Data Extraction Table					
Author/Year	Publication	Level of Evidence	Participants/Timing of Screening	Inclusion/Exclusion Criteria For 5 Primary Studies	Primary Findings
Burton & Tyson, 2015[26] (studies included in systemic review*)	Systemic review-30 papers and 27 depression screening tools with 8 tools also screening for anxiety.	A	(n=607 in 4 studies evaluating PHQ-9)	Inc: Articles from 5 databases evaluating 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. Exc: Studies of tools not designed as a screening tool and intended to make a full assessment of mood or diagnosis, studies of tools assessing constructs such as quality of life, involved validation of language translation of a tool, abstracts or conference papers and where < ½ participants had suffered a stroke or data from stroke survivors could not be extracted.	PHQ-9 detected major depression but was less sensitive in detecting milder symptoms. PHQ-9 CP scores in 3 studies ranged from 6-10; specificity ranged from 69-100% and specificity ranged from 78-89%.
de Man-Van Ginkel et al 2012a[29] *	Observational longitudinal study design.	B	(n=55 non-aphasic stroke patients < 2 weeks post stroke	Inc: Ischemic or hemorrhagic stroke patients able to communicate Exc: Positive for cognitive disorder per MMSE, too ill or communication disorders per clinical judgement of researcher or nurse caring for pt. For inpatients aphasia determined w/Frenchay Aphasia Screening Test, subjects discharged ≤ 14 days post stroke	PHQ-9 interrater reliability and internal consistency was strong with ICC=.90 and overall Cronbach's α=.79. Pearson's correlation 0.7 (p<.001) demonstrating moderate concurrent validity. Optimum CP score ≥ 10 when screening for MDD w/ sensitivity of 100%, specificity of 86%. At this CP, PPV = 50% NPV of 100%.
de Man-Van Ginkel et al 2012b[30] *♦	Prospective study design. (All patients screened with PHQ-2. Only patients who scored ≥ 2 on PHQ-2 were screened with PHQ-9	B	n=164 non-aphasic stroke patients screened 6-8 weeks post stroke	Inc: Diagnosis of stroke Exc: w/o serious cognitive disorders per MMSE score ≥189 or with communicative disorders per Frenchay Aphasia Screening Test with scores of ≥17 for pts< 60 yrs. ≥16 pts ≥61 yrs. and ≤70 yrs. and ≥15pts ≤71 yrs.	PHQ-9 alone CP ≥10 with sensitivity of 0.80% and specificity of 78%. PPV=34%, NPV=97%. PLR=3.71, NLR.24. Administering the PHQ-9 only to patients scoring ≥ 2 on the PHQ-2 improved PHQ-9's sensitivity to 87% and improved identification of depression, specificity=20%, PPV=32%, NPV=78%. PLR=1.08. NLR=.67.
Meader et al 2014[24] (studies included in meta-analysis ♦)	Meta-analysis of 24 studies,) using 1/3 depression screening tools: PHQ-9, CESD, HDRS (n=2907)	A	(n=552 stroke patients in 3 studies screening with PHQ-9)	Inc: Articles from 7 databases determining validity of case identification instruments for depression Exc: Studies not explicitly stating comparison to DSM or w/o ICD diagnosis of depression or without w/o sufficient data to be extracted.	PHQ-9 sensitivity=86, specificity=79%. PPV=67%. NPV = 92%. PHQ-9 (CUI + 0.58) indicated high utility for case finding in clinical practice (ruling in a diagnosis with minimal false positives)
Moriarty et al 2015[25] (studies included in meta-analysis ♥)	Meta-analysis of 36 studies with total n=21,292 patients.	A	Two studies using PHQ-9 in stroke population (n=388 stroke patients).	Inc: Articles from 3 databases and reverse citation search for any design evaluating PHQ-9 vs DSM or ICD criteria for MDD. Studies had to report enough data to calculate 2 x 2 contingency table, adult populations, depression diagnosed by SCID, MINI. Exc: Teens/adolescents or no diagnosis of MDD	PHQ-9 pooled sensitivity for CP ≥ 10 =78% and pooled specificity was 87%. DOR=26.27. At this CP, PHQ-9 better suited to screen for MDD in primary care vs secondary care settings
Prisnie et al 2016[27]	Validation study of 4 depression screening tools (PHQ-2, HADS, GDS-15, PHQ-9) vs DSM-IV SCID	B	(n=72 stroke patients) and (n=49 TIA patients). Time of screening unknown	Inc: Diagnosis stroke or TIA per specialized neurologist using AHA diagnostic criteria Exc: < 18 yrs., developmentally delayed, not fluent in English, hearing impaired, prior dx of Dementia	Of 4 tools being evaluated, PHQ-9 CP ≥ 13, achieved best balance of sensitivity 81.8% and 97.1% specificity with AUC= 89.5%. ACC of 95.6%, PPV=75%, NPV=98%. Also, had the smallest NLR at this CP (NLR=0.187)
Turner et al 2014[28] *♦♥	Validation study of 6 depression screening tools: (PHQ-2, PHQ-9, HADS-D, BDI-II, Distress Thermometer, K-10) vs DSM-IV SCID	B	(n=72 stroke patients ≥ 3 weeks post stroke	Inc: Aged ≥18, ≥ 3 weeks post confirmed stroke, able to attend health site for screening. Subjects w/dysarthria or expressive dysphasia but with adequate receptive and expressive communication strategies were included. Exc: Subjects who were not able to read or understand English or those with severe cognitive or physical impairment	PHQ-9 AUC=0.82. Cronbach's α=0.82. PHQ-9's best sensitivity was 85% at cut off point ≥ 6 vs ≥ 8 or ≥9 and best specificity was 78% at CP ≥ 9 vs ≥ 6 ≥ 8. Results support previous literature PHQ-9 appropriate to screen for depression in non-aphasic stroke patients.
Williams et al 2005[31] *♦♥	Randomized treatment trial for PSD. Participants screened with PHQ-9 and those with > or more symptoms of depression administered DSM-IV SCID. ROC analysis to determine sensitivity and specificity of PHQ-9.	B	n=316 stroke patients, some with mild aphasia screened for depression 4 and 8 weeks post stroke	Inc: Subjects not depressed matched 1:1 at site of enrollment to depressed subjects. Exc: Subjects w/more than moderate aphasia per NIHSS Scale language item score >1 or cognitive impairment per modified 6-item MMS>3	PHQ-9 with CP ≥ 10 had 91% sensitivity and 89% specificity for MDD and 78% sensitivity and 96% specificity for any depression diagnosis. PHQ-9 able to discriminate between depressed and non-depressed patients regardless of ethnicity, gender, age. AUC=96%

PHQ-9=Patient Health Questionnaire-9 item, PHQ-2-Patient Health Questionnaire-2 item, HADS-D=Hospital Anxiety and Depression Scale, DBI=Beck Depression Inventory II, K10=Kessler 10, CESD=Center for Epidemiological Studies Depression Scale, HDRS-Hamilton Depression Rating Scale, MMSE=Mini-Mental State Exam, CP=cut point, PPV=positive predictive value, NPV=negative predictive value, AUC=area under the curve, LR-=negative likelihood ratio, PLR=positive likelihood ratio, NLR=negative likelihood ratio, ACC=overall accuracy (% correctly classified), CUI=Clinical Utility Index, ROC= receiver operating characteristic curve, TIA=transient ischemic attack, CI=confidence interval, DOR=diagnostic odds ratio, MDD-major depressive disorder, DSM SCID=structured clinical interview for DSM-IV