

## Appendix 2. Risk of Bias Tool

Modified Newcastle-Ottawa Scale from Wells, G. et al, The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp) 2013;

Selection	
<p>1. Definition of ACL injured population</p> <p>Low risk of bias = a High risk of bias = b</p>	<p>a) Clearly described if the inclusion/exclusion criteria of an ACL injured person stated <b>both</b> of the following criteria:</p> <ul style="list-style-type: none"> <li>i) Diagnosed ACL injury with clinical/imaging or surgical confirmation (e.g. Lachman's or pivot shift test ± MRI/arthroscopic confirmation),</li> <li>ii) Reports of surgical or non-surgical management</li> </ul> <p>b) Not described OR used minimal criteria for inclusion/exclusion.</p>
<p>2. Source population</p> <p>Low risk of bias = a High risk if = b, c</p>	<p>a) A consecutive sample or random selection from a source population that is well described and representative of the condition under study (e.g. surgeon's clinic, outpatient clinic).</p> <p>b) A consecutive sample or random selection from a population that is not highly representative of the condition under study.</p> <p>c) Cannot be defined or enumerated (i.e. volunteering or self-recruitment).</p>
<p>3. Typical of the average ACL injured population (representativeness of cohort)</p> <p>Low risk of bias = a High risk if = b</p>	<p>a) Truly representative of the average ACL injured person in the community if <b>all</b> of the following criteria are present:</p> <ul style="list-style-type: none"> <li>i) Including men and women,</li> <li>ii) Typical age range at time of ACL injury/surgery (mean age = 16-35),</li> <li>iii) If surgery, then 'typical' surgical procedure (arthroscopic and not synthetic graft*)</li> </ul> <p><i>*If non-surgical management then N/A for this point</i></p> <p>b) Above criteria are not present then not truly representative of the average ACL injured population.</p>
<p>4. Sample size</p> <p>Low risk of bias = a High risk if = b, c</p>	<p>a) Power analysis completed and sample size adequate to detect meaningful difference.</p> <p>b) Power analysis completed but sample size not adequate to detect meaningful difference.</p> <p>c) No power analysis completed.</p>

Exposure	
<p>5. Methods for assessment of functional performance (i.e. ascertainment of exposure)</p> <p>Low risk of bias = a High risk if = b, c</p>	<p>a) Well described methods for functional tests - including an appropriately trained or appropriate profession as assessor <b>AND</b> describes or cites reliability.</p> <p>b) Well described methods for functional tests including an appropriately trained or appropriate profession as assessor) <b>OR</b> describes or cites reliability.</p> <p>c) Not described.</p>
<p>6. Demonstration that outcome of interest was not present at ascertainment of exposure (i.e. outcome that is compared to exposure)</p> <p>Low risk = a High risk if = b</p>	<p>a) <b>True</b> if baseline score of outcome of interest for both exposed/non-exposed (poor/good functional performance) has been accounted for (for example as a covariate or change in score or not present at ascertainment of exposure).</p> <p>b) No demonstration that the baseline score of outcome of interest has been accounted for.</p>
Comparability	
<p>7. Comparability of cohorts on the basis of the design or analysis</p> <p>Low risk = a High risk if = b</p>	<p>a) Comparability exists if study cohort (exposed/non-exposed) was a priori matched for <b>at least one covariate</b>, or confounding controlled for in statistical analysis.</p> <p>Covariate examples:</p> <ul style="list-style-type: none"> <li>i) Age,</li> <li>ii) BMI,</li> <li>iii) Sex</li> </ul> <p>b) Study not controlled in design or analysis and no confounders acknowledged.</p> <p>(Statements of no differences between groups or that differences were not statistically significant are not sufficient for establishing comparability)</p>

Outcome	
<p>8. Validity and reliability of outcome(s) of interest</p> <p>Low risk = a High risk if = b</p>	<p>a) Outcome measure(s) of interest are clearly described, and references other article(s) which found outcome measure to be valid &amp; reliable <b>OR</b> demonstrates the outcome measure(s) of interest are valid and reliable. (note all outcome(s) of interest must be valid and reliable for (a))</p> <p>b) If outcome measure(s) of interest were not explained in reproducible detail, or validity and reliability not proven/reported.</p>
<p>9. Assessment of outcome(s) of interest</p> <p>Low risk = a High risk if = b</p>	<p>a) Assessor has suitable qualification to interpret findings (e.g. musculoskeletal radiologist) <b>AND</b> blind to participant baseline exposure/non-exposure. <i>*N/A: Blinding not needed for self-reported outcomes</i></p> <p>b) Poor or no description.</p>
<p>10. Adequacy of follow-up of cohorts</p> <p>Low risk = a High risk if = b</p>	<p>a) Adequacy of follow-up if <b>either</b> of the following are satisfied:</p> <ul style="list-style-type: none"> <li>i) &lt;15% lost to follow up + description of those lost,</li> <li>ii) &lt;5% lost to follow up with no description</li> </ul> <p>b) &gt;15% lost to follow up or not explicitly stated with number of participants lost to follow-up <b>OR</b> characteristics of those lost to follow-up were not described.</p>

\*risk of bias assessed from published paper, not considering extra data if provided by the authors