This article has been accepted for publication in IEEE Transactions on Medical Imaging. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TMI.2026.3658770

Editorial
AI Reviewer (AIR) Trial for Responsible, Secure, and Efficient Peer Review

Ge Wang, *Life Fellow, IEEE*, Tolga Çukur, *Senior Member, IEEE*, Uwe Kruger, *Senior Member, IEEE*, Jennifer Ferina, *Member, IEEE*, and Hongming Shan, *Senior Member, IEEE*

Peer review is central to the integrity of scientific publishing. At IEEE Transactions on Medical Imaging (TMI), thousands of reviewers and editors work each year to ensure that accepted papers meet our high standards of significance, innovation, evaluation, and reproducibility (SIER) [1]. Yet the rapid growth in submissions, the increasing complexity of papers, and the decreasing availability of reviewers place mounting pressure on the TMI peer review system.

Recently, large AI models have begun to enter the peer review process. Several studies have shown that AI systems can generate review-like feedback that many authors and editors find useful, while raising serious questions about quality, bias, transparency, and integrity. A large-scale empirical analysis in NEJM AI found that GPT-4—based feedback on research papers was often judged as helpful, and in some cases more beneficial than traditional reviews, by a substantial set of users [2]. At ICLR 2024, a leading AI conference, a quasi-experimental study estimated that at least 15–16% of reviews contained substantial AI-assisted content and that such reviews were associated with slightly higher scores and acceptance rates for borderline papers in that setting [3]. Conceptual and survey articles have begun to map the opportunities and risks of "AI-assisted peer review," including potential improvements in consistency and efficiency, as well as concerns about opacity, bias, and over-reliance [4], [5]. Medical journals are responding as well. A recent analysis of the top 100 medical journals reported that the majority now provide explicit guidance on the use of AI in peer review, reflecting both growing uptake and the need for governance [6]. Recent surveys on automated scholarly paper review have synthesized the growing literature, cataloging tools, use cases, and open challenges [7].

Together, these works show that AI-assisted peer review is already influencing decisions in large venues and that structured, responsible frameworks are urgently needed. In this context and following our "AI for TMI" (AI4TMI) initiative [8], here we introduce the AI Reviewers (AIR) project as our effort to explore AI as a secure, transparent, and rigorously evaluated assistant in the TMI editorial workflow. We agree with Mann et al. that the ethical and governance dimensions of AI-assisted peer review depend as much on oversight and transparency as on technical performance [9]. AIR is our approach to learning from all these experiences while meeting TMI-specific requirements on security, evaluation, and governance.

I. AI Reviewer (AIR) for TMI

The AI Reviewer (AIR) project, supported by IEEE, aims to develop and evaluate AI systems as assistants in the TMI editorial workflow. The design is intentionally multi-team, multi-stage, and conservative in scope. Five independent research teams are drawn from the TMI editorial board, covering Asia, Europe, and America. Each team develops or adapts large AI models (open- or closed-source) to generate structured peer-review reports that follow a harmonized template including a summary of the manuscript, major and minor comments, evaluation along TMI's SIER dimensions, and an editorial recommendation.

The proceeds in four stages. (1) Model Development: Teams build AIR systems and provide workflows that can run within a secure environment and output review reports in a standardized format. (2) Evaluation on Public Data: AIR systems are first evaluated on publicly available, anonymized manuscripts and reviews (e.g., OpenReview), where each paper has multiple human- and AI-generated reviews. A panel of experienced Associate Editors scores individual reviews for specificity, correctness, and usefulness in addressing the SIER dimensions, blinded to whether the reviews were human- or AI-generated. (3) Evaluation on TMI Data under Firewall Protection: The best-performing AIR model(s) from Stage 2 are then integrated and tested on TMI manuscripts in retrospective and prospective

Ge Wang is with Department of Biomedical Engineering, School of Engineering, Biomedical Imaging Center, Center for Computational Innovations, Center for Biotechnology and Interdisciplinary Studies, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: gewang@ieee.org)

Tolga Çukur is with the Dept. of Electrical-Electronics Engineering and National Magnetic Resonance Research Center (UMRAM), Bilkent University, Ankara, Turkey, 06800 (e-mail: cukur@ee.bilkent.edu.tr)

Jennifer Ferina is an independent researcher, Seattle, WA USA. e-mail (jferina@ieee.org)

Uwe Kruger is with Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: uwe.kruger@gmail.com)

Hongming Shan is with Institute of Science and Technology for Brain-Inspired Intelligence, MOE Frontiers Center for Brain Science, Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence, Fudan University, Shanghai 200433, China (e-mail: hmshan@ieee.org)

double-blind studies conducted inside the IEEE/TMI firewall, under IRB and IEEE data-use approvals. One AI-generated review is added to the set of human reviews. Then, editors and authors provide feedback, and performance is assessed using the same rubric as that used in Stage 2. (4) Deployment and Dissemination: Contingent on satisfactory results, the top AIR implementations will be integrated into the TMI editorial workflow as optional assistants for editors entirely inside secure infrastructure. Throughout all stages, AIR is advisory: final editorial decisions remain in human hands.

The evaluations of AIR are structured around two core objectives. The first is to assess whether AIR can meaningfully contribute to substantive peer review while remaining consistent with established editorial standards. This involves evaluating AIR's ability to identify technical strengths and weaknesses, provide feedback that is specific, correct, and constructive, and raise issues that are relevant to editorial decision-making. The second objective is to assess AIR's practical utility and reliability within the editorial workflow. To operationalize these objectives, AIR will be evaluated using the following success criteria:

- Review quality metrics (specificity, correctness, constructiveness),
- Consistency with SIER dimensions [1],
- Reduced burdens to reviewers and editors,
- Absence of adverse effects (*e.g.,* , over-reliance, gaming), and
- Author and editor feedback.

These criteria will be examined through structured comparisons with human-generated reviews and editor assessments, as well as through measures of reviewer and editor burdens, adverse effects, and qualitative feedback from editors and authors. Importantly, both positive and negative outcomes are considered informative for guiding editorial policy and making responsible deployment.

In the revision phase, the data from human reviewers and AI Reviewer will be paired, facilitating supervised and contrastive learning and enhancing both human and AI review quality. At the same time, human review data can serve as a constraint for AI Reviewer, preventing AI from developing biases or hallucinations, much like how external data guides information retrieval in RAG systems. This approach helps ensure that AI's evaluations remain accurate and reliable. By enabling such interactions at the revision phase, the editorial process will become more refined, bringing AI Reviewer closer to practical application and optimizing the overall quality of the peer review.

The transparency in this editorial refers to the disclosure and governance of editorial processes, rather than symmetry of tool access between editors and authors. Specifically, we distinguish three complementary dimensions of transparency:

1) Process transparency: clear disclosure that AIR is used, when it is used, and how its outputs are incorporated into editorial decision-making.
2) Governance transparency: documented evaluation protocols, success metrics, oversight mechanisms, and performance reporting (including negative results and limitations).
3) Accessibility transparency: clarification that during the trial phase AIR is not exposed to authors due to confidentiality, security, and policy constraints. Any future forms of controlled access may be considered subject to empirical evidence and policy alignment.

II. Benefits and Risks

The AIR project could potentially offer several benefits. First, AIR promises reduced reviewer burden. TMI receives about 4,000 submissions per year. Even a modest reduction of 3 hours per paper in human (assuming about 30% of the submissions for external review) effort would correspond to more than 3,600 reviewer-hours saved annually. Also, the time associate editors saved should be substantial in the cases of triaged articles. Second, AIR improves consistency and coverage. By enforcing a standardized structure and prompting explicit consideration of SIER criteria, AIR may help reduce variability between reviews, flag missing elements (such as absent external validation or incomplete method descriptions), and provide more uniform baseline feedback. Third, AIR is faster. While AIR may mitigate certain human conflicts of interest (e.g., personal, institutional, or competitive conflicts), it is still susceptible to data- or model-driven biases. Detecting, measuring, and characterizing such biases is among the goals of the AIR evaluation framework. Structured AI reports can help Associate Editors synthesize multiple reviews and identify key points more quickly, potentially shortening turnaround times while maintaining or enhancing decision quality. In this sense, AIR is primarily as a consistency check, not as a replacement for expert scientific judgment. With this framing, AIR represents our initial effort in responsible publishing innovation. With AIR, TMI can move from ad-hoc, undisclosed use of external AI tools to a controlled, auditable, and secure framework that sets a positive example for the wider community.

At the same time, AIR is explicitly recognized as a high-risk and experimental initiative whose limitations must be articulated with equal clarity. AI-generated reviews may miss subtle methodological flaws, over-emphasize superficial elements, or inadvertently propagate biases contained in their training data. Our multi-stage evaluation framework is designed to measure and understand these issues. Also, there is a risk that editors may over-rely on AI feedback, particularly when its language is fluent and confident. Recent observations in the broader community indicate that some authors may attempt to influence AI reviewers through hidden text prompts or other adversarial strategies. Clearly, any deployment of AIR must incorporate safeguards against manipulation and be accompanied by policies on misconduct. Surveys and commentaries reveal divergent perspectives on the appropriateness of AI in peer review, ranging from optimism for improved efficiency to concerns about diminishing scholarly responsibility. Given these considerations, AIR will remain optional, transparent, and strictly subordinate to human judgment. AIR will not serve as a sole or primary basis for editorial decisions, and its use will be continually reassessed through empirical evidence and community feedback. Again, negative or mixed findings from these evaluations will be treated as equally informative for shaping future policy. We note that AIR is not currently available for author pre-submission use, and in future we expect the potential trade-offs of such access, including possible quality improvements versus risks of adversarial behaviors driven by misaligned incentives.

III. Data Security and Technical Infrastructure

A defining feature of AIR, compared with many existing uses of AI in peer review, is its strict security model. Manuscripts and reviews are processed on a dedicated GPU workstation deployed inside the IEEE firewall and processed in the secure mode in compliance with IEEE policy [10] so that data will not be used for training other models. OpenAI and other major AI companies support the secure mode, which is designed to meet industry-standard security and data-protection practices. Furthermore, access is restricted to authorized personnel, with technical and procedural safeguards. All AIR workflows are containerized or otherwise encapsulated to ensure reproducibility and auditability. Our AIR design has placed the highest priority to address a central concern in the broader literature on AI-assisted peer review: how to gain potential benefits of AI without compromising confidentiality, intellectual property, and compliance with journal policies.

IV. Call for Participation

The AIR project is our community effort. We invite authors, reviewers, and editors to participate in evaluation studies, perform single-blinded scoring of AI and human reviews, and share qualitative feedback. Authors and editors will have opportunities to indicate their willingness and contribute perspectives on how AI-generated comments affect the peer-review process and experience, especially in prospective pilot studies. We will stay in close collaboration with IEEE publishing experts and other IEEE journals to discuss about governance, ownership, and potential refinement and extension of AIR, even beyond TMI. We will report progress through TMI, its sponsoring societal channels, conferences, and journals on both successes and limitations to inform the academic publishing ecosystem.

AI tools are already influencing peer review across disciplines, often in ways that are invisible and unregulated. By AIR, TMI aims to move from fragmented, ad-hoc use to a structured, secure, and evidence-based framework for AI-assisted peer review. Our goals are modest but important: to reduce human burden, improve consistency, and support more timely and informed editorial decisions, without compromising confidentiality, integrity, or human responsibility. Through multi-stage evaluation, strong data security, and active community participation, we will rigorously assess whether and how AI can serve as a trustworthy assistant in the TMI editorial workflow.

We welcome your feedback as we embark on this exciting next step in TMI's evolution.

References

[1] H. Shan, U. Kruger, and G. Wang, "Criteria for TMI papers—significance, innovation, evaluation, and reproducibility," *IEEE Transactions on Medical Imaging*, vol. 44, no. 12, pp. 4746–4748, 2025.

[2] W. Liang *et al.*, "Can large language models provide useful feedback on research papers? a large-scale empirical analysis," *NEJM AI*, vol. 1, no. 8, p. AIoa2400196, 2024.

[3] G. R. Latona, M. H. Ribeiro, T. R. Davidson, V. Veselovsky, and R. West, "The AI review lottery: Widespread ai-assisted peer reviews boost paper scores and acceptance rates," *arXiv:2405.02150*, 2024.

[4] A. Checco, L. Bracciale, P. Loreti, S. Pinfield, and G. Bianchi, "AI-assisted peer review," *Humanities and Social Sciences Communications*, vol. 8, no. 1, pp. 1–11, 2021.

[5] J. Zou, "ChatGPT is transforming peer review—how can we use it responsibly?" *Nature*, vol. 635, no. 8037, pp. 10–10, 2024.

[6] Z.-Q. Li, H.-L. Xu, H.-J. Cao, Z.-L. Liu, Y.-T. Fei, and J.-P. Liu, "Use of artificial intelligence in peer review among top 100 medical journals," *JAMA Network Open*, vol. 7, no. 12, p. e2448609, 2024.

[7] Z. Zhuang, J. Chen, H. Xu, Y. Jiang, and J. Lin, "Large language models for automated scholarly paper review: A survey," *Information Fusion*, p. 103332, 2025.

[8] G. Wang, "Flagship toward the future," *IEEE Transactions on Medical Imaging*, vol. 44, no. 3, pp. 1113–1114, 2025.

[9] S. P. Mann *et al.*, "Ai and the future of academic peer review," *arXiv:2509.14189*, 2025.

[10] IEEE Author Center, "Submission and peer review policies," https://journals.ieeeauthorcenter.ieee.org/become-an-ieee-journal-author/publishing-ethics/guidelines-and-policies/submission-and-peer-review-policies/, 2026, accessed: 2026-01-04.