**Reviewer ID:** Amy Kim, Kyle Diab

**Type of Outcome Measure:** Patient Health Questionnaire-9 (PHQ-9)  

<table>
<thead>
<tr>
<th>Author ID Year</th>
<th>Study Design</th>
<th>Setting</th>
<th>Population (sample size, age) and Group</th>
</tr>
</thead>
</table>
| 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 ± standard deviation was 36.9±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 ±16.6y (18-88y)  
Traumatic SCI patients recruited between February 2008 and December 2010  
Cervical = 95  
Thoracic = 32  
Lumbar = 11  
Sacral = 4 |
| 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 |
| 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±13.44y (range: 18-90y)  
Mean age at time of injury = 31.8±13.62y  
Traumatic SCI patients who participated in the NSCID from Oct |
### 1. RELIABILITY

<table>
<thead>
<tr>
<th>Author ID</th>
<th>Internal Consistency</th>
<th>Test-retest, Inter-rater, Intra-rater</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bombardier et al. 2004</td>
<td>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).</td>
<td>No data available</td>
</tr>
<tr>
<td>Richardson and Richards 2008</td>
<td>Alpha coefficients revealed good internal consistency for the PHQ-9 scale and for the subscales across groups. Alpha coefficients:</td>
<td>No data available</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total 9-item scale</th>
<th>Affective subscale</th>
<th>Somatic subscale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 year post-injury</td>
<td>.84</td>
<td>.81</td>
</tr>
<tr>
<td>5 years post-injury</td>
<td>.87</td>
<td>.82</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Author ID</th>
<th>Validity</th>
</tr>
</thead>
</table>
| Bombardier et al. 2004 | 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).

*Sample size indicated by subscript number after rho symbol (ρ).

There were significant inverse correlations between higher depressive scores as determined by the PHQ-9 and SWLS (ρ_{144}=-.51; P<.001) and subjective health (ρ_{144}=.50; P<.001). There were significant positive correlations with greater difficulty in daily role functioning (ρ_{38}=0.62; P<.001).

**Calculated sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio and negative likelihood ratio.**

**Refer to paper for definitions of each.**

Sensitivity indicators of probable Major Depressive Disorder (MDD) (>80%): depressed mood (93.8%), disturbed sleep (89.5%), decreased energy (87.5%), anhedonia (84.4%) and feelings of failure (80.2%).

Items with high specificity (>90%): psychomotor changes (97.7%), difficulty concentrating (93.8%), feelings of failure (92.8%), appetite changes (92.2%) and depressed mood (90.9%).

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. 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).

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).

Bombardier et al. 2012

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 ρ=0.37; P<.001).
The PHQ-9 was inversely correlated with the Euro-QOL current health state thermometer (Spearman ρ=-0.38; P<.001).
Greater depression severity on the PHQ-9 was negatively correlated with overall quality of life since injury on the Life-1 (Spearman ρ=-0.38; P<.001).
The relationship between depression severity and difficulty with daily role functioning was also significant (Spearman ρ=0.37; P<.001).

The agreement between the PHQ-9 ≥11 and the SCID* was moderate, with k 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 probable MDD.
without MDD by the SCID with a high degree of accuracy.

*Structured Clinical Interview for the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (SCID MDD)

The SCID MDD module was used as the criterion standard to diagnose major depression

Based on the Youden Index, the diagnostic accuracy of the PHQ-9 was optimized at a cutoff of PHQ-9 ≥11.

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%.

Richards on and Richards 2008

Among persons 1 year postinjury, both affective and somatic subscores showed a significant inverse correlation with satisfaction with life (ρ = -.463, P < .001, and ρ = -.346, P < .001, respectively).

Significant negative correlations were also found between SWLS scores and factor subscores at 5 years postinjury (ρ = -.415, P < .001 for the somatic subscore; ρ = -.456, P < .001 for the affective subscore) and at 15 years postinjury (ρ = -.404, P < .001, for the affective subscore; ρ = -.248, P < .001, for the somatic subscore).

Regarding the 25 years postinjury group, the affective subscale also correlated significantly, and in a negative direction, with satisfaction with life (ρ = -.368, P < .001). A significant negative relationship was also found with the somatic subscale for the 25 year postinjury group (ρ = -.255, P < .001).

Graves & Bombardier 2008

The relative efficiency will represent the proportion of information available in the shorter scales relative to the information available in the 9-item scale.

2-item test = 0.46
3-item test = 0.67
9-item test = 1.05 (for men), 0.88 (for women)

Positive Predictive Value for 3-item screening test with a total score cutoff of:
3 = 0.56
4 = 0.77

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.

For the 3-item screening test with a total score cutoff of 3:
Specificity = 0.93
Sensitivity = 0.87

For the 3-item screening test with a total score cutoff of 4:
Specificity = 0.95
Sensitivity = 0.82

Krause et al. 2009

Spearman Rank correlations between PHQ-9 and:
- Major depressive disorder: 0.530
- PHQ-9 > 15: 0.505
- PHQ-9 > 10: 0.692
- Older Adult Health and Mood Questionnaire (OAHMQ): 0.781
- Satisfaction with Life Scale (SWLS): -0.477
  (P < .0001 for all the above)

Williams et al. 2009

**Content Validity:** Rasch analysis suggests the PHQ-9 is a unidimensional measure of depression.

3. RESPONSIVENESS – no data available
### 4. FLOOR/CEILING EFFECT

<table>
<thead>
<tr>
<th>Author ID</th>
<th>Floor/ceiling effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Williams et al. 2009</td>
<td>Floor effects: 22% of participants reported no depressive symptoms.</td>
</tr>
</tbody>
</table>

### 5. INTERPRETABILITY

<table>
<thead>
<tr>
<th>Author ID</th>
<th>Interpretability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krause et al. 2009</td>
<td>Mean (SD) PHQ-9 score: 5.57 (5.74)</td>
</tr>
<tr>
<td>Bombardier et al. 2004</td>
<td>Mean PHQ-9 score: 5.48 (95% CI: 5.07-5.88)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>PHQ-9 SCI Norms:</strong></th>
<th><strong>Definition:</strong></th>
<th><strong>Rate, n(%)</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>No depressive symptoms</td>
<td>PHQ-9 ≥ 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>