# Global initiative aims to enhance inclusivity in medical AI datasets



A collaborative research initiative involving over 350 professionals from 58 countries has produced a comprehensive set of recommendations aimed at promoting inclusivity in datasets employed for training medical artificial intelligence (AI) systems. This initiative, named the STANDING Together project, appears in a report published in The Lancet Digital Health, highlighting the potential for these systems to deliver safe and effective healthcare solutions accessible to all members of society.

The STANDING Together project utilised a Delphi approach, enhanced by a public consultation and an international interview study, to gather input from global experts and establish 29 consensus recommendations. The project has explicitly divided its findings into two sets of recommendations: the first comprising 18 guidelines aimed at improving the documentation of health datasets, and the second consisting of 11 recommendations to enhance the utilisation of these datasets.

The initial recommendations focused on assisting dataset curators in fostering transparency regarding the composition and limitations of health datasets. Notable suggestions include providing a summary of the dataset in plain language, elucidating the dataset’s creation background and purpose, revealing the origin of the data, and detailing the demographics represented, including any known missing groups and the reasons for their absence.

The second set of recommendations seeks to address algorithmic bias within health datasets. This involves ensuring sufficient dataset information to support traceability and auditability, identifying contextualised groups who may face risks from AI technologies, and justifying the appropriate application of datasets to their intended populations. Furthermore, reporting on the performance of AI health technologies for specific groups of interest is suggested for enabling comparisons with aggregate performance metrics.

Dr Xiao Liu, an associate professor of AI and Digital Health Technologies at the University of Birmingham, served as the chief investigator for the study. He stressed the potential for data to “magnify societal biases” and emphasised the importance of addressing root causes to foster health equity, stating, “To create lasting change in health equity, we must focus on fixing the source, not just the reflection.”

The trends surrounding AI implementation in healthcare also point to challenges pertaining to regulation and the management of bias. A discussion panel held last year under the banner of HTN Now featured industry specialists exploring these themes, including Puja Myles from the MHRA, Shanker Vijayadeva from NHS England, and Ricardo Baptista Leite from HealthAI. They addressed inquiries surrounding responsible AI practices and the necessity for inclusive and equitable AI technologies, particularly pertaining to their deployment in the NHS and their potential role in transitioning from reactive to proactive health care.

A recent poll conducted on LinkedIn revealed that among over 100 participants, a majority of 52 percent identified regulation as their primary concern regarding AI in healthcare. Additional questions surrounding the barriers to responsible AI yielded a notable focus on inadequate regulation, which garnered 38 percent of the votes, whilst a lack of standardised data followed closely at 33 percent.

In parallel, Health Level Seven International (HL7) has circulated new guidelines concerning the AI and machine learning data lifecycle to aid developers in enhancing the quality and trustworthiness of interoperable data for AI applications. This guidance includes case studies illustrating AI’s influence on healthcare decision-making amid various patient scenarios.

Moreover, NHS England has released advisory content on evaluating AI projects, drawing lessons from the four-year ‘Artificial Intelligence in Health and Care Award’ programme, designed to facilitate the development and deployment of promising AI technologies. Concurrently, NICE has introduced a new reporting standard intended to bolster transparency in cost-effectiveness analyses of AI treatments, aiming to expedite patient access to promising AI-enabled healthcare solutions.

Source: [Noah Wire Services](https://www.noahwire.com)

## Bibliography

1. <https://www.med-technews.com/medtech-insights/digital-in-healthcare-insights/standing-together-initiative-outlines-new-ai-medical-standar/> - This article explains the STANDING Together initiative, its international scope, and the comprehensive set of recommendations to ensure medical AI systems are developed with diverse and inclusive health datasets.
2. <https://www.eurekalert.org/news-releases/1068414> - This source details the STANDING Together recommendations published in The Lancet Digital Health, emphasizing the need for diverse and representative datasets to reduce AI bias and ensure equity in health.
3. <https://www.hdrmidlands.org.uk/hdrproject/standing-together/> - This page describes the STANDING Together project's mission to develop standards for diverse, inclusive, and generalizable datasets for training and testing AI systems in healthcare.
4. <https://www.datadiversity.org/the-data-gap> - This article highlights the findings of the STANDING Together project regarding the lack of diversity in health datasets and provides guidance on dataset documentation to promote transparency.
5. <https://www.med-technews.com/medtech-insights/digital-in-healthcare-insights/standing-together-initiative-outlines-new-ai-medical-standar/> - This source mentions Dr. Xiaoxuan Liu's role as the lead researcher and his comments on addressing societal biases in health data to foster health equity.
6. <https://www.eurekalert.org/news-releases/1068414> - This article includes quotes from Sir Jeremy Farrar and Dominic Cushnan, emphasizing the importance of transparent and representative datasets for the responsible development and use of AI in healthcare.
7. <https://www.hdrmidlands.org.uk/hdrproject/standing-together/> - This page outlines the project's approach, including the involvement of various stakeholders and the funding by The Health Foundation and the NHS AI Lab.
8. <https://www.datadiversity.org/the-data-gap> - This source provides a summary of the reviews and findings from different disease areas, highlighting the need for diverse and representative data in AI healthcare technologies.
9. <https://www.eurekalert.org/news-releases/1068414> - This article discusses the recommendations for improving dataset documentation and addressing algorithmic bias, ensuring that AI systems are safe and effective for everyone.
10. <https://www.med-technews.com/medtech-insights/digital-in-healthcare-insights/standing-together-initiative-outlines-new-ai-medical-standar/> - This source details the collaborative nature of the STANDING Together project, involving over 350 people from 58 countries, and the support from various international institutions.
11. <https://www.hdrmidlands.org.uk/hdrproject/standing-together/> - This page explains the project's focus on ensuring that no one is left behind by developing standards for AI datasets that are inclusive and representative of diverse populations.
12. <https://htn.co.uk/2025/01/09/research-study-involving-350-experts-from-58-countries-offers-recommendations-to-ensure-inclusivity-of-datasets-for-medical-ai/> - Please view link - unable to able to access data