### Quality Assessment

#### Paper ID: 

#### Reviewer: 

#### Study design: 

<table>
<thead>
<tr>
<th>Question</th>
<th>Response</th>
<th>Scoring</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. A clearly stated aim</td>
<td>Did they have a &quot;study question&quot; or &quot;main aim&quot; or &quot;objective&quot;? The question addressed should be precise and relevant in light of available literature. To be scored adequate the aim of the study should be coherent with the &quot;Introduction&quot; of the paper.</td>
<td>☐ 1. adequate  ☐ 0. inadequate  ☐ 0. not reported</td>
</tr>
<tr>
<td>2. Inclusion of consecutive patients</td>
<td>Did the authors say: “consecutive patients” or “all patients during period from … to …” or “all patients fulfilling the inclusion criteria”.</td>
<td>☐ 1. adequate  ☐ 0. inadequate  ☐ 0. not reported</td>
</tr>
<tr>
<td>3. A description of inclusion and exclusion criteria</td>
<td>Did the authors report the inclusion and exclusion criteria?</td>
<td>☐ 1. adequate  ☐ 0. inadequate  ☐ 0. not reported</td>
</tr>
<tr>
<td>4. Inclusion of patients</td>
<td>Did the authors report how many eligible patients agreed to participate (i.e. gave consent)?</td>
<td>☐ 1. adequate  ☐ 0. inadequate  ☐ 0. not reported</td>
</tr>
<tr>
<td>5. Prospective collection of data. Data were collected according to a protocol established before the beginning of the study.</td>
<td>Did they say “prospective” or “follow-up”? The study is NOT PROSPECTIVE when: • chart review, or database review • &quot;retrospective&quot;</td>
<td>☐ 1. adequate  ☐ 0. inadequate  ☐ 0. not reported</td>
</tr>
<tr>
<td>6. Outcome measures</td>
<td>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?</td>
<td>☐ 1. adequate  ☐ 0. inadequate  ☐ 0. not reported</td>
</tr>
<tr>
<td>7. Was the used OA classification shown to be valid and reliable?</td>
<td>To be scored as adequate, the following classifications or indications could be used: • Clinical: ACR criteria, osteotomy, total knee arthroplasty, unilateral knee arthroplasty • Radiographic OA: Kelgren &amp; Lawrence, Fairbank, Ahlback, IKDC grading system, OARSI grading system. • MRI: use of description of definite osteophyte formation and cartilage loss • Arthroscopic: Outerbridge classification • Combination of above-mentioned classifications/ indications.</td>
<td>☐ 1. adequate  ☐ 0. inadequate  ☐ 0. not reported</td>
</tr>
</tbody>
</table>
To be scored as *inadequate*:
- Use of self-formulated classifications
- Use of modified classifications

| 8. Unbiased assessment of the study outcome and determinants | To be judged as *adequate* the following 2 aspects had to be positive:
- Outcome and determinants had to be measured independently
- Both for cases and controls the outcome and determinants had to be assessed in the same way | □ 1. adequate □ 0. inadequate □ 0. not reported |

| 9. Were the determinant measures used accurate (valid and reliable)? For studies where the determinant measures are shown to be valid and reliable, the question should be answered *adequate*. For studies which refer to other work that demonstrates the determinant measures are accurate, the question should be answered as *adequate*. For example: a meniscus rupture had to be scored during arthroscopy or on MRI; activity level had to be measured with a validated questionnaire. | | □ 1. adequate □ 0. inadequate □ 0. not reported |

| 10. Follow-up period appropriate to the aim of the study | Did they report the follow-up period? To be judged as *adequate*:
- 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. | □ 1. adequate □ 0. inadequate □ 0. not reported |

| 11. Loss to follow-up | To be judged as *adequate* the following 2 aspects had to be positive:
- Did they report the losses to follow-up?
- Was the loss to follow-up less than 20% | □ 1. adequate □ 0. inadequate □ 0. not reported |

| 12. Adequate Statistical analyses | To be judged as *adequate* the following 3 aspects had to be positive:
- 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)
- Was there adjustment for the following confounders:
  a. Age
  b. Gender
  c. BMI
  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 *inadequate*.
- Did they show variance in the reported outcome (for example SD, CI) | □ 1. adequate □ 0. inadequate □ 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