

Reviewer ID: Amy Kim, Kyle Diab, Risa Fox			
Type of Outcome Measure: Patient Health Questionnaire-9 (PHQ-9)			Total articles: 7
Author ID Year	Study Design	Setting	Population (sample size, age) and Group
Bombardier et al. 2004	Cross-sectional survey	Not specified	N=849 (645M, 204F) Age >17yrs. 1 year post-SCI Mean age at the time of injury $\pm$ standard deviation was 36.9 $\pm$ 15.0 years  Recruited from 16 Model Spinal Cord Injury Systems throughout the USA. Patients were injured between Aug 30, 2000 and Apr 1, 2003. 47.6% AIS A complete  45.5% paraplegia
Bombardier et al. 2012	Blinded comparison of the PHQ-9 with the major depression module of the Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (SCID)	Inpatient rehabilitation units at the University of Washington Medical Center, Seattle, Washington; Harborview Medical Center, Seattle, Washington; the Texas Institute for Rehabilitation and Research, Houston, Texas; and the University of Michigan Health System, Ann Arbor, Michigan	N=142 M=111, F=31 Mean Age = 42.2 $\pm$ 16.6y (18-88y)  Traumatic SCI patients recruited between February 2008 and December 2010  Cervical = 95 Thoracic = 32 Lumbar = 11 Sacral = 4
Graves & Bombardier 2008	Retrospective analysis	National Spinal Cord Injury Database (NSCID)	N=3652 (M=2863; F=789) Mean age at time of interview = 41.4 $\pm$ 13.44y (range: 18-90y) Mean age at time of injury = 31.8 $\pm$ 13.62y  Traumatic SCI patients who participated in the NSCID from Oct 2000 through April 2003.
Krause et al. 2009	Follow-up survey	Hospital in the Southeastern United States	727 SCI subjects mean age: 47.9 70.2% male 75.8% White 53.3% cervical injury Average number of years since injury = 18.2.  A total of 1,385 participants were enrolled in the original study in 1997–1998. Participants were then contacted in 2007–2008

			to participate in a follow-up survey. At that time, 306 were deceased, 34 could not be located, and 5 were eliminated. Responses were received by 727 participants, yielding an adjusted response rate of 69.5% percent.
Richardson and Richards 2008	Retrospective analysis	National Spinal Cord Injury Database (NSCID)	2570 participants  1 year postinjury: 682 subjects (535 M, 147F) mean age: 38.66±15.32 5 years postinjury: 517 subjects (402M, 115F) mean age: 40.26±14.53 15 years postinjury: 653 subjects (518M, 135F) mean age: 42.72±10.09 25 years postinjury: 718 subjects (558M, 160F) mean age: 49.49±8.60
Summaka et al. 2019	Cross-sectional study  (PHQ9-Arabic version)	3 Lebanese rehabilitation centers (participants from Beirut, Mount Lebanon, North/South, and Bekaa regions)	N=51 (51M) Mean age: 37.2 ± 12.6 Mean HRDS (Hamilton Depression Rating Scale) score: 10.82 ± 7.3 Type of injury: paraplegia (37), tetraplegia (14) Cause of injury: war and explosions (27), motor vehicle accident (7), falls (7), disease (7), other (3)
Williams et al. 2009	Methodological study. Factor analysis and Rasch rating scale analysis.	Recruit from outpatient clinics, the Midwest Regional SCI Care System and community advertisements.	N = 202 people with SCI 77% male Mean (SD) age = 42.6 (13.9) years All participants were at least 1 year after injury, with a range of 1 to 44 years; the sample was a median of 7 years after injury.

**1. RELIABILITY**

Author ID	Internal Consistency	Test-retest, Inter-rater, Intra-rater
Summak a et al. 2019	Good internal consistency $\alpha=0.711$ (number of items = 9)	Test-retest reliability assessed using ICC with 95% confidence interval. The results showed that ICC = 0.88 (0.711-0.955), $P<0.001$ , which reflects a strong reproducibility of the PHQ-9-A total scale.
Bombardier et al. 2004	Overall $\alpha=0.87$ .  Corrected item total correlations ranged from 0.72 (depressed mood) and 0.69 (feelings of failure) to 0.45 (psychomotor agitation/depression) and 0.48 (suicidal ideation).	No data available
Richardson and Richards 2008	Alpha coefficients revealed good internal consistency for the PHQ-9 scale and for the subscales across groups.  Alpha coefficients:	No data available

		Total 9-item scale	Affective subscale	Somatic subscale	
	1 year post-injury	.84	.81	.74	
	5 years post-injury	.87	.82	.78	
	15 years post-injury	.87	.84	.77	
	25 years post-injury	.83	.70	.70	
Krause et al. 2009	The internal consistency of the full scale, as measured by Cronbach's alpha = 0.89			No data available	

**2. VALIDITY**

Author ID	Validity
Summaka et al. 2019	<p>Significant correlation was found between the PHQ-9 and the Hamilton Depression Rating Scale                      Convergent validity: <math>r = 0.713</math> (<math>P &lt; 0.001</math>)                      Discriminatory validity: showed statistical difference between depressed SCI persons and non-depressed SCI subjects (<math>11.8 \pm 5.2</math> vs <math>5.8 \pm 4.5</math>; <math>P &lt; 0.001</math>)</p>
Bombardier et al. 2004	<p>Spearman correlations and chi-square tests to compare PHQ-9 values to those of quality of life, subjective health and difficulty in role functioning from other established measures (Short Form-36, Satisfaction With Life Scale).</p> <p>*Sample size indicated by subscript number after rho symbol (<math>\rho</math>).</p> <p>There were significant inverse correlations between higher depressive scores as determined by the PHQ-9 and SWLS (<math>\rho_{144} = -.51</math>; <math>P &lt; .001</math>) and subjective health (<math>\rho_{144} = .50</math>; <math>P &lt; .001</math>). There were significant positive correlations with greater difficulty in daily role functioning (<math>\rho_{638} = .62</math>; <math>P &lt; .001</math>).</p> <p><b>Calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio and negative likelihood ratio.</b>                      **Refer to paper for definitions of each.</p> <p>Sensitive indicators of probable Major Depressive Disorder (MDD) (&gt;80%): depressed mood (93.8%), disturbed sleep (89.5%), decreased energy (87.5%), anhedonia (84.4%) and feelings of failure (80.2%).</p> <p>Items with high specificity (&gt;90%): psychomotor changes (97.7%), difficulty concentrating (93.8%), feelings of failure (92.8%), appetite changes (92.2%) and depressed mood (90.9%).</p> <p>All symptoms had low PPV (40.8% to 67.9%), suggesting that a large proportion of those reporting a particular item will not have MDD.</p>

	<p>NPV was higher (92.5% to 99.1%; i.e. the probability of not having MDD was high with a negative response to an item).</p> <p>Likelihood ratios for a positive response were high (5:1 for sleep disturbance to 18:1 for psychomotor changes). Likelihood ratios for a negative test were lower (0.07:1 for depressed mood to 0.64:1 for psychomotor changes).</p>
<p>Bombardier et al. 2012</p>	<p>Significant correlation between the PHQ-9 total score and each of the compared measures with the same underlying construct: Higher PHQ-9 scores were positively correlated with poorer subjective health on the Medical Outcomes Study Short Form-1 (SF-1) (Spearman <math>\rho=0.37</math>; <math>P&lt;.001</math>) The PHQ-9 was inversely correlated with the Euro-QOL current health state thermometer (Spearman <math>\rho=-0.38</math>; <math>P&lt;.001</math>) Greater depression severity on the PHQ-9 was negatively correlated with overall quality of life since injury on the Life-1 (Spearman <math>\rho=-0.38</math>; <math>P&lt;.001</math>) The relationship between depression severity and difficulty with daily role functioning was also significant (Spearman <math>\rho=0.37</math>; <math>P&lt;.001</math>)</p> <p>The agreement between the PHQ-9 <math>\geq 11</math> and the SCID* was moderate, with <math>k</math> of 0.50. The area under the curve value of 0.92 was excellent, indicating that the PHQ-9 total score correctly discriminated between those with and without MDD by the SCID with a high degree of accuracy.</p> <p>*Structured Clinical Interview for the <i>Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition</i> (SCID MDD)</p> <p>The SCID MDD module was used as the criterion standard to diagnose major depression</p> <p>Based on the Youden Index, the diagnostic accuracy of the PHQ-9 was optimized at a cutoff of PHQ-9 <math>\geq 11</math>.</p> <p>At this cutoff: PHQ-9 identified 24.6% of the sample as having MDD. The PHQ-9 detected 100% of those with a diagnosis of MDD (sensitivity) and had a specificity of 84%.</p>
<p>Richards on and Richards 2008</p>	<p>Among persons 1 year postinjury, both affective and somatic subscores showed a significant inverse correlation with satisfaction with life (<math>\rho = -.463</math>, <math>P &lt; .001</math>, and <math>\rho = -.346</math>, <math>P &lt; .001</math>, respectively).</p> <p>Significant negative correlations were also found between SWLS scores and factor subscores at 5 years postinjury (<math>\rho = -.415</math>, <math>P &lt; .001</math> for the somatic subscore; <math>\rho = -.456</math>, <math>P &lt; .001</math> for the affective subscore) and at 15 years postinjury (<math>\rho = -.404</math>, <math>P &lt; .001</math>, for the affective subscore; <math>\rho = -.248</math>, <math>P &lt; .001</math>, for the somatic subscore),</p> <p>Regarding the 25 years postinjury group, the affective subscale also correlated significantly, and in a negative direction, with satisfaction with life (<math>\rho = -.368</math>, <math>P &lt; .001</math>). A significant negative relationship was also found with the somatic subscale for the 25 year postinjury group (<math>\rho = -.255</math>, <math>P &lt; .001</math>).</p>
<p>Graves &amp; Bombardier 2008</p>	<p>The relative efficiency will represent the proportion of information available in the shorter scales relative to the information available in the 9-item scale.</p> <p>2-item test = 0.46 3-item test = 0.67 9-item test = 1.05(for men), 0.88(for women)</p> <p>Positive Predictive Value for 3-item screening test with a total score cutoff of: 3 = 0.56 4 = 0.77</p> <p>The squared correlation coefficient between the total scores on the 3-item scale and the 9-item scale is 0.794, meaning that the 3-item score accounts for approximately 79% of the variance in the 9-item total score.</p>

	<p>For the 3-item screening test with a total score cutoff of 3:                  Specificity = 0.93                  Sensitivity = 0.87</p> <p>For the 3-item screening test with a total score cutoff of 4:                  Specificity = 0.95                  Sensitivity = 0.82</p>																								
Krause et al. 2009	<p><u>Spearman Rank correlations between PHQ-9 and:</u>                  Major depressive disorder: 0.530                  PHQ-9&gt;15: 0.505                  PHQ-9&gt;10: 0.692                  Older Adult Health and Mood Questionnaire (OAHMQ): 0.781                  Satisfaction with Life Scale (SWLS): -0.477                  (P&lt;.0001 for all the above)</p>																								
Williams et al. 2009	<p><b>Content Validity:</b> Rasch analysis suggests the PHQ-9 is a unidimensional measure of depression.</p>																								
<p><b>3. RESPONSIVENESS</b> – no data available</p>																									
<p><b>4. FLOOR/CEILING EFFECT</b></p>																									
<b>Author ID</b>	<b>Floor/ceiling effect</b>																								
Williams et al. 2009	Floor effects: 22% of participants reported no depressive symptoms.																								
<p><b>5. INTERPRETABILITY</b></p>																									
<b>Author ID</b>	<b>Interpretability</b>																								
Summak a et al 2019	Mean (SD) PHQ-9 score: 7.2 (5.2)																								
Krause et al. 2009	Mean (SD) PHQ-9 score: 5.57 (5.74)																								
Bombardier et al. 2004	<p>Mean PHQ-9 score: 5.48 (95% CI: 5.07-5.88)</p> <table border="1"> <thead> <tr> <th colspan="3">PHQ-9 SCI Norms:</th> </tr> <tr> <th>Diagnostic Category/Label</th> <th>Definition: (PHQ-9 =)</th> <th>Rate, n(%)</th> </tr> </thead> <tbody> <tr> <td>No depressive symptoms</td> <td>0</td> <td>199 (23.4)</td> </tr> <tr> <td>Minimal depressive symptoms</td> <td>1 to 4</td> <td>294 (34.6)</td> </tr> <tr> <td>Mild depressive symptoms</td> <td>5 to 9</td> <td>170 (20.0)</td> </tr> <tr> <td>Moderate depressive symptoms</td> <td>10 to 14</td> <td>101 (11.9)</td> </tr> <tr> <td>Moderate/severe depressive symptoms</td> <td>15 to 19</td> <td>48(5.7)</td> </tr> <tr> <td>Severe depressive symptoms</td> <td>20 to 27</td> <td>37 (4.4)</td> </tr> </tbody> </table>	PHQ-9 SCI Norms:			Diagnostic Category/Label	Definition: (PHQ-9 =)	Rate, n(%)	No depressive symptoms	0	199 (23.4)	Minimal depressive symptoms	1 to 4	294 (34.6)	Mild depressive symptoms	5 to 9	170 (20.0)	Moderate depressive symptoms	10 to 14	101 (11.9)	Moderate/severe depressive symptoms	15 to 19	48(5.7)	Severe depressive symptoms	20 to 27	37 (4.4)
PHQ-9 SCI Norms:																									
Diagnostic Category/Label	Definition: (PHQ-9 =)	Rate, n(%)																							
No depressive symptoms	0	199 (23.4)																							
Minimal depressive symptoms	1 to 4	294 (34.6)																							
Mild depressive symptoms	5 to 9	170 (20.0)																							
Moderate depressive symptoms	10 to 14	101 (11.9)																							
Moderate/severe depressive symptoms	15 to 19	48(5.7)																							
Severe depressive symptoms	20 to 27	37 (4.4)																							