Singapore to build first global medical AI foundation model

09 September 2025 | News

The data will span Africa, the Middle East, South America, Southeast Asia, the Western Pacific, and the Caucasus region.

A global research consortium of over 100 study groups in more than 65 countries has launched the Global RETFound initiative, a collaborative effort to develop the first globally representative Artificial Intelligence (AI) foundation model in medicine, using 100 million eye images.

The project addresses growing concerns about AI bias in healthcare while demonstrating how international collaboration can advance medical AI development in an equitable way. One of the largest medical AI collaborations ever undertaken, it produces a medical dataset that is geographically and ethnically diverse. The data will span Africa, the Middle East, South America, Southeast Asia, the Western Pacific, and the Caucasus region. The consortium welcomes additional researchers and institutions to join their collaborative effort towards more inclusive medical AI development.

Led by researchers from the National University of Singapore Yong Loo Lin School of Medicine (NUS Medicine), Moorfields NHS Foundation Trust, University College London (UCL), and the Chinese University of Hong Kong (CUHK), the consortium will develop its model using an unprecedented dataset of over 100 million color fundus photographs (photos of the back of the eye), sourced from more than 65 countries. The global initiative builds on the success of RETFound, the first foundation model for retinal and systemic disease detection.

While RETFound demonstrated potential for medical AI applications, the next global model will expand the training data to encompass every continent except Antarctica. "Current foundational models are trained on data that is geographically and demographically 'narrow', which limits their effectiveness and can perpetuate existing health inequalities," explained Dr. Yih Chung Tham, Assistant Professor at NUS Medicine, and a NUS Presidential Young Professor, one of the project key leads.

"The Global RETFound Consortium addresses this challenge through innovative approaches that enable broad international participation at unprecedented scale, while maintaining data privacy protections."

The project features a flexible, two-pronged data sharing framework that accommodates varying technical capacities and regulatory requirements. The first approach involves local fine-tuning of generative AI models at individual institutions, with only model weights shared centrally, ensuring no patient data leaves the originating site. The second pathway enables direct sharing of de-identified data through secure infrastructure for institutions that do not have local GPU resources or technical expertise. By combining real and synthetic data generation techniques, we can build a diverse, globally representative dataset without compromising security.