T20 Policy Brief

Task Force 05

INCLUSIVE DIGITAL TRANSFORMATION

Ethical Governance of Artificial Intelligence in Healthcare: Balancing Innovation with Ensuring Ethical Integrity and Gender Equity

Viola Savy Dsouza, PhD Scholar, Department of Health Information, Prasanna School of Public Health, Manipal Academy of Higher Education (India)

Philip Hines, Engagement Manager, Life Sciences and Health Policy, IQVIA (United Kingdom)

Jestina Rachel Kurian, PhD Scholar, Department of Data Science, Prasanna School of Public Health, Manipal Academy of Higher Education (India)

Elstin Anbu Raj S, Research Fellow, Department of Health Information, Prasanna School of Public Health, Manipal Academy of Higher Education (India)

Angela Brand, International Faculty, Prasanna School of Public Health, Manipal Academy of Higher Education (India), Professor for Public Health Genomics, Faculty of Health, Medicine and Life Sciences, Maastricht University (The Netherlands) and Professorial Fellow, United Nations University - Maastricht Economic and Social Research Institute on Innovation and Technology (The Netherlands)

Sarah Rickwood, Vice President, EMEA Thought Leadership, IQVIA (United Kingdom)

Helmut Brand, International Faculty, Prasanna School of Public Health, Manipal Academy of Higher Education (India) and Jean Monnet Professor of European Public Health, the Care and Public Health Research Institute, Faculty of Health Medicine and Life Sciences, Maastricht University (The Netherlands)

TF05

Abstract

Artificial Intelligence (AI) holds the promise of transforming various sectors, notably healthcare. AI's capabilities extend to improving diagnostic precision, tailoring treatment plans, and optimizing administrative processes. However, this potential comes with significant ethical challenges, particularly patient autonomy, data privacy, bias, and the risk of amplifying existing inequalities. This policy brief explores the prospects and ethical challenges AI introduces in healthcare. It underscores the necessity for developing robust ethical governance frameworks for deploying AI in collaboration with diverse stakeholders. These should result in pragmatic tools for ethical adherence assessment, stakeholder engagement, and considering the broader implications for public health. This policy brief then applies these findings to a critical area where ethical governance is needed: gender-related issues. AI can empower women and foster gender equality, but without careful oversight, it might perpetuate gender stereotypes and disparities. To tackle this, the policy brief proposes an ethical governance framework drawing inspiration from established frameworks and/or models: (i) UNESCO's Recommendation on the ethics of artificial intelligence, (ii) WHO's guidance on Ethics & Governance of Artificial Intelligence for Health. This framework enshrines principles such as respect for human autonomy, harm prevention, fairness, and transparency.

Keywords: Artificial Intelligence, AI, gender equity, ethical governance, healthcare, public health

Diagnosis of the Issue

Artificial Intelligence (AI) has emerged as a transformative force, with the potential to revolutionize various sectors, including healthcare, by enhancing diagnostic precision, tailoring treatment plans, and optimizing administrative processes(Alowais et al. 2023). Yet, it brings in ethical concerns, including challenges to patient autonomy, data privacy, and potential biases that may worsen social disparities(Dsouza et al. 2023). Moreover, technological gaps widen the digital divide, as 327 million women in low- and middle-income countries (LMICs) remain offline(United Nations 2019). This policy brief aims to explore the prospects and ethical challenges AI introduces in healthcare and advocates the development of pragmatic tools for ethical adherence assessment, stakeholder engagement, and broader public health good. The governance of AI in healthcare varies significantly across G20 countries reflecting their distinct healthcare systems, societal values, and investment priorities in AI technology.

Preceding the AI summit 2023, the USA issued an executive order for AI regulation, emphasizing disclosure of safety test results and commitments from developers to prevent harmful biological use (The White House 2023). Conversely, the UK advocates for regulation proportionate to risk to uphold public trust, balancing AI benefits with ethical concerns (Department of Science Innovation and Technology 2023). The EU, through the EU AI Act, and the General Data Protection Regulation (GDPR) has initiated regulation for AI use (European Parliament 2023; Wolford 2024). India has introduced the "Ethical Guidelines for AI in Biomedical Research and Health," while addressing challenges of data privacy and bias. In Brazil, significant AI investments and the "National Strategy for AI" prioritize ethical considerations (DHR-ICMR Artificial Intelligence Cell 2023; Filgueiras and Junquilho 2023). These examples illustrate the global diversity in

responses to AI's impact on healthcare, underscoring the importance of context driven solutions to address ethical issues for AI in health.

In 2020, the G20 Digital Economy Task Force (DETF) emphasized human-centric, data-driven, and evidence-informed policies to enhance standards of living, economic competitiveness, and societal inclusion. Through DETF, G20 members acknowledged the necessity of tackling digital divides and crafting innovative, agile, and flexible strategies suited for the digital era (G20 Research Group 2020). In parallel, the WHO and G20 India presidency launched the Global Initiative on Digital Health (GIDH) during the G20 Summit's Health Minister's Meeting. This initiative aligns with the Global Strategy on Digital Health 2020–2025, aiming to expedite the transformation of global health systems by consolidating standards, best practices, and resources. The GIDH signifies the joint commitment of G20 to utilize AI and digital health technologies for accessible, quality-assured solutions. It emphasizes international collaboration, highlighting the G20's role in fostering the ethical use of AI in healthcare (World Health Organization 2023). Additionally, Brasil's G20 presidency aligns with digital initiatives for health through its highlights data sharing as a strategic asset to address global crises, prioritizing digital technology for the Global South (Ingram and Vora 2024).

The strengths of these efforts lie in their commitment to international cooperation and the adoption of evidence-informed digital policies to address challenges but the current forms of AI governance remain fragmented addressing specific challenges such as explainability of AI, inclusive and diverse training data or self-governance for private developers (World Health Organization 2021). Moreover, they do not include deployable ethical governance frameworks that bridge digital divides, reduce inequalities and build public trust in these digital systems. The solutions put forth in this policy brief will address the need for a nuanced understanding of the ethical challenges and opportunities

presented by AI in healthcare, particularly when it comes to addressing gender-related issues.

Recommendations

As AI permeates social and economic structures, G20 countries as global leaders have moral responsibility to use the present window of opportunity to establish a unified approach as a foundation for AI-driven transformation. But a pro innovation approach to AI regulation relies on responsiveness to emerging concerns among which gender equity needs to be prioritised. To balance innovation with ethical integrity and gender equity presented by AI in healthcare, the following recommendations are proposed, with a focus on ensuring inclusivity and fairness (figure 1):

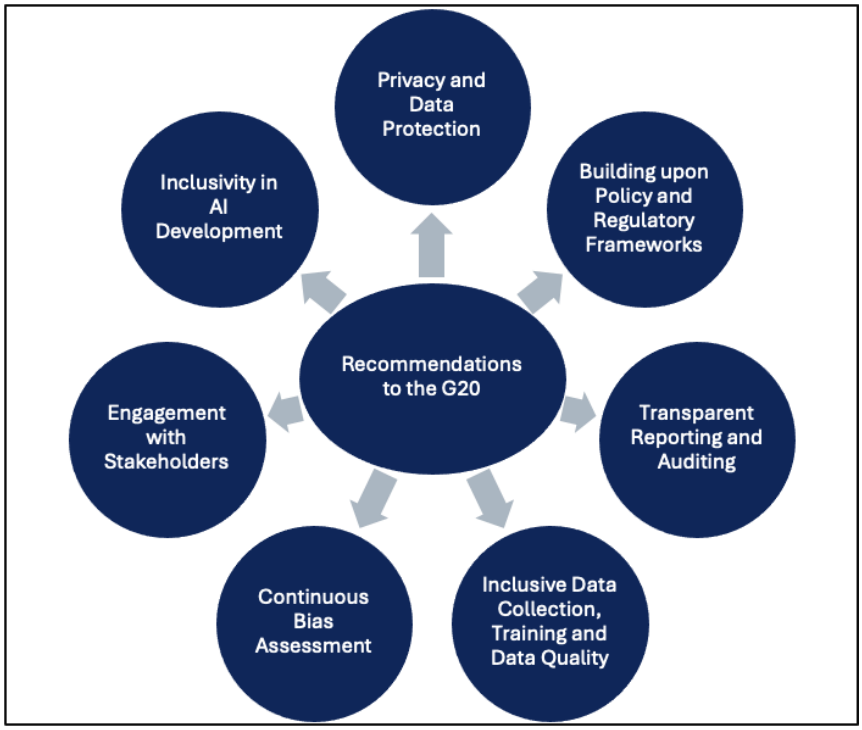

FIGURE 1: Ethical Governance Framework Recommendation

1. Transparent Reporting and Auditing: Develop standardized reporting and auditing of AI applications in healthcare, encompassing data collection, model training, validation, and deployment processes. These standards should be developed through representative collaboration, ensuring alignment with evolving ethical principles and patient rights and updated to keep pace with rapid advancements. In this direction, the EU's AI Act mandates certain AI systems to undergo a 'Fundamental Rights Impact Assessment' similar to efforts in Brasil advocating for transparency, requiring providers to conduct a preliminary risk assessment and apply proportionate regulation. Similarly, the Ethical Impact Assessment (EIA) process, as demonstrated by UNESCO, aids AI project teams in evaluating impacts and identifying relevant harm prevention actions (United Nations Educational Scientific and Cultural Organization 2023). The AI Audit Challenge 2023 introduces Meerkat, an open-source Python library facilitating participatory algorithmic audits over unstructured data, involving non-technical stakeholders. These initiatives, underline the pivotal role of transparent reporting and auditing in developing safe, fair, and equitable AI applications.

2. Representative Data Collection for Equity and Quality**:** Bias in training data can lead to serious predisposition in health outcomes and quality of health care (Norori et al. 2021). Therefore it is imperative to develop explicit directives and establish metrics to measure the inclusivity in data collection, ensuring representation across diverse demographics, geographies, and socio-economic backgrounds to mitigate bias and improve AI system accuracy. This approach aligns seamlessly with the UNESCO Recommendation on the Ethics of Artificial Intelligence, emphasizing the need for AI to promote social justice, fairness, and non-discrimination, ensuring accessibility for all. Establishing standards to uphold data quality is also essential for unlocking the full potential, performance, and relevance of AI systems, enabling value delivery, driving

innovation, ensuring ethical outcomes and trustworthiness. Research on reducing gender bias in AI systems emphases corpus-level constraints in addressing gender bias (Shrestha and Das 2022).

3. Inclusivity in AI Development: Additional to representative data, the incorporation of diversity and inclusion (D&I) principles at all AI development stages is vital (Shams, Zowghi, and Bano 2023). This concept operates at three levels: technical, (examining algorithmic fairness); community, (assessing diversity in development teams); and user, (focusing on intended users, stakeholder feedback, and responsible research and innovation principles). As an example of this, the Women4Ethical AI platform supports gender equality in AI design and deployment, aligning with the call for inclusive AI research and design (The United Nations Educational Scientific and Cultural Organization 2023). Incorporating a gender lens into data, algorithms, and healthcare practices, from deploying AI solutions to broader systemic changes is essential to achieving health equity.

4. Continuous Bias Assessment: Develop methodologies for continuous assessment of bias throughout the AI development cycle to ensure equity in AI applications, especially in healthcare. This proactive approach allows for the early identification and mitigation of biases before they become entrenched in AI systems, maintaining the integrity of AI technologies and ensuring they serve the public benefit effectively. The concept of data drift, including concept drift, highlights the dynamic nature of data and the challenges it poses to AI systems. For instance, investigation into sex disparities in published AI models for disease detection revealed that the models performed worse in detecting liver and cardiac disease in women compared to men (Straw and Wu 2022). This stresses the importance of continuous assessment with consideration for

intersectionality to avoid reinforcing bias particularly in the context of gender bias inherent in the health system (Villines 2021).

5. Engagement with Stakeholders: Inclusive engagement with stakeholders, spanning healthcare professionals, patients, and ethicists, is indispensable for crafting AI applications that ensure gender equality and sensitivity to the diverse needs and concerns of various social groups. This approach necessitates dedicating time and resources to foster inclusive development from the outset, integrating ethical considerations and maintaining organizational support throughout the development lifecycle. By identifying and rectifying stakeholder exclusion, AI technologies can be tuned to physiological and socio-cultural factors that are overlooked for women in conventional health care.

6. Building upon Policy and Regulatory Frameworks: Enhance policy and regulatory frameworks in G20 countries to address the gender biases and disparities in AI algorithms and healthcare practices, ensuring equitable access and outcomes for all individuals. Accountability and countermeasures against discrimination or bias, driven by existing gender inequalities, must be explicit and actionable within member state's AI transformation. The USA's Federal Register's AI Accountability Policy Request for Comment is a good example of the role of government policy in the AI accountability ecosystem. Tools for AI regulation can be sectoral or horizontal in the broad sense, incorporating outcome-based accountability measures for stakeholders, including users (National Telecommunications and Information Administration 2023). The OECD emphasizes a risk management approach and regulatory sandboxes for testing innovative products and services with appropriate oversight and safeguards, and ethical requirements could also be trialled here (Organisation for Economic Co-operation and Development 2020).

7. Privacy and Data Protection: Data protection and privacy should stand as cornerstones for the ethical utilization of AI in health, necessitating responsible and careful handling of sensitive information with compliance. The USA's National Institute of Standards and Technology advocates for robust controls, encryption, data integrity mechanisms, access restrictions, and appropriate data retention (Joint Task Force 2020). Argentina and Brasil, have enacted legislations to safeguard personal data in AI systems Prioritizing data privacy fosters trust, aligning with ethical AI development (Agency for Access to Public Information 2023; International Association of Privacy Professionals 2020). In the EU, GDPR and the AI Act both detail stringent privacy requirements for AI use. The regulatory approach must enshrine protections against unauthorised access and respect of individual patient rights.

Outcomes' Scenario

A protective transformation for AI in health

As G20 countries are forging regulations to catch up with rapid advances in AI, this policy brief recommends key ethical governance principles to ensure gender equity and integrity of AI in health. The commitment to principles such as development and use of representative high-quality data or bias assessment activities will yield accurate models, leading to equitable healthcare outcomes. Implementation of privacy, transparency, accountability and representative stakeholder engagement should serve to enhance public trust and adoption of AI.

Nonetheless, navigating complex, and differing legal frameworks to meet regulatory requirements poses challenges. Firstly, divergence in requirements may hamper the global use of AI and policymakers need to use international fora like the G20 to harmonise AI regulation. Secondly, there are difficulties in adhering to multiple requirements; for example, adherence to privacy laws or protection of classified information may hinder transparency, necessitating a delicate balance between disclosure and compliance. Meeting these requirements is also resource intensive, involving significant investments in infrastructure and personnel. Therefore, regulations should be implemented proportionately in crises management scenarios or low-risk applications. Nevertheless, ethically governed AI should be mandated from an equity, health, and commercial standpoint. This requires continued political will, dedicated resources, technical innovation and systemic changes.

References

Agency for Access to Public Information. 2023. “Program for Transparency and Protection of Personal Data in the Use of Artificial Intelligence.” Official Portal of the Argentine State. August 30, 2023. https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-161-2023-389231/texto.

Alowais, Shuroug A., Sahar S. Alghamdi, Nada Alsuhebany, Tariq Alqahtani, Abdulrahman I. Alshaya, Sumaya N. Almohareb, Atheer Aldairem, et al. 2023. “Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical Practice.” *BMC Medical Education* 23 (1): 689. https://doi.org/10.1186/s12909-023-04698-z.

Department of Science Innovation and Technology. 2023. “A Pro-Innovation Approach to AI Regulation.” March 2023. www.gov.uk/official-documents.

DHR-ICMR Artificial Intelligence Cell. 2023. “Ethical Guidelines for Application of Artificial Intelligence in Biomedical Research and Healthcare.” Indian Council of Medical Resesarch. 2023. https://main.icmr.nic.in/sites/default/files/upload_documents/Ethical_Guidelines_AI_Healthcare_2023.pdf.

Dsouza, Viola Savy, Sanjay Pattanshetty, Julien Venne, and Helmut Brand. 2023. “Responsible Artificial Intelligence (AI) Governance Using a Relational Governance Framework.” Observer Research Foundation. February 9, 2023. https://www.orfonline.org/search?q=Responsible+Artificial+Intelligence+%28AI%29+governance+using+a+relational+governance+framework.

European Parliament. 2023. “Artificial Intelligence Act,” June. https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf.

Filgueiras, Fernando, and Tainá Aguiar Junquilho. 2023. “The Brazilian (Non)Perspective on National Strategy for Artificial Intelligence.” *Discover Artificial Intelligence* 3 (1): 7. https://doi.org/10.1007/s44163-023-00052-w.

G20 Research Group. 2020. “Ministerial Declaration G20 Digital Economy Ministers Meeting.” G20 Information Centre. July 22, 2020. http://www.g20.utoronto.ca/2020/2020-g20-digital-0722.html.

Ingram, George, and Priya Vora. 2024. “Brazil’s G20 Opportunity: A Consensus on Digital Data.” Brookings. January 9, 2024. https://www.brookings.edu/articles/brazils-g20-opportunity-a-consensus-on-digital-data/.

International Association of Privacy Professionals. 2020. “Brazilian General Data Protection Law.” October 2020. https://iapp.org/resources/article/brazilian-data-protection-law-lgpd-english-translation/.

Joint Task Force. 2020. “Security and Privacy Controls for Information Systems and Organizations.” Gaithersburg, MD. https://doi.org/10.6028/NIST.SP.800-53r5.

National Telecommunications and Information Administration. 2023. “AI Accountability Policy Request for Comment.” US Department of Commerce. December 4, 2023. https://www.federalregister.gov/documents/2023/04/13/2023-07776/ai-accountability-policy-request-for-comment.

Norori, Natalia, Qiyang Hu, Florence Marcelle Aellen, Francesca Dalia Faraci, and Athina Tzovara. 2021. “Addressing Bias in Big Data and AI for Health Care: A Call for Open Science.” *Patterns* 2 (10): 100347. https://doi.org/10.1016/j.patter.2021.100347.

Organisation for Economic Co-operation and Development. 2020. “Trustworthy AI in Health.” April 1, 2020. https://www.oecd.org/health/trustworthy-artificial-intelligence-in-health.pdf.

Shams, Rifat Ara, Didar Zowghi, and Muneera Bano. 2023. “AI and the Quest for Diversity and Inclusion: A Systematic Literature Review.” *AI and Ethics*, November. https://doi.org/10.1007/s43681-023-00362-w.

Shrestha, Sunny, and Sanchari Das. 2022. “Exploring Gender Biases in ML and AI Academic Research through Systematic Literature Review.” *Frontiers in Artificial Intelligence* 5 (October). https://doi.org/10.3389/frai.2022.976838.

Straw, Isabel, and Honghan Wu. 2022. “Investigating for Bias in Healthcare Algorithms: A Sex-Stratified Analysis of Supervised Machine Learning Models in Liver Disease Prediction.” *BMJ Health & Care Informatics* 29 (1): e100457. https://doi.org/10.1136/BMJHCI-2021-100457.

The United Nations Educational Scientific and Cultural Organization. 2023. “Artificial Intelligence: UNESCO Launches Women4Ethical AI Expert Platform to Advance Gender Equality,” April. https://www.unesco.org/en/articles/artificial-intelligence-unesco-launches-women4ethical-ai-expert-platform-advance-gender-equality.

The White House. 2023. “FACT SHEET: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence.” October 30, 2023. https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/#:~:text=The%20Executive%20Order%20establishes%20new,around%20the%20world%2C%20and%20more.

United Nations. 2019. “In Tech-Driven 21st Century, Achieving Global Development Goals Requires Closing Digital Gender Divide.” UN News Global Perspective Human Stories. March 15, 2019. https://news.un.org/en/story/2019/03/1034831.

United Nations Educational Scientific and Cultural Organization. 2023. “Ethical Impact Assessment. A Tool of the Recommendation on the Ethics of Artificial Intelligence.” https://doi.org/10.54678/YTSA7796.

Villines, Zawn. 2021. “What to Know about Gender Bias in Healthcare.” Medical News Today. October 25, 2021. https://www.medicalnewstoday.com/articles/gender-bias-in-healthcare.

Wolford, Ben. 2024. “What Is GDPR, the EU’s New Data Protection Law?” GDPR EU. 2024. https://gdpr.eu/what-is-gdpr/.

World Health Organization. 2021. “Ethics and Governance of Artificial Intelligence for Health: WHO Guidance.” 2021. https://iris.who.int/bitstream/handle/10665/341996/9789240029200-eng.pdf?sequence=1.

World Health Organization. 2023. “WHO Launches a New Global Initiative on Digital Health Supported by the G20 Presidency.” August 19, 2023. https://www.who.int/news/item/19-08-2023-who-launches-a-new-global-initiative-on-digital-health-at-the-g20-summit-in-india.

G20 BRASIL 2024
BUILDING A JUST WORLD AND A SUSTAINABLE PLANET

T20 Brasil 2024
Let's rethink the world

Let's rethink the world

CEBRI
Rethink Tank

FUNDAÇÃO ALEXANDRE DE GUSMÃO

ipea
Instituto de Pesquisa Econômica Aplicada