

# Setting guidelines to report the use of AI in clinical trials

Delivering the potential of artificial intelligence in clinical decision-making will require testing interventions in well-designed randomized clinical trials and reporting these results in a standardized and transparent fashion.

Artificial intelligence (AI) methods developed in the past decades have made invaluable contributions to biomedical research. Recent technological progress in machine learning and deep learning algorithms, and their application in solving clinical questions, are expanding the possibilities for enhancing healthcare delivery and hold great promise for transforming clinical research. Several proof-of-concept studies have illustrated the capacity of these models, when trained on sufficiently large datasets, to attain image-based [diagnosis](#) accuracy compatible with clinical applications or to select optimal [treatment](#) regimens for hospitalized patients, among other tasks. However, the utility of most of these algorithms remain largely theoretical, and they have mostly been tested in controlled settings that cannot recapitulate the complexities of the real world.

At this crossroads, when the value of AI approaches for patient management is put to the test, it is crucial that steps be taken to ensure the highest quality in the reporting of prospective randomized clinical trials that involve AI-based interventions. With this goal in mind, the CONSORT-AI and SPIRIT-AI Steering Groups coordinated a Delphi process involving multiple stakeholders — trialists, statisticians, clinical and translational researchers, patients, regulators and editors — to elaborate specific guidelines aimed at increasing transparency of study protocols and reporting for randomized clinical trials involving AI. The resulting checklists, [CONSORT-AI](#) and [SPIRIT-AI](#), published in this issue along with their respective explanatory documents, represent an extension to the parental CONSORT and SPIRIT guidelines that have raised the

impact and quality of study protocols and reporting for randomized clinical trials.

The new extensions incorporate a series of items that were inadequately covered or not covered at all by current guidance, such as how AI is integrated in the clinical pathway or aspects related to code availability or assessment of performance. As with the original clinical guidelines, the CONSORT-AI and SPIRIT-AI extensions provide a set of principles in a burgeoning field of research, and will evolve and be revised as technological advances and clinical needs require. We at *Nature Medicine* consider it crucial to follow a standardized, transparent and rigorous report procedure for AI interventions in clinical research to ensure the correct steps are taken and the field advances in the right direction. Therefore, consistent with our mission to nurture high-quality reporting of clinical research, we endorse the CONSORT-AI and SPIRIT-AI guidelines and will require that submissions of manuscripts describing the results of clinical trials using AI algorithms in the clinical decision-making process be reported in accordance with these standards. An example using the CONSORT-AI extension can be seen in the [ADVICE4U](#) study in this issue — a randomized non-inferiority trial comparing insulin dosing for youths with type 1 diabetes calculated by an AI-based decision-support system with that of physicians. The study has been reported in accordance with the new guidelines, including completion of the CONSORT-AI checklist.

In the process of elaborating the extensions, it became clear that the incorporation of AI technologies in clinical care also creates new challenges that will need to be overcome to narrow the gap between simulated and real-world medical

AI. Some of these challenges are tackled in a series of commissioned Comments in this issue. Eric Topol provides a [snapshot](#) of where medical AI stands and highlights the limitations and challenges that these algorithms and study designs need to overcome to be effectively implemented in clinical care. Melissa McCradden and colleagues propose a [step-wise process](#) to ensure a positive impact of this technology in patient care, which includes considerations of data access and protection, model performance and deployment, and oversight. For these models to ensure safety and fairness in clinical care, it is also essential that the design of algorithms be devoid of any bias that might introduce structural inequity in the predictions, as Kellie Owens and Alexis Walker [write](#). Similarly, Atul Butte and colleagues also propose a framework, the [MI\\_CLAIM](#) guidelines, describing the minimal reporting elements that are required for ensuring the transparency, reproducibility and utility of AI algorithms in medicine.

Clinical research is on the brink of a new phase where innovation has enormous potential to bring new opportunities for advancing healthcare delivery. There are, however, risks that need to be anticipated, and necessary steps must be taken to ensure that AI-enabled solutions prioritize the needs of all patients and, in doing so, earn the trust of users. The CONSORT-AI and SPIRIT-AI guidelines lay the foundation for a responsible and transparent evaluation of these tools, and we look forward to seeing how the promise of AI-enhanced healthcare will be fulfilled. □

Published online: 9 September 2020  
<https://doi.org/10.1038/s41591-020-1069-z>