Evaluation of Compliance with Consolidated Standards of Reporting Trials for Artificial Intelligence in Imaging Randomized Controlled Trials: A Systematic Review

Fatema Aftab[1], Ibrahim Riaz, Wameth Alaa Jamel, Panagiotis Bompolas, Makoto Shiraishi, Gbenga Adesoye, Kian Daneshi, Ankur Khajuria

1.Cambridge University Hospitals NHS Foundation Trust

Purpose

The Consolidated Standards of Reporting Trials for Artificial Intelligence (CONSORT-AI) guidelines were developed to improve the reporting quality of randomized controlled trials (RCTs) involving AI. Underreporting of CONSORT-AI undermines transparency, reproducibility and evidence-based decision-making in imaging AI research. This systematic review evaluates how well these guidelines are followed in imaging AI RCTs, identifying gaps in reporting standards.

Methods

Multiple databases including MEDLINE and Cochrane Database of Systematic Reviews were searched comprehensively on 9 December 2024 for the medical subject headings 'imaging', 'artificial intelligence' and 'randomised control trial'. Limitations were entered for English language, human studies and RCTs. Articles on the clinical use of diagnostic and interventional radiology were manually searched. A scoring system awarded 1 point for each applicable CONSORT-AI sub-criterion met, and overall compliance was calculated as a percentage.

Results

32 papers met the inclusion criteria. The overall average compliance of the included studies with the CONSORT-AI guidelines was 56.7%. Lowest reported sub-criteria were the accessibility of the AI code (21.9%), the discussion of important unintended effects in each study arm (28.1%), and the handling of poor-quality data (30%). Highest reported sub-criteria were declaration of funding sources (90.3%) and the inclusion of the trial registry number (93.8%). No statistically significant positive correlation was found between CONSORT scores and journal impact factors ($R^2 = 0.216$, $p = 0.332$) or the number of authors ($R^2 = 0.033$, $p = 0.177$).

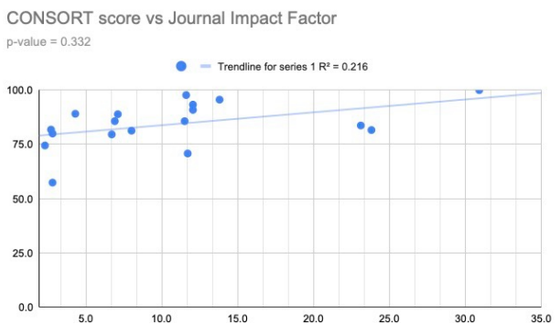


Figure 1: Relationship between CONSORT-AI Compliance Score and Journal Impact Factor

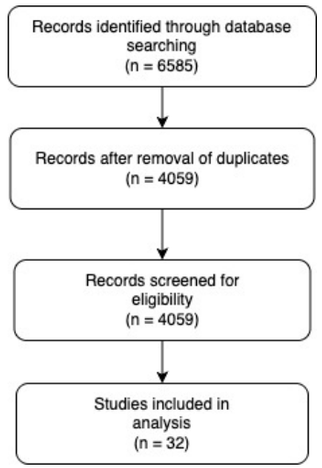


Figure 2. PRISMA Flowchart for systematic study selection

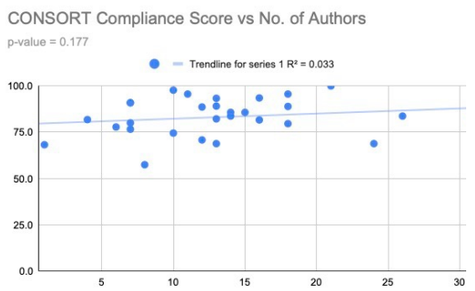


Figure 3. Relationship between CONSORT-AI Adherence and Number of Authors

Discussion

This systematic review—the first of its kind in diagnostic and interventional imaging AI—found that RCTs in this domain demonstrate poor adherence (56.7%) to CONSORT-AI reporting guidelines. Key deficiencies were identified in the reporting of data quality handling, algorithm transparency, and generalisability, with only 30% addressing poor-quality data management and just 21.9% reporting code accessibility. Over half of studies originated from Asia, limiting global applicability. These gaps threaten the transparency, reproducibility, and clinical safety of AI tools. Future efforts should focus on enhancing guideline awareness, promoting multicenter and diverse datasets, and encouraging code sharing to strengthen reporting standards.

Conclusion

Compliance with CONSORT-AI is suboptimal. Journals should urgently address this by requiring mandatory adherence to CONSORT-AI checklists as part of the manuscript submission process.

References

Scan QR Code for References