**Appendix E – Critical Appraisal Skills Program (CASP) Checklist for Quality Assessment of Systematic Reviews**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Bias Risk:**  **Green – low**  **Red – high** | **Systematic Review** | | | | | | |
| **Barteit et al.**6 | **Kovoor et al.**7 | **Laverdiere et al.**8 | **Mao et al.**9 | **Ong et al.**10 | **Polce et al.**11 | **Williams et al.**12 |
| 1. Did the review address a clearly focused question? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 2. Did the authors look for the right type of papers? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 3. Do you think all the important, relevant studies were included? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 4. Did the review’s authors do enough to assess quality of in the included studies? | Yes | Yes | No | Yes | Yes | Yes | No |
| 5. If the results of the review have been combined, was it reasonable to do so? | Yes | Yes | Yes | Yes | NA | Yes | Yes |
| 6. What are the overall results of the review? | -XR beneficial for medical education  -XR noninferior compared to conventional methods  -Greater enthusiasm and enjoyment learning with XR  -Many studies small-scale and short-term pilots | -Confirmed the validity and effectiveness of AR to supplement fundamental laparoscopic skills acquisition and complete surgical procedures  -For laparoscopic skills, achieved more realism, face validity, and construct validity  -For entire surgical procedures, it achieved greater content validity  -Effective in education/mentoring  -AR similar to more effective when compared to other XR | -Increasing interest in AR among orthopedic surgeons  -Current AR studies show similar or better outcomes with AR compared with traditional techniques  -Still many challenges to overcome for widespread use | Results in note below1 | Results in note below2 | -Most level one  studies demonstrate objective improvements with VR compared with traditional training  -Considerable heterogeneity in simulator constructs, training paradigms, and outcome measures | -Competency, surgical opinion, and postoperative complication rates were in favor of AR  -Operative duration appears to increase |
| 7. How precise are the results? | NA | NA | NA | iVR vs control (post intervention procedural time): pooled standardized mean difference for time to completion was -0.90 (95% CI = 1.33 to -047, I 2 = 1%, P < 0.0001) | NA | Results in note below3 | NA |
| 8. Can the results be applied to the local population? | Yes | Yes | Yes | Yes | No | Yes | Yes |
| 9. Were all important outcomes considered? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| 10. Are the benefits worth the harms and costs? | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| **Overall Risk of Bias** | **Low** | **Low** | **Low** | **Low** | **Low** | **Low** | **Low** |
| RCT/Comparative Study Included | Logishetty 235  Rai42 | Andersen25 | None | Blumstein26  Hooper27  Logishetty34  Lohre36  Lohre 237  Orland28,41  Xin 228 | Rai42 | Hooper27  Logishetty 235  Xin 228 | Al Janabi48  Andersen25  Logishetty 235 |
| 1iVR performed 18% to 43% faster on procedural time to completion compared to control; iVR demonstrated greater post-intervention scores on procedural checklists and greater implant placement accuracy compared to control; Increase in cognitive load (reaction time) in iVR vs VR; Positive feedback about usability of iVR; Residents enjoyed iVR training significantly more than didactic training (4.8/5 vs. 3.3/5 points, P < 0.001). Immersive VR training was estimated to be 34.1 times more cost-effective than traditional training methods  2Education: simulators demonstrated efficacy and validity in improving surgical performance and reducing complication rates. Ophthalmoscopy simulators demonstrated efficacy and validity. Evidence in improving ophthalmoscopy skills in the clinical setting.  Diagnostics: studies demonstrated proof-of-concept in presenting ocular imaging data on extended reality platforms and validity in assessing the function of patients with ophthalmic diseases.  Therapeutics: heads-up surgical systems had similar complication rates, procedural success rates, and outcomes in comparison with conventional ophthalmic surgery.  3Time-to-task completion (VR vs control) (n=6):  -VR training favored, avg mean difference of 282.25 seconds (95% CI, 2133.64 to 230.87, P = 0.002)  -Sub-analysis stratified by length of training interval: avg mean difference of 245.24 seconds (95% CI, 294.76 to 4.28, P = 0.07) and 2137.74 seconds (95% CI, 2191.39 to 84.09, P < 0.001) for the short-term and immediate posttesting subgroups, respectively.  Objective performance scores (n=3): -avg mean difference of 1.24 points (95% CI, 0.18 to 2.30) indicating VR training significantly favored (P = 0.02) | | | | | | | |