

## Appendix Table 1

### Quality Assessment

Paper ID:

Reviewer:

Study design:

Question	Response	Scoring
1. A clearly stated aim	Did they have a “study question” or “main aim” or “objective”? The question addressed should be precise and relevant in light of available literature. To be scored <i>adequate</i> the aim of the study should be coherent with the “Introduction” of the paper.	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
2. Inclusion of consecutive patients	Did the authors say: “consecutive patients” or “all patients during period from ... to...” or “all patients fulfilling the inclusion criteria”.	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
3. A description of inclusion and exclusion criteria	Did the authors report the inclusion and exclusion criteria?	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
4. Inclusion of patients	Did the authors report how many eligible patients agreed to participate (i.e. gave consent)?	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
5. Prospective collection of data. Data were collected according to a protocol established before the beginning of the study.	Did they say “prospective” or “follow-up”?  The study is NOT PROSPECTIVE when: <ul style="list-style-type: none"> <li>• chart review, or database review</li> <li>• “retrospective”</li> </ul>	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
6. Outcome measures	Did they report the OA outcome; clinical OA, osteotomy, total knee arthroplasty, unilateral knee arthroplasty, radiographic OA, OA findings on MRI, OA findings during arthroscopy?	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
7. Was the used OA classification shown to be valid and reliable?	To be scored as <i>adequate</i> , the following classifications or indications could be used: <ul style="list-style-type: none"> <li>• Clinical: ACR criteria, osteotomy, total knee arthroplasty, unilateral knee arthroplasty</li> <li>• Radiographic OA: Kellgren &amp; Lawrence, Fairbank, Ahlback, IKDC grading system, OARSI grading system.</li> <li>• MRI: use of description of definite osteophyte formation <u>and</u> cartilage loss</li> <li>• Arthroscopic: Outerbridge classification</li> <li>• Combination of above-mentioned classifications/ indications.</li> </ul>	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported

	<p>To be scored as <i>inadequate</i>:</p> <ul style="list-style-type: none"> <li>• Use of self-formulated classifications</li> <li>• Use of modified classifications</li> </ul>	
8. Unbiased assessment of the study outcome and determinants	<p>To be judged as <i>adequate</i> the following 2 aspects had to be positive:</p> <ul style="list-style-type: none"> <li>• Outcome and determinants had to be measured independently</li> <li>• Both for cases and controls the outcome and determinants had to be assessed in the same way</li> </ul>	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
9. Were the determinant measures used accurate (valid and reliable)?	<p>For studies where the determinant measures are shown to be valid and reliable, the question should be answered <i>adequate</i>. For studies which refer to other work that demonstrates the determinant measures are accurate, the question should be answered as <i>adequate</i>.</p> <p>For example: a meniscus rupture had to be scored during arthroscopy or on MRI; activity level had to be measured with a validated questionnaire.</p>	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
10. Follow-up period appropriate to the aim of the study	<p>Did they report the follow-up period? To be judged as <i>adequate</i>:</p> <ul style="list-style-type: none"> <li>• the follow-up should be sufficiently long to allow the assessment of the main outcome: for radiographic OA a minimum of 4 years and for OA findings on MRI or during arthroscopy a minimum of 2 years.</li> </ul>	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
11. Loss to follow-up	<p>To be judged as <i>adequate</i> the following 2 aspects had to be positive:</p> <ul style="list-style-type: none"> <li>• Did they report the losses to follow-up?</li> <li>• Was the loss to follow-up less than 20%</li> </ul>	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported
12. Adequate Statistical analyses	<p>To be judged as <i>adequate</i> the following 3 aspects had to be positive:</p> <ul style="list-style-type: none"> <li>• There must be a description of the relationship between the determinant and OA outcome or a description of the comparison (with information about the statistical significance)</li> <li>• Was there adjustment for the following confounders: <ul style="list-style-type: none"> <li>a. Age</li> <li>b. Gender</li> <li>c. BMI</li> </ul> <p>If the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses, the question should be answered <i>inadequate</i>.</p> </li> <li>• Did they show variance in the reported outcome (for example SD, CI)</li> </ul>	<input type="checkbox"/> 1. adequate <input type="checkbox"/> 0. inadequate <input type="checkbox"/> 0. not reported

Abbreviations: ACR: American College of Rheumatology; BMI: body mass index; CI: confidence interval; IKDC; International Knee Documentation Committee; MRI: magnetic resonance imaging; OA: osteoarthritis; SD: standard deviation