Table C-10. Risk of bias of data reported in studies included for Key Question 1C

Study	Psychometric Property	Risk-of-Bias Considerations[a]	Risk-of-Bias Category	Comments
Finger et al. 2014[6]	Internal consistency reliability	Percentage of missing items given? Adequate sample size? Checked unidimensionality? Separate internal consistency reliability statistic for each unidimensional subscale? Any important flaws?	Low	Authors provided separate internal consistency reliability statistics for both item difficulty and person ability
Finger et al. 2014[6]	Separation reliability	Percentage of missing items given? Adequate sample size? Checked unidimensionality? Separate internal consistency reliability statistic for each unidimensional subscale? Any important flaws?	Low	Used a statistical measure for this psychometric property.
Finger et al. 2014[6]	Face validity	Did they assess whether all items are relevant to what they are trying to measure? Did they assess whether all items are relevant for the purpose of the instrument? Did they assess whether the items comprehensively reflect what they are trying to measure? Any important flaws?	Low	Unclear whether the list of key activities was comprehensive, but probably it was.
Finger et al. 2014[6]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Low	Used a statistical measure for this psychometric property.
Finger et al. 2014[7]	Internal consistency reliability	Percentage of missing items given? Adequate sample size? Checked unidimensionality? Separate internal consistency reliability statistic for each unidimensional subscale? Any important flaws?	Moderate	Authors provided separate internal consistency reliability statistics for both item difficulty and person ability. Only 40 patients tested, which may be too low.
Finger et al. 2014[7]	Separation reliability	Percentage of missing items given? Adequate sample size? Checked unidimensionality? Separate internal consistency reliability statistic for each unidimensional subscale? Any important flaws?	Moderate	Used a statistical measure for this psychometric property. Only 40 patients tested, which may be too low.
Finger et al. 2014[7]	Face validity	Did they assess whether all items are relevant to what they are trying to measure? Did they assess whether all items are relevant for the purpose of the instrument? Did they assess whether the items comprehensively reflect what they are trying to measure? Any important flaws?	Low	Unclear whether all items were truly important to the patients
Finger et al. 2014[7]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Moderate	Used a statistical measure for this psychometric property. Only 40 patients, which may be too low.

Table C-10. Risk of bias of data reported in studies included for Key Question 1C (continued)

Study	Psychometric Property	Risk-of-Bias Considerations[a]	Risk-of-Bias Category	Comments
Bittner et al. 2011[2]	Test-retest reliability	Percentage of missing items given? Adequate sample size? At least 2 measurements available? Were administrations independent? Was time interval stated? Were patients stable in the interim? Was time interval appropriate? Were test conditions similar for the 2 measurements? Any important flaws?	Moderate	Only 20 patients, and some did not have a retinal condition of interest.
Bittner et al. 2011[2]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Moderate	Used a statistical measure for this psychometric property. Only 20 patients, which may be too low.
McKnight and Babcock-Parziale 2007[8]	Responsiveness	Percentage of missing items given? Adequate sample size? Longitudinal design with at least 2 measurements? Time interval stated? Hypotheses about changes made a priori? Comparator instrument to determine true responsiveness? Any important flaws?	Moderate	Authors did not indicate whether the measured pre-post difference was sufficiently large for the measures to be considered responsive.
McKnight and Babcock-Parziale 2007[8]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Low	Used a statistical measure for this psychometric property.
Roman et al. 2007[9]	Test-retest reliability	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Low	Did not report whether the reported SD of 1.41 was for the 61 patients enrolled or the 36 patients for whom data were reported.
Roman et al. 2007[9]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Moderate	Unclear why sensitivity data were only reported for 36 of 61 enrolled patients.
Kiser et al. 2006[10]	Test-retest reliability	Percentage of missing items given? Adequate sample size? At least two measurements available? Were administrations independent? Was time interval stated? Were patients stable in the interim? Was time interval appropriate? Were test conditions similar for the two measurements? Any important flaws?	Low	4 had optic neuropathies, but this only represents 6% of the low-vision patients

Table C-10. Risk of bias of data reported in studies included for Key Question 1C (continued)

Study	Psychometric Property	Risk-of-Bias Considerations[a]	Risk-of-Bias Category	Comments
Kiser et al. 2006[10]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Moderate	Some patients could not complete 2 of the 3 tests; missing data was a problem. Unclear whether the correlation among the remaining patients is relevant.
Babcock-Parziale et al. 2005[11]	Internal consistency reliability	Percentage of missing items given? Adequate sample size? Checked unidimensionality? Separate internal consistency reliability statistic for each unidimensional subscale? Any important flaws?	Low	Authors provided separate internal consistency reliability statistics for both item difficulty and person ability.
Babcock-Parziale et al. 2005[11]	Separation reliability	Percentage of missing items given? Adequate sample size? Checked unidimensionality? Separate internal consistency reliability statistic for each unidimensional subscale? Any important flaws?	Low	Used a statistical measure for this psychometric property.
Babcock-Parziale et al. 2005[11]	Face validity for VA-13	Did they assess whether all items are relevant to what they are trying to measure? Did they assess whether all items are relevant for the purpose of the instrument? Did they assess whether the items comprehensively reflect what they are trying to measure? Any important flaws?	High	Pretest assessment relies on patients' potentially faulty memory of their abilities before treatment. Sparse assessment of whether this instrument applies well to patients with very low vision
Babcock-Parziale et al. 2005[11]	Face validity for Functional Low-Vision Observer Rated Assessment (FAST)	Did they assess whether all items are relevant to what they are trying to measure? Did they assess whether all items are relevant for the purpose of the instrument? Did they assess whether the items comprehensively reflect what they are trying to measure? Any important flaws?	Moderate	Sparse assessment of whether this instrument applies well to patients with very low vision
Babcock-Parziale et al. 2005[11]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Low	Used a statistical measure for this psychometric property.
Babcock-Parziale et al. 2005[11]	Responsiveness	Percentage of missing items given? Adequate sample size? Longitudinal design with at least 2 measurements? Time interval stated? Hypotheses about changes made a priori? Comparator instrument to determine true responsiveness? Any important flaws?	Moderate	No comparator instrument to determine true responsiveness.

Table C-10. Risk of bias of data reported in studies included for Key Question 1C (continued)

Study	Psychometric Property	Risk-of-Bias Considerations[a]	Risk-of-Bias Category	Comments
Kiser et al. 2005[4]	Test-retest reliability	Percentage of missing items given? Adequate sample size? At least 2 measurements available? Were administrations independent? Was time interval stated? Were patients stable in the interim? Was time interval appropriate? Were test conditions similar for the 2 measurements? Any important flaws?	Low	3 had optic neuropathies, but this represents only 5% of the patient with low vision
Stelmack et al. 2002[12]	Construct validity	Percentage of missing items given? Adequate sample size? Was an accepted statistical measure used, with standard thresholds for acceptability? If authors reported a comparator measure, would this comparator be expected to correlate with the tested measure? Any important flaws?	Low	Used a statistical measure for this psychometric property, and provided data separately for person ability and item difficulty.
Stelmack et al. 2002[12]	Responsiveness	Percentage of missing items given? Adequate sample size? Longitudinal design with at least 2 measurements? Time interval stated? Hypotheses about changes made a priori? Comparator instrument to determine true responsiveness? Any important flaws?	Moderate	No comparator instrument to determine true responsiveness.

VA-13=Veteran's Administration-13

Risk-of-bias considerations were based on the COSMIN manual.[5]