Anxiety and Depression Scale. *Aging and Mental Health* **8**, 397–399.

Tang W, Wong E, Chiu H, Ungvari G (2007). Rasch analysis of the scoring scheme of the HADS Depression subscale in Chinese stroke patients. *Psychiatry Research* **150**, 97–103.

Tennant A, Conaghan PG (2007). The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Care and Research* **57**, 1358–1362.

Tham W, Kwan Y, Ang A (2012). Performance of the Patient Health Questionnaire-2 (PHQ-2) as a screening tool for depression in stroke patients. *Biological Psychiatry* **71**, 1415.

Turner A, Hambridge J, White J, Carter G, Clover K, Nelson L, Hackett M (2012). Depression screening in stroke: a comparison of alternative measures with the structured diagnostic interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (major depressive episode) as criterion standard. *Stroke* **43**, 1000–1006.

Turner-Stokes L, Kalmus M, Hirani D, Clegg F (2005). The Depression Intensity Scale Circles (DISCs): a first evaluation of a simple assessment tool for depression in the context of brain injury. *Journal of Neurology, Neurosurgery and Psychiatry* **76**, 1273–1278.

Watkins C, Daniels L, Jack C, Dickinson H, van den Broek M (2001*a*). Accuracy of a single question in screening for depression in a cohort of patients after stroke: comparative study. *British Medical Journal* **323**, 1159.

Watkins C, Leathley M, Daniels L, Dickinson H, Lightbody C, van den Broek M, Jack C (2001*b*). The Signs of Depression Scale in stroke: how useful are nurses' observations? *Clinical Rehabilitation* **15**, 447–457.

Watkins C, Lightbody C, Sutton C, Holcroft L, Jack C, Dickinson H, van den Broek M, Leathley M (2007). Evaluation of a single-item screening tool for depression after stroke: a cohort study. *Clinical Rehabilitation* **21**, 846–852.

WHO (1992). *Schedules for Clinical Assessment in Neuropsychiatry* (*SCAN*). World Health Organization: Geneva.

Williams L, Brizendine E, Plue L, Bakas T, Tu W, Hendrie H, Kroenke K (2005). Performance of the PHQ-9 as a screening tool for depression after stroke. *Stroke* **36**, 635–638.

Williams L, Ghose S, Swindle R (2004). Depression and other mental health diagnoses increase mortality risk after stroke. *American Journal of Psychiatry* **161**, 1090–1095.

Wing J, Cooper J, Sartorius N (1974). *Measurement and Classification of Psychiatric Symptoms*. Cambridge University Press: Cambridge, UK.

Yesavage J, Brink T, Rose T, Lum O, Huang V, Adey M, Leirea O (1983). Development and validation of a geriatric depression screening scale: a preliminary report. *Journal of Psychiatric Research* **17**, 37–49.

Zigmond AS, Snaith RP (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica* **67**, 361–370.

Zung W (1965). A self-rating depression scale. *Archives of General Psychiatry* **12**, 63–70.