Radiology AI Evidence Scorecard

Rate each question Strong, Partial or Weak. Don't add them up: the pattern matters more than a total.

Vendor: _____ **Product:** _____

Study reviewed: _____ **Date:** _____

#	Question	Strong	Partial	Weak
1	**Study design** **Strong:** randomized, or prospective in routine practice. **Red flag:** retrospective or simulation results sold as real-world performance. Notes:	☐	☐	☐
2	**Real-world population** **Strong:** consecutive patients at realistic prevalence. **Red flag:** enriched case sets with no real-world follow-up. Notes:	☐	☐	☐
3	**Independent data** **Strong:** external testing across several sites. **Red flag:** results only from the sites that built the model. Notes:	☐	☐	☐
4	**Fit to your setting** **Strong:** several scanner vendors, exclusions explained. **Red flag:** one scanner vendor that isn't yours. Notes:	☐	☐	☐
5	**Outcomes that matter** **Strong:** detection, recall and interval cancers with follow-up. **Red flag:** AUC only. Notes:	☐	☐	☐
6	**Human and AI workflow** **Strong:** how and when readers saw the AI, and who decided. **Red flag:** “radiologists used the AI”, nothing more. Notes:	☐	☐	☐
7	**Funding and authors** **Strong:** independent funding with disclosures. **Red flag:** vendor-funded, vendor-authored, no independent study. Notes:	☐	☐	☐
8	**Final and transparent** **Strong:** final, preregistered, data or protocol available. **Red flag:** preliminary results presented as conclusive. Notes:	☐	☐	☐

Based on CLAIM 2024, CONSORT-AI and STARD-AI reporting guidelines.
Full guide: blog.medicai.io/en/evaluate-radiology-ai-vendor-evidence/

MEDICAI
medicai.io