[1] Child Health Research Foundation, Dhaka, Bangladesh

Cite this as: *BMJ* 2024;384:q27
http://dx.doi.org/10.1136/bmj.q27
Published: 09 January 2024

DECOLONISING HEALTH AND MEDICINE

Code sharing and artificial intelligence can help decolonise public health modelling

Code sharing and artificial intelligence have potential to empower researchers from low and middle income countries to develop their use of public health modelling

Yogesh Hooda, [1] Senjuti Saha[1]

Steps towards democratisation of efforts in global health and capacity strengthening in low and middle income countries (LMICs) have made significant progress in recent years.[1-4] These efforts have empowered local researchers with the skills to collect, process, and analyse data, bridging the data generation gap between the global north and south. But glaring disparities remain in the realm of modelling and prediction studies, which are predominantly led by countries in the global north, often with limited transparency in sharing the underlying codes.[5] This absence of mandates to share codes when publishing in journals exacerbates the challenge for researchers from LMICs, who are often located in the global south. Artificial intelligence (AI) and ethical code sharing practices can narrow this divide, empowering researchers in resource constrained settings to benefit from public health modelling.

Public health modelling for predictive or analytical purposes directly influences public health policy and decision making in LMICs—this makes it imperative that models are relevant and effective in the context of these countries.[6,7] Modelling studies, requiring expertise in mathematics, statistics, and programming, are uniquely challenging in the context of decolonisation owing to the need to navigate intricate sociopolitical dynamics and diverse health systems, while also respecting local knowledge and cultural practices. The complex web of assumptions in these models, which are intrinsic to the modelling process, are often not explicitly detailed in publications, making them nearly impossible to reproduce by local scientists. These assumptions pose a barrier to scientists in the global south, limiting their contribution to modelling and manuscript writing.

Furthermore, when models are developed by scientists in the global north, the codes are typically owned by their creators. The ownership of codes by authors primarily in the global north creates a hierarchy of access to knowledge, hindering researchers from LMICs from developing and applying their own models. This stifles innovation and reinforces the notion that knowledge production and technological advancements are centralised in the global north.

The absence of code sharing mandates in many academic journals is a major obstacle to tackling this problem. Journals have a pivotal role in shaping research practices. Organisations including the International Committee of Medical Journal Editors[8] and the Committee on Publication Ethics[9] can help rectify this by mandating code sharing as a standard practice. Open source code sharing in publicly available repositories can promote transparency, reproducibility, and inclusivity in health modelling research.[10] It can ensure that researchers worldwide, including those located in LMICs, have access to the tools and methods used for health modelling, thereby empowering them to adapt and improve on existing models.

Code sharing mandates, however, cannot extend back to the large number of public health modelling studies that have already been published. In this context, AI presents an opportunity for scientists in the global south to leapfrog the barriers to public health modelling. Large language models like GPT-4, Bard and Llama-2 are AI programs that have the capacity to generate codes based on written prompts. Using descriptions of a model, often found in the methods section of papers, these tools can regenerate existing models from descriptions with high accuracy.[11,12] But the design of complex models can require engineering of the prompts and repeated tuning. Large language models can also be used to improve code readability and provide detailed explanation of the codes when they are shared, without the requirement of modelling, coding, and decoding expertise. Thus, large language models have the potential to empower researchers from LMICs by providing them access to modelling tools that can simplify complex modelling techniques and provide a foundation for further customisation and expansion.

For effective use of AI for public health modelling, scientists from LMICs need a foundational understanding of AI technologies and must proactively learn to comprehend their functions.[13,14] Equally crucial is an awareness of the limitations of AI based tools. These tools are often trained on existing datasets that might not fully represent the unique contexts of LMICs and typically lack nuanced medical expertise. Additionally, AI models can perpetuate biases present in their training data, leading to skewed or inaccurate results, particularly in diverse settings that differ substantially from the original data. This underscores the need for careful, context aware application and continuous evaluation of AI tools in public health endeavours. To test the practicality and effectiveness of such a learning

approach, a workshop called "AI in Public Health" was designed and offered in Bangladesh in October 2023,[15] conducted by the Child Health Research Foundation and supported by the Bill and Melinda Gates Foundation. The training modules focused on ethical AI use and its applications in various areas of public health modelling. This module drew huge interest and showed the potential of such training in empowering researchers from LMICs. Within four days, participants were able to use GPT to generate complex codes for modelling the potential impact of introducing typhoid conjugate vaccination in Bangladesh. Without the aid of such an AI based tool, this could have taken weeks to months. Participants also learnt to generate codes for time series prediction of locally relevant epidemics like dengue fever.

Decolonisation of global health extends beyond bridging the gaps of advanced research equipment and funding between the global north and the global south. It necessitates restructuring power dynamics in research, dismantling colonial legacies, and reimagining public health landscapes. Tackling disparities in public health modelling and prediction studies is a crucial step in this direction. Ethical use of AI, coupled with code sharing, can foster a more inclusive, equitable, and effective global health research approach, ensuring that health modelling serves the needs of LMICs and advances global health.

Provenance: Commissioned; not externally peer reviewed.

Competing interests: None declared

1 Abimbola S, Asthana S, Montenegro C, etal. Addressing power asymmetries in global health: Imperatives in the wake of the COVID-19 pandemic. *PLoS Med* 2021;18:e1003604. doi: 10.1371/journal.pmed.1003604 pmid: 33886540

2 Saha S, Pai M. Can COVID-19 innovations and systems help low- and middle-income countries to re-imagine healthcare delivery?*Med* 2021;2:-73. doi: 10.1016/j.medj.2021.02.008 pmid: 33686383

3 Khan M, Abimbola S, Aloudat T, Capobianco E, Hawkes S, Rahman-Shepherd A. Decolonising global health in 2021: a roadmap to move from rhetoric to reform. *BMJ Glob Health* 2021;6:e005604. doi: 10.1136/bmjgh-2021-005604 pmid: 33758016

4 Adhikari S, Torres I, Oele ELancet Global Health. The way forward in decolonising global health. *Lancet Glob Health* 2023;11:e982. doi: 10.1016/S2214-109X(23)00275-9 pmid: 37349039

5 Chersich MF, Blaauw D, Dumbaugh M, etal. Local and foreign authorship of maternal health interventional research in low- and middle-income countries: systematic mapping of publications 2000-2012. *Global Health* 2016;12:. doi: 10.1186/s12992-016-0172-x pmid: 27338707

6 Kretzschmar M. Disease modeling for public health: added value, challenges, and institutional constraints. *J Public Health Policy* 2020;41:-51. doi: 10.1057/s41271-019-00206-0 pmid: 31780754

7 Metcalf CJE, Edmunds WJ, Lessler J. Six challenges in modelling for public health policy. *Epidemics* 2015;10:-6. doi: 10.1016/j.epidem.2014.08.008 pmid: 25843392

8 ICMJE. Recommendations. https://www.icmje.org/recommendations/browse/

9 Committee on Publication Ethics. Data and reproducibility. https://publicationethics.org/data

10 Gomes DGE, Pottier P, Crystal-Ornelas R, etal. Why don't we share data and code? Perceived barriers and benefits to public archiving practices. *Proc Biol Sci* 2022;289:20221113. doi: 10.1098/rspb.2022.1113 pmid: 36416041

11 Chen M, Tworek J, Jun H, etal. Evaluating large language models trained on code.2021. https://doi.org/10.48550/arXiv.2107.03374

12 Li Y, Choi D, Chung J, etal. Competition-level code generation with AlphaCode. *Science* 2022;378:-7. doi: 10.1126/science.abq1158 pmid: 36480631

13 Peters D, Vold K, Robinson D, etal. Responsible AI—two frameworks for ethical design practice. *IEEE Trans Technol Soc* 2020;1:-47doi: 10.1109/TTS.2020.2974991.

14 Mhlanga D. Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning.2023. doi: 10.1007/978-3-031-37776-1_17

15 Child Health Research Foundation. AI in public health workshop. https://chrfbd.org/blogs/f0cc0f49-f22e-426e-b63f-e432c89d0d0e