Big Data and AI in Sexual and Reproductive Health: A Comment

Mahesh Karra[2]
Saumya RamaRao[1]

1: Population Council, New York, NY. saumya.ramarao@gmail.com
2: Frederick S. Pardee School of Global Studies, Boston University, Boston, MA. mvkarra@bu.edu

Abstract
Big data and artificial intelligence (AI) have the potential to transform sexual and reproductive health (SRH), offering new avenues to enhance access, efficiency, and personalization in healthcare. AI-driven tools can provide opportunities to improve diagnostics and optimize resource allocation. Through data-driven insights, healthcare providers can better understand population trends, predict health risks, and tailor interventions for diverse communities, ultimately advancing gender equality and empowerment. However, the integration of AI into SRH also presents significant challenges. Ethical concerns such as informed consent, data privacy, and transparency are critical to ensuring that AI applications do not violate individual autonomy and rights. The digital divide—disparities in technology access between different regions and populations—further risks exacerbating inequalities in SRH services and provision of care. Moreover, there is a need for robust governance frameworks and global data protection laws to regulate the use of AI in SRH, and in healthcare more broadly. Programs and policies must focus on bridging these gaps, emphasizing equity and ethical considerations while leveraging AI's potential to enhance SRH services and support the vision of sexual and reproductive health and rights for all.

Introduction
The 1994 International Conference on Population and Development (ICPD) established a comprehensive Programme of Action (PoA) that focused on promoting sexual and reproductive health and rights (SRHR), gender equality, and women's empowerment worldwide (UNFPA 2014). Technology and innovation have played an instrumental role in achieving these goals by enhancing access to sexual and reproductive health information and services (World Health Organization 2024b). The PoA envisaged that technology could contribute to achieving ICPD's goals in a number of ways. That expectation has been met through technology such as telemedicine, mobile health apps, and online education platforms that have provided women and girls with critical reproductive health care. Additionally, innovative contraceptive technologies, data-driven monitoring systems, and AI tools have been introduced to facilitate personalized care and informed decision-making. Social media and digital advocacy have further raised awareness of SRHR, mobilizing communities and driving policy changes to support women's rights globally.

Despite being far-sighted, it would have been difficult for ICPD's stakeholders to imagine the world of possibility that has been opened with the Age of Computing, where information is collected, processed, and evaluated at larger volumes and with greater speed and power than ever before (Kumar 2015; Sagiroglu and Sinanc 2013; Magalhães et al. 2022). The convergence between the production of "big data" and the computational methods that that have been developed to interpret it has led to a profound shift in how analytics can inform research, policy, and practice. The integration of big data, machine learning (ML), and artificial intelligence (AI) methodologies has revolutionized how researchers examine complex problems and generate inferences, particularly in health (Yaqoob et al. 2016; Ismail and Kumar 2021). As a result, outputs from these approaches have provided policymakers with more opportunities to incorporate evidence into decision-making and policy processes (Höchtl, Parycek, and Schöllhammer 2016). Moreover, policymakers may be able to leverage data-driven insights to more efficiently respond to problems, tailor interventions, and monitor outcomes in real-time, which in turn may promote transparency and accountability (Hacker and Petkova 2017).

At the same time, the expansion of big data and AI brings a range of new challenges for researchers, policymakers, and practitioners alike. A central assumption underlying the use of big data, and the computational methods required to generate inferences from it, is that it will help to address long-standing issues related to access, equity, and discrimination in a cost-effective manner. Concerns over the ethical and responsible use of data, particularly with regards to privacy, consent, and bias, have been raised (Magalhães et al. 2022; Borda et al. 2022; Richards and King 2014). These concerns are especially relevant to the use of big data in sexual and reproductive health (SRH) research and practice, which comprise a range of sensitive issues that may put vulnerable populations at risk. By the same token, policymakers grapple with the need to establish regulatory frameworks to safeguard individual rights and promote equitable and responsible access to and use of data (Giest 2017). To this end, normative guidance bodies have begun to raise similar considerations with a wide variety of stakeholders in the public and private sectors (World Health Organization 2024b).

Equally, concerns have been raised about the resource needs and approaches by which inferences with big data are even produced. An increasing reliance on poorly understood models of prediction has fueled debates over algorithmic bias, lack of transparency, and limited accountability, as decision-making processes that are informed by "black box" approaches may, in fact, exacerbate systemic inequalities and reinforce existing biases (Sullivan 2022). Moreover, AI systems that are trained on biased data may, in turn, (re)produce biased inferences and further disadvantage those populations who would likely benefit the most (Hyde, Bachura, and Harrison 2023). Relatedly, limited access to technology (e.g. phone ownership), data infrastructure, and computation power in resource constrained settings may further widen the digital divide and compound inequities when applying AI systems to SRH issues in these settings (World Health Organization 2024b).

In acknowledging the potential of AI, the field will also have to ensure that existing inequities are not exacerbated due to a lack of understanding of how AI systems in SRH can and should operate. To this end, the level of technical innovation in these domains will require that researchers and policy-oriented consumers of AI invest more extensively in digital literacy and methods training to ensure that they have the skills needed to keep pace with an increasingly data-driven world. Equally important, we as a community of SRH researchers, practitioners, and advocates have to ensure that users of SRH services are oriented to the potential use of AI in their care so that individual rights and personal autonomy are respected.

Big Data, AI, and SRH

This commentary discusses the role of big data and AI in SRH research, policy, and practice. We first define the concepts of "big data," "machine learning," and "artificial intelligence," and their applications in the context of SRH. There is a common misperception that big data, as it applies to public health, refers to the simple aggregation of multiple data sources (such as the Demographic and Health Surveys) while using the same suite of empirical methods that are typically used for standard statistical inference (e.g. regression analyses, etc.). This is not necessarily the case – we clarify that big data constitutes both the volume of data that is used for inference as well as the methodologies that are necessary for inferring relationships and processes of interest given the size and complexity of the data. We then review the potential advantages that AI-driven predictive and generative analytics can offer in the context of SRH by producing inferences from large-scale health data, identifying at-risk populations, and informing resource allocation efforts to improve access to and the distribution of SRH services[1]. We also discuss the extent to which AI-powered approaches can support the design of person-centered SRH interventions that can be more effectively tailored to individual preferences. Third, we highlight key pitfalls and challenges to using big data analytic approaches in the context of SRH research and policy. Finally, we present the ethical implications of deploying AI solutions for conducting SRH research and informing programs and policy. We infer the extent to which AI-aided approaches can be effectively and feasibly implemented in resource-constrained settings.

Defining Big Data, and AI in Sexual and Reproductive Health

Big Data, AI, and Machine Learning

A fundamental challenge to understanding "big data" and "artificial intelligence" is that both terms are defined inconsistently and therefore interpreted differently in their applications across disciplines (Finnegan et al. 2023). Generally, however, "big data" refers to the collection, storage, and management of large amounts of structured or unstructured volumes of information that may have been collated from multiple traditional and non-traditional data sources (e.g. mobile phones, wearables, satellite imagery, among others). This definition aligns closely with the use of the term "big data" in social science research to be the integration and repurposing of multiple sources of data, including data which was not originally designed (or meant to) be repurposed (Kashyap 2021; Salganik 2019). As a result, the performance requirements for data management and inference become salient when designing a system that can effectively process data that are generated in large volumes and at significant speeds, and are highly varied in their content type, complexity, and potential value to the user (Magolous and Lorica 2009; Ishwarappa and Anuradha 2015). Given the circumstances under which big data need to be handled, new methods of computation and analysis of such data, which include data mining, predictive analytics, and artificial intelligence (AI) approaches, may be warranted. By the same token, new guidelines and protections related to the ethics of data protection and management, consent and confidentiality, and the integration of traditional and non-traditional data into existing systems also need to be considered (Cangiano 2024).

AI, in the broadest sense, is defined by the creation of artificial (machine) systems that can execute tasks and produce outputs that typically would require human intelligence (Fetzer 1990). Key tasks include learning, reasoning and problem-solving, perception, natural language processing (NLP), and decision-making. AI approaches can typically be classified into two broad types: 1) artificial general

[1] For example, the use of AI in imaging is emerging as a diagnostic tool for breast cancers (Taylor et al. 2023).

intelligence, in which machines are designed to replicate the human mind and perform any intellectual tasks that a human can perform; or 2) artificial narrow intelligence, in which machines are designed to perform one particular task extremely well. Most AI approaches and applications in health, to date, are considered to be examples of artificial narrow intelligence (22). Notable AI approaches include:

1. **Machine Learning,** where systems learn from data and improve their performance over time without being explicitly programmed. This involves using algorithms to identify patterns and make predictions. Two key types of machine learning paradigms used to train algorithms are supervised and unsupervised learning, which differ in how data is labeled and the type of tasks they perform. In supervised learning, models are trained on labeled data and learn to map inputs to known outputs for tasks like classification and regression. In unsupervised learning, models are given unlabeled data and must discover patterns or groupings without specific guidance or governing structure, as in clustering or dimensionality reduction.
2. **Natural Language Processing (NLP)**, systems where machines are trained to understand, interpret, and respond to human language, as seen in applications like chatbots, virtual assistants, and language translation tools.
3. **Image Sensing and Signal Processing,** in which machines are trained to interpret and make decisions based on visual data from images, videos, or real-world scenes.
4. **Robotics,** in which AI is integrated with machines to automate tasks, often mimicking human actions or performing tasks beyond human capability.
5. **Expert Systems,** in which AI systems rely are designed to simulate the decision-making ability of a human expert to solve specific problems. These systems rely on knowledge engineering approaches that are founded on a knowledge base and a reasoning engine.

Taken together, big data and AI approaches allow for the analysis of large, complex datasets at unprecedented speed and scale, enabling users to uncover empirical insights that were previously hidden or out of reach for traditional methods of inference to process. These approaches, in turn, allow for more accurate prediction and real-time decision-making in an increasingly data-driven world.

Big Data and AI in Sexual and Reproductive Health

The datafication of sexual and reproductive health (SRH) has taken place at great speed in the last decade and has involved the collection of large-scale health data from a variety of sources, including digital platforms, wearables, and health apps to track fertility, contraception use, and sexual health behaviors (Hammond and Moretti 2023). SRH data is also being collected in a variety of contexts via Internet searches, social media, electronic health records, and even online transactions. This shift provides a wide range of alternative sources that can be used to complement traditional SRH datasets, which can serve to facilitate personalized care, public health monitoring, and deeper insights into SRH trends and outcomes across diverse populations.

To date, big data and AI approaches have been applied to SRH research and policy in several ways. Algorithms that use predictive analytics with data from large populations have been developed to forecast fertility trends and risks of conditions like sexually transmitted infections or complications during pregnancy (Stockman et al. 2022; Ridgway et al. 2018; Chung et al. 2021). By analyzing large datasets, family planning providers can more effectively tailor counseling, screen and triage reproductive health conditions (e.g. breast cancer), facilitate assisted reproduction and fertility treatments, and support maternal health programs with minimal client data (Taylor et al. 2023; Raef and Ferdousi 2019; Beim, Elashoff, and Hu-Seliger 2013; Opatunji and Sowunmi 2024; Goueth et al. 2022). Machine learning and other AI approaches have also been used to support public health surveillance strategies, in which data-driven methods facilitate the tracking of SRH indicators and outcomes on a large scale, helping governments to understand the spread of diseases, changes in sexual and reproductive health outcomes in populations (e.g. fertility, unintended pregnancy, birth complications), and the impact of programs and interventions (Wilde et al. 2024; Koivu and Sairanen 2020; Prince 2008). Relatedly, analyses with big data can provide deeper insights into sexual and reproductive behavior, where patterns in social media or mobile app data can reveal trends in contraceptive use and service utilization, SRH risks, and barriers to SRH services (Finnegan et al. 2023; Li et al. 2023).

Big Data in SRH: Advantages

The application of big data and AI to SRH research and practice can confer a number of advantages. In research, big data approaches facilitate the management and analysis of massive datasets that are collected from diverse sources and in multiple formats, providing a more comprehensive understanding of SRH trends. AI tools can process and analyze this data at speeds and scales that were previously not possible, amplifying the researcher's toolkit with new methods to infer relationships that may not have been evident through traditional research approaches (World Health Organization 2024b). This leads to more accurate predictions of conditions and risks at the individual and population levels, which can ultimately guide more informed research outcomes and spur new innovation in SRH.

In practice, AI-driven tools can be used to more effectively personalize SRH care, offering tailored recommendations for meeting contraceptive needs and family planning demand, (in)fertility treatment (Raef and Ferdousi 2019), or pregnancy care based on individual health data (Koivu and Sairanen 2020; Kazantsev, Ponomareva, and Kazantsev 2014; Davidson and Boland 2021). This enables healthcare providers to deliver more precise interventions to improve patient outcomes. AI also has the potential to both mitigate the effects of climate change on SRHR. Climate-induced displacement and natural disasters often disrupt access to SRH services, including contraception, maternal care, and safe abortion services. AI-powered tools can help bridge gaps in SRHR access in climate-affected regions (e.g. through telemedicine or chatbot platforms), but their effectiveness depends on digital access and literacy.

At a population level, big data approaches allow policymakers to track aggregate SRH indicators in real time, which can serve to inform evidence-based decisions and efficiently allocate resources (Mamo et al. 2024; Kebede et al. 2023; Hossain et al. 2022). These approaches also improve access to care, as AI-powered platforms, chatbots, and mobile apps can provide real-time, tailored health information and guidance, particularly in underserved areas (Shankar, Dixit, and Howard 2023; Green et al. 2019). Importantly, the use of AI can also bring together multiple stakeholders who operate in SRH, including individual users, providers and health workers, higher-level implementing partners who facilitate service provision, and policy stakeholders who are responsible for the regulation and management of health systems (World Health Organization 2024b; Vo et al. 2023). Taken together, big data and AI have significant potential to synthesize evidence, enhance data-driven decision-making, optimize resource use, and improve the quality and reach of SRH services.

Big Data in SRH: Challenges

SRH behaviors and decisions are inherently personal matters (World Health Organization 2024b). Additionally, some SRH behaviors and outcomes also carry social and cultural stigma, which increases the need for privacy and confidentiality. For example, the provision of safe abortion services, the use of contraceptives by youth, a person's HIV status, or sexual intercourse by unmarried individuals can be socially and/or culturally taboo in many settings.

Four main challenges arise while applying big data and AI to SRH research and practice. The first challenge is related to ensuring the right to privacy and confidentiality of information of research participants and service users. The risk of data breaches is a big concern as it can not only expose the identity of specific individuals but also potentially sensitive aspects of their behavior, choices, and health conditions. Increasingly, information that is collected during interactions between an individual client and a health care provider is stored in electronic medical records. The advantage of electronic medical records is that they can follow the client across the different points of care to more effectively facilitate a continuity of service. However, this advantage has to be weighed against the cost of potential data breaches and loss of privacy and, in turn, individual autonomy. Health system users may also interact with digital applications, such as chatbots, to educate themselves on specific SRH matters or to seek medical advice, especially if they are reluctant to see a human service provider out of concerns of embarrassment, lack of access, or poor quality of care. Unlike human service providers, chatbots can provide non-judgmental and confidential services, often within the private spaces that the user is able to afford (Mills et al. 2023). Just like electronic medical records, chatbots collect information from the user to tailor information that is appropriate to the user's needs and original queries. This may be in addition to other personal identifiable information (PII) that may be gathered, such as a user's IP address, location, and other demographic information, that a user submits to access the chatbot. Data held by a chatbot therefore have to be securely stored so that they are not accessed by unauthorized third parties, which leads to a loss of confidentiality and privacy.

The second challenge is related to the limited knowledge of health consumers about big data and AI and how they are being deployed in the context of their health care. The lack of orientation about big data and AI often makes consumers ill-equipped to make choices or understand the implications of their choices, thereby limiting their autonomy in care-seeking. The choice for clients to have their SRH care to be provided by a human versus a chatbot (or other AI application), and the disadvantages and disadvantages of each option, is one such example. Another illustration of choice relates to clients having their mammograms being interpreted by an AI application alone, versus a radiologist in addition to an AI application, for diagnosing breast cancer (Taylor et al. 2023). Additionally, SRH consumers may not be aware of the types of data that would be collected, stored, secured, shared, and used in their interactions with AI applications. For example, consumers may be aware of their medical data being collected but may be less aware of other types of PII that are also simultaneously collected, including contact information (e-mail addresses, phone numbers, home addresses), location data (GPS, IP addresses), biometric data (voice prints, facial recognition data, fingerprints for authentication), metadata from the device used to access the application (device type, operating system), online identifiers (usernames, account IDs, cookies to track online activity), online behavioral data (browsing history, search queries, shopping behavior), and financial behavior (payment details, transaction histories), among others. As a result, they may not be aware of how a comprehensive picture of their health, choices, and behaviors is being developed and is being made available to others who might have access to their information. More broadly, consumers of AI especially may not be aware of the risks of loss of privacy should the data be exposed, and experiential learning through trial and error alone is likely not sufficient to educate and prepare individuals and communities for the use of big data and AI in their care. To address these and other related concerns, governments are moving towards implementing more robust data protection, regulation, and privacy laws such as the General Data Protection Regulation (GDPR), which was adopted by the European Union, and India's Digital Personal Data Protection Act (DPDP Act).

The third challenge is centered around the quality of and inherent biases within the data that are used to train and calibrate AI applications and algorithms. If producers of the training data are not aware of the conditions under which the data were collected and curated, then the structural biases that are generated under these conditions can feed into the assumptions of the AI algorithm, leading to misinformed outputs, inaccurate diagnoses, and poor treatment management options. Often, the structural biases that training data may exhibit are the same biases that exist in the socio-economic and contextual environments under which the data were produced. For example, there are differences in preferences, values, and knowledge on various SRH topics by gender, age, place of residence, socioeconomic status, and cultural background. Thus, if an urban, male AI designer were to develop a chatbot to provide menstrual health counseling, then the AI responses may likely reflect the designer's own inherent menstrual health knowledge, preferences, values, and assumptions in the absence of conscientious and proactive bias correction. In other instances, there may be no issues with data quality, but the training data may originate from a social or cultural context that is applied to a vastly different geographical setting or population and may therefore not be relevant for the audience for whom the AI application is intended to serve.

The fourth challenge, which is related to the third, is that there is a potential risk in amplifying myths and misconceptions that arise either from malicious intent by bad actors or from poor-quality training and calibration data. Myths and misconceptions abound around many aspects of SRH. Well known examples include myths and misconceptions about contraception, including how specific methods work or their perceived effects on future fertility and health. Such misconceptions about contraception continue to persist among the youth; for example, a qualitative study of young men and women from one county in Kenya reported that young men were more prone to have erroneous information about contraceptives than young women (Mwaisaka et al. 2020). Other common misconceptions pertain to the transmission of HIV or other sexually transmitted diseases (Sano et al. 2016), menstruation and ovulatory cycles (Zegeye et al. 2022), and abortion (Sorhaindo and Loi 2022), among others. With the proliferation of information sources of varying quality on the public domain, big data applications and AI that scrape these sources have the potential to amplify factually incorrect information. When actively promoted by malicious actors, disinformation can not only cause confusion but can further stigmatize individuals in their reproductive health decisions and utilization of services. For example, anti-abortion websites can actively promote misleading information about abortion, making it difficult for the public to be able to distinguish fact from opinion or propaganda. The amplification of erroneous information can also be exacerbated due to poor quality or unvetted training data. Specifically, if AI applications are trained on a data set which has inherent flaws, then these same errors will continue to be propagated even as new data are added to the database. Training on high-quality, context-specific datasets—ideally developed with domain experts—would improve reliability and relevance.

AI-driven tools are often framed around their ability to personalize experiences by offering tailored health information, improving decision-making efficiency, or predicting individual risk factors. However, the aggregate impact of AI in SRH depends on accessibility, equity in deployment, and systemic adoption. Without careful implementation, AI solutions could inadvertently reinforce existing disparities rather than improve outcomes across diverse populations. Moreover, while AI solutions may function well in controlled settings or pilot programs, real-world scalability remains a challenge. Models trained in one context may not generalize well to others, further limiting scalability.

The Ethics of Big Data and AI in SRH

The ethical concerns arising from the use of big data and AI in SRH are closely related to the challenges noted in the earlier section. These concerns have been noted by ethicists and scientists alike (Blasimme and Vayena 2019; Williamson and Prybutok 2024) and can be grouped into two broad categories. The first category represents an individual and community perspective and relates to how individuals and communities understand, engage, and participate with big data and AI. The second category represents areas that developers and programmers have to consider while building AI applications and related analytical tools.

The first ethical concern is related to equity of access to digital devices and AI-driven SRH services (Richards and King 2014). Although global mobile phone penetration is high and growing, ownership and usage of mobile internet is not similar across the regions of the world or even within subpopulations of a country. As of 2023, an estimated 54 percent of the global population own a smartphone; relatedly, 57 percent of the world's population, or 4.6 billion people, have access to mobile internet (Association 2023). Access to mobile broadband is almost ubiquitous, with 95 percent of the global population living within coverage of a mobile broadband network. In spite of improved access worldwide, many "digital divides" remain; for example, gender gaps in mobile internet use persist in low- and middle-income countries, with women being 15 percent less likely to use mobile internet services compared to men (Association 2024). These data indicate that it is likely that sections of the population will be left behind even as access to digital services expand. These population segments may also be the most vulnerable in other aspects as well, including access to AI-driven SRH services. Ensuring that these groups are not deprived of key SRH services, even as more services continue to become digitized and (exclusively) provided over digital platforms, is a concern that needs to be addressed.

The second ethical concern, which is related to the first one, is about the processes that are needed to obtain informed consent from research participants and health system users before they participate in the research or receive care. These individuals will need to fully understand not only the type of research that they will participate in or care that they are about to receive, but also about how their data will be used, stored, and shared. For example, it will be important to get informed consent from users of health systems if they desire to receive part of their care through an AI application. In a similar fashion, they will also have to know how their data might be shared and used. The example of the AI company Deep Mind analyzing electronic health records of 1.6 million patients of the UK's National Health Service is an illustration of the ethical challenges related to informed consent (Blasimme and Vayena 2019). In this instance, while the patients may have consented to care at the NHS, they had not consented to have their data be used to help Deep Mind create an application and profit from their data. This case highlights the extent to which the NHS patients had been deprived of their autonomy over how their own personal health data is shared and used. Additional ethical concerns are related to data security, including the number of entities that have access to personal data; should there be a breach in the process, then participant/patient data can be revealed with potential privacy loss across multiple systems that may jointly share access. AI systems can be complex, but explaining how it works, including the implications of the use of the data, will be critical for meaningful participation and engagement. The informed consent process has to be thorough and broken down into understandable concepts so that individuals can truly consent after being informed.

The next two ethical concerns are related to the development of the AI application, including the datasets on which they are trained. AI systems can perpetuate and, in some cases, exacerbate biases if the training data were not representative of the populations for whom they are ultimately intended. Such bias and discrimination may lead to inaccurate results and recommendations, with potentially negative consequences for underrepresented and vulnerable groups (Richards and King 2014; Blasimme and Vayena 2019; Williamson and Prybutok 2024). For example, if data that included contraceptive information from high income countries was used to train an AI application to deliver services in low- and middle-income settings, then not only would the recommendations not fit the new context, but they could also be detrimental to the populations being served. For example, the AI application may recommend a particular contraceptive which may be contraindicatory to clients in the new setting. As a first step to address these concerns, the WHO has issued guidance on the importance of high quality, context-appropriate data to ensure that wellbeing, safety, and ethics are assured (World Health Organization 2024a).

The last ethical concern is related to the lack of transparency in AI algorithms and decision-making processes, commonly referred to as the "black box." Developers are sometimes reluctant to reveal the underlying decision-making algorithms if they believe it will reduce their competitiveness in the marketplace. At other times, developers themselves may not be able to identify how the assumptions that undergird the models yield the outcomes that are produced, particularly as algorithms continue to change as the application self-learns and generates new content. Yet, algorithms can be based on fundamentally incorrect assumptions, which in turn can lead to poor quality care for SRH users. Normative bodies such as the WHO and national regulators are increasingly asking developers to adopt ethical principles for the use of AI in health. One such principle that was proposed by the WHO is to "ensure transparency, explainability and intelligibility," The intention of this principle is to ensure that all health stakeholders are able to understand how the AI is designed, deployed, and the conditions for its use (World Health Organization 2024a).

The trade-off between providing some content versus a generic non-response in sensitive exchanges via chatbots or digital assistants is complex. On the one hand, even a partial or imperfect response may be more helpful than no information. On the other, an inaccurate or inappropriately phrased response could cause harm or erode trust. Ethical guidelines should dictate when an AI should engage versus defer to human experts. Relatedly, clear standards and protocols for handling sensitive situations do not currently exist. The development of such guidelines, through collaboration between AI researchers, public health experts, and ethicists, is crucial. These guidelines should define best practices for response thresholds, disclaimers, escalation pathways to human support, and safeguards against misinformation. Moreover, users should be made aware when they are interacting with an AI, what its limitations are, and when it is appropriate to seek human advice. This can mitigate potential harm and ensure informed decision-making.

Conclusions

From a skeptic's point of view, interest around big data and AI is arguably not new and yet another example of (unfounded?) excitement around technological innovations. Skepticism around AI's potential to transform SRH service delivery is also valid, particularly given the history of technological hype cycles in health—most notably and recently, the enthusiasm around m-Health. The experience with m-Health provides important lessons for the field: digital tools like SMS reminders have shown promise in improving adherence and knowledge, but their broader, sustained impact at scale has been more difficult to establish (McCool et al. 2020; Gagnon et al. 2016). However, there are reasons to believe that AI's role in SRH could be distinct from prior waves of technological enthusiasm, particularly when recognizing the impacts of digital tools in other sectors.[2] First, AI's potential for personalization and real-time adaptation, e.g. through chatbots, predictive analytics, and decision-support tools, differentiates it from static m-Health interventions. This is particularly relevant in contexts where social stigma or misinformation act as barriers to care. Second, AI's ability to analyze vast amounts of data—from medical records to social determinants of health—opens possibilities to design more targeted and proactive interventions. AI applications in SRH can therefore be more responsive to queries in real-time, benefitting from improvements in computing power, the availability of large-scale datasets, and advancements in natural language processing (NLP), which were not as developed during the m-Health boom.

More broadly, challenges around individual-level prediction remain. In SRHR, where contextual and sociocultural factors play a significant role, over-reliance on AI-driven predictions without robust validation can lead to incorrect risk assessments or inappropriate recommendations. To navigate these tensions, hybrid models that balance AI-driven personalization with broader public health strategies should be considered. AI should complement—not replace—existing health systems, and its use for both individual and systemic benefits should be continuously evaluated. Additionally, AI deployment should be accompanied by rigorous impact assessments to measure not only efficiency gains but also equity and long-term population health outcomes.

[2] In health, for example, electronic medical records (EMR) and District Health Information Software 2 (DHIS2) open-source software is increasingly being adopted for the collection and management of data in many LMICs. Outside of health, transfers and welfare payments from governments to individuals via mobile phones has increased. India's Aadhaar program is an example of the adoption of biometric identification at scale—Aadhaar provides citizens with a unique ID that combines biographical information with biometric information, which can be used for the identification of those who are most in need of governmental assistance for the efficient distribution of public goods and services (Rao and Nair 2019).

While it is still too early to know how big data and AI will shape SRH research, policy, and practice, the use of AI and data-driven approaches in the field is inevitable, particularly as governments look for more cost-effective ways to deliver high-quality services. In line with the goals of the ICPD PoA, AI and technology offer tremendous potential to improve service delivery, expand access, and optimize resource allocation. To harness this potential, it is crucial that practitioners, academics, communities, and health stakeholders rapidly familiarize themselves with these emerging technologies. Governments also need to take steps to implement data protection laws that align with these advancements, and additional research will be necessary to demonstrate the health impacts of AI and big data in SRH (Nascimento et al. 2016). At the same time, ethical concerns—particularly around informed consent, equity, and transparency—must be addressed, and it is important to remain cautious about the responsible use of AI and ensure that these tools are deployed in ways that protect privacy, promote equity, and avoid exacerbating disparities in healthcare. This can be achieved with robust regulation, effective governance frameworks, well-informed stakeholders, and partnerships between users, developers, and service providers.

References

Association, GSM. 2023. “The State of Mobile Internet Connectivity Report 2023.” London, United Kingdom: GSM Association..

———. 2024. “The Mobile Gender Gap Report 2024.” London, United Kingdom: GSM Association.

Beim, Piraye Yurttas, Michael Elashoff, and Tina T. Hu-Seliger. 2013. “Personalized Reproductive Medicine on the Brink: Progress, Opportunities and Challenges Ahead.” *Reproductive BioMedicine Online* 27 (6): 611–23. https://doi.org/10.1016/j.rbmo.2013.09.010.

Blasimme, Alessandro, and Effy Vayena. 2019. “The Ethics of AI in Biomedical Research, Patient Care and Public Health.” SSRN Scholarly Paper. Rochester, NY. https://doi.org/10.2139/ssrn.3368756.

Borda, Ann, Andreea Molnar, Cristina Neesham, and Patty Kostkova. 2022. “Ethical Issues in AI-Enabled Disease Surveillance: Perspectives from Global Health.” *Applied Sciences* 12 (8): 3890. https://doi.org/10.3390/app12083890.

Cangiano, Alessio. 2024. “Navigating Megatrends: The ICPD Programme of Action for a Sustainable Future.” New York, NY: United Nations Population Fund.

Chung, Esther H., Chaitanya R. Acharya, Benjamin S. Harris, and Kelly S. Acharya. 2021. “Development of a Fertility Risk Calculator to Predict Individualized Chance of Ovarian Failure after Chemotherapy.” *Journal of Assisted Reproduction and Genetics* 38 (11): 3047–55. https://doi.org/10.1007/s10815-021-02311-0.

Davidson, Lena, and Mary Regina Boland. 2021. “Towards Deep Phenotyping Pregnancy: A Systematic Review on Artificial Intelligence and Machine Learning Methods to Improve Pregnancy Outcomes.” *Briefings in Bioinformatics* 22 (5): bbaa369. https://doi.org/10.1093/bib/bbaa369.

Fetzer, James H. 1990. “What Is Artificial Intelligence?” In *Artificial Intelligence: Its Scope and Limits*, edited by James H. Fetzer, 3–27. Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-009-1900-6_1.

Finnegan, Amy, Saisahana Subburaj, Kelly Hunter, Priya Parkash, Elizabeth Shulman, Janel Ramkalawan, and Megan J Huchko. 2023. “Big Data, Machine Learning and Contraceptive Use: A Scoping Review.” *Oxford Open Digital Health* 1 (January):oqad002. https://doi.org/10.1093/oodh/oqad002.

Gagnon, Marie-Pierre, Patrice Ngangue, Julie Payne-Gagnon, and Marie Desmartis. 2016. “M-Health Adoption by Healthcare Professionals: A Systematic Review.” *Journal of the American Medical Informatics Association* 23 (1): 212–20. https://doi.org/10.1093/jamia/ocv052.

Giest, Sarah. 2017. “Big Data for Policymaking: Fad or Fasttrack?” *Policy Sciences* 50 (3): 367–82. https://doi.org/10.1007/s11077-017-9293-1.

Goueth, Rose C., Kristin G. Maki, Ayo Babatunde, Karen B. Eden, and Blair G. Darney. 2022. “Effects of Technology-Based Contraceptive Decision Aids: A Systematic Review and Meta-Analysis.” *American Journal of Obstetrics and Gynecology* 227 (5): 705-713.e9. https://doi.org/10.1016/j.ajog.2022.06.050.

Green, Eric P, Alexandra Whitcomb, Cynthia Kahumbura, Joseph G Rosen, Siddhartha Goyal, Daphine Achieng, and Ben Bellows. 2019. “‘What Is the Best Method of Family Planning for Me?’: A Text Mining Analysis of Messages between Users and Agents of a Digital Health Service in Kenya.” *Gates Open Research* 3 (May):1475. https://doi.org/10.12688/gatesopenres.12999.1.

Hacker, Philipp, and Bilyana Petkova. 2017. “Reining in the Big Promise of Big Data: Transparency, Inequality, and New Regulatory Frontiers.” *Northwestern Journal of Technology and Intellectual Property* 15:1.

Hammond, Natalie, and Angelo Moretti. 2023. "Data like Any Other? Sexual and Reproductive Health, Big Data and the Sustainable Development Goals." *Sexualities*, May, 13634607231152599. https://doi.org/10.1177/13634607231152599.

Höchtl, Johann, Peter Parycek, and Ralph Schöllhammer. 2016. "Big Data in the Policy Cycle: Policy Decision Making in the Digital Era." *Journal of Organizational Computing and Electronic Commerce* 26 (1–2): 147–69. https://doi.org/10.1080/10919392.2015.1125187.

Hossain, Md Ismail, Md Jakaria Habib, Ahmed Abdus Saleh Saleheen, Md Kamruzzaman, Azizur Rahman, Sutopa Roy, Md Amit Hasan, et al. 2022. "Performance Evaluation of Machine Learning Algorithm for Classification of Unintended Pregnancy among Married Women in Bangladesh." *Journal of Healthcare Engineering* 2022 (May):e1460908. https://doi.org/10.1155/2022/1460908.

Hyde, Steven J., Eric Bachura, and Joseph S. Harrison. 2023. "Garbage in, Garbage out: A Theory-Driven Approach to Improve Data Handling in Supervised Machine Learning." In *Methods to Improve Our Field*, edited by Aaron D. Hill, Aaron F. McKenny, Paula O'Kane, and Sotirios Paroutis, 14:101–32. Research Methodology in Strategy and Management. Emerald Publishing Limited. https://doi.org/10.1108/S1479-838720220000014006.

Ishwarappa, and J. Anuradha. 2015. "A Brief Introduction on Big Data 5Vs Characteristics and Hadoop Technology." *Procedia Computer Science*, International Conference on Computer, Communication and Convergence (ICCC 2015), 48 (January):319–24. https://doi.org/10.1016/j.procs.2015.04.188.

Ismail, Azra, and Neha Kumar. 2021. "AI in Global Health: The View from the Front Lines." In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–21. CHI '21. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3411764.3445130.

Kashyap, Ridhi. 2021. "Has Demography Witnessed a Data Revolution? Promises and Pitfalls of a Changing Data Ecosystem." *Population Studies* 75 (sup1): 47–75. https://doi.org/10.1080/00324728.2021.1969031.

Kazantsev, Alexander, Julia Ponomareva, and Pavel Kazantsev. 2014. "Development and Validation of an AI-Enabled mHealth Technology for in-Home Pregnancy Management." In *2014 International Conference on Information Science, Electronics and Electrical Engineering*, 2:927–31. https://doi.org/10.1109/InfoSEEE.2014.6947804.

Kebede, Shimels Derso, Daniel Niguse Mamo, Jibril Bashir Adem, Birhan Ewunu Semagn, and Agmasie Damtew Walle. 2023. "Machine Learning Modeling for Identifying Predictors of Unmet Need for Family Planning among Married/in-Union Women in Ethiopia: Evidence from Performance Monitoring and Accountability (PMA) Survey 2019 Dataset." *PLOS Digital Health* 2 (10): e0000345. https://doi.org/10.1371/journal.pdig.0000345.

Koivu, Aki, and Mikko Sairanen. 2020. "Predicting Risk of Stillbirth and Preterm Pregnancies with Machine Learning." *Health Information Science and Systems* 8 (1): 14. https://doi.org/10.1007/s13755-020-00105-9.

Kumar, Mallick, Pradeep. 2015. *Research Advances in the Integration of Big Data and Smart Computing*. IGI Global.

Li, Xuan, Hanxiyue Zhang, Shuangyu Zhao, and Kun Tang. 2023. "Predicting Risky Sexual Behavior Among College Students Through Machine Learning Approaches: Cross-Sectional Analysis of Individual Data From 1264 Universities in 31 Provinces in China." *JMIR Public Health and Surveillance* 9 (1): e41162. https://doi.org/10.2196/41162.

Magalhães, Jorge Lima de, Luc Quoniam, Zulmira Hartz, Henrique Silveira, and Priscila da Nobrega Rito. 2022. "Knowledge Management in Big Data Times for Global Health: Challenges for Quality in One Health." In *Handbook of Research on Essential Information Approaches to Aiding*

Global Health in the One Health Context, 149–63. IGI Global. https://doi.org/10.4018/978-1-7998-8011-0.ch008.

Magolous, Roger, and Ben Lorica. 2009. "Introduction to Big Data." Sebastopol, CA: O'Reilly Media.

Mamo, Daniel Niguse, Yosef Haile Gebremariam, Jibril Beshir Adem, Shimels Derso Kebede, and Agmasie Damtew Walle. 2024. "Machine Learning to Predict Unintended Pregnancy among Reproductive-Age Women in Ethiopia: Evidence from EDHS 2016." *BMC Women's Health* 24 (1): 57. https://doi.org/10.1186/s12905-024-02893-8.

McCool, Judith, Rosie Dobson, Naomi Muinga, Chris Paton, Claudia Pagliari, Smisha Agawal, Alain Labrique, Helen Tanielu, and Robyn Whittaker. 2020. "Factors Influencing the Sustainability of Digital Health Interventions in Low-Resource Settings: Lessons from Five Countries." *Journal of Global Health* 10 (2): 020396. https://doi.org/10.7189/jogh.10.020396.

Mills, Rhiana, Emily Rose Mangone, Neal Lesh, Diwakar Mohan, and Paula Baraitser. 2023. "Chatbots to Improve Sexual and Reproductive Health: Realist Synthesis." *Journal of Medical Internet Research* 25 (1): e46761. https://doi.org/10.2196/46761.

Mwaisaka, Jefferson, Lianne Gonsalves, Mary Thiongo, Michael Waithaka, Hellen Sidha, Alfred Agwanda, Carol Mukiira, and Peter Gichangi. 2020. "Exploring Contraception Myths and Misconceptions among Young Men and Women in Kwale County, Kenya." *BMC Public Health* 20 (1): 1694. https://doi.org/10.1186/s12889-020-09849-1.

Nascimento, Bruno R., Maria Carmo P. Nunes, Eduardo L. V. Lopes, Vitória M. L. R. Rezende, Taylor Landay, Antonio L. P. Ribeiro, Craig Sable, and Andrea Z. Beaton. 2016. "Rheumatic Heart Disease Echocardiographic Screening: Approaching Practical and Affordable Solutions." *Heart* 102 (9): 658-664 7p. https://doi.org/10.1136/heartjnl-2015-308635.

Opatunji, Florence Omowunmi, and Christiana Olanrewaju Sowunmi. 2024. "Family Planning Services Digitalization: Implications to Nursing." *Nternational Journal of Public Health Pharmacy and Pharmacology* 9 (1): 14–26. https://doi.org/10.37745/ijphpp.2013/vol9n11426.

Prince, Martin. 2008. "Glucocorticosteroids for Primary Biliary Cirrhosis." Edited by Erik Christensen and Christian Gluud. *Cochrane Database of Systematic Reviews*, no. 1 (September). http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=chh&AN=CD003778&site=ehost-live&scope=site.

Raef, Behnaz, and Reza Ferdousi. 2019. "A Review of Machine Learning Approaches in Assisted Reproductive Technologies." *Acta Informatica Medica* 27 (3): 205–11. https://doi.org/10.5455/aim.2019.27.205-211.

Rao, Ursula, and Vijayanka Nair. 2019. "Aadhaar: Governing with Biometrics." *South Asia: Journal of South Asian Studies* 42 (3): 469–81. https://doi.org/10.1080/00856401.2019.1595343.

Richards, Neil M., and Jonathan H. King. 2014. "Big Data Ethics." *Wake Forest Law Review* 49:393.

Ridgway, Jessica P., Ellen A. Almirol, Alvie Bender, Andrew Richardson, Jessica Schmitt, Eleanor Friedman, Nicola Lancki, et al. 2018. "Which Patients in the Emergency Department Should Receive Preexposure Prophylaxis? Implementation of a Predictive Analytics Approach." *AIDS Patient Care and STDs* 32 (5): 202–7. https://doi.org/10.1089/apc.2018.0011.

Sagiroglu, Seref, and Duygu Sinanc. 2013. "Big Data: A Review." In *2013 International Conference on Collaboration Technologies and Systems (CTS)*, 42–47. https://doi.org/10.1109/CTS.2013.6567202.

Salganik, Matthew J. 2019. *Bit by Bit: Social Research in the Digital Age*. Princeton University Press.

Sano, Yujiro, Roger Antabe, Kilian Nasung Atuoye, Lucia Kafui Hussey, Jason Bayne, Sylvester Zackaria Galaa, Paul Mkandawire, and Isaac Luginaah. 2016. "Persistent Misconceptions about HIV Transmission among Males and Females in Malawi." *BMC International Health and Human Rights* 16 (1): 16. https://doi.org/10.1186/s12914-016-0089-8.

Shankar, Lalita, Anvita Dixit, and Susan Howard. 2023. "A Design-Led Theory of Change for a Mobile Game App (Go Nisha Go) for Adolescent Girls in India: Multimix Methodology Study." *JMIR Formative Research* 7 (1): e43085. https://doi.org/10.2196/43085.

Sorhaindo, Annik, and Ulrika Rehnstrom Loi. 2022. "Interventions to Reduce Stigma Related to Contraception and Abortion: A Scoping Review." *BMJ Open* 12 (11): e063870. https://doi.org/10.1136/bmjopen-2022-063870.

Stockman, Jeni, Jonathan Friedman, Johnna Sundberg, Emily Harris, and Lauren Bailey. 2022. "Predictive Analytics Using Machine Learning to Identify ART Clients at Health System Level at Greatest Risk of Treatment Interruption in Mozambique and Nigeria." *JAIDS Journal of Acquired Immune Deficiency Syndromes* 90 (2): 154. https://doi.org/10.1097/QAI.0000000000002947.

Sullivan, Emily. 2022. "How Values Shape the Machine Learning Opacity Problem." In *Scientific Understanding and Representation*. Routledge.

Taylor, Clayton R., Natasha Monga, Candise Johnson, Jeffrey R. Hawley, and Mitva Patel. 2023. "Artificial Intelligence Applications in Breast Imaging: Current Status and Future Directions." *Diagnostics* 13 (12): 2041. https://doi.org/10.3390/diagnostics13122041.

UNFPA. 2014. "Programme of Action, Adopted at the International Conference on Population and Development." Cairo, Egypt: United Nations Population Fund.

Vo, Vinh, Gang Chen, Yves Saint James Aquino, Stacy M. Carter, Quynh Nga Do, and Maame Esi Woode. 2023. "Multi-Stakeholder Preferences for the Use of Artificial Intelligence in Healthcare: A Systematic Review and Thematic Analysis." *Social Science & Medicine* 338 (December):116357. https://doi.org/10.1016/j.socscimed.2023.116357.

Wahl, Brian, Aline Cossy-Gantner, Stefan Germann, and Nina R Schwalbe. 2018. "Artificial Intelligence (AI) and Global Health: How Can AI Contribute to Health in Resource-Poor Settings?" *BMJ Global Health* 3 (4): e000798. https://doi.org/10.1136/bmjgh-2018-000798.

Wilde, Joshua, Wei Chen, Sophie Lohmann, and Jasmin Abdel Ghany. 2024. "Digital Trace Data and Demographic Forecasting: How Well Did Google Predict the US COVID-19 Baby Bust?" *Population and Development Review* 50 (S1): 421–46. https://doi.org/10.1111/padr.12647.

Williamson, Steven M., and Victor Prybutok. 2024. "Balancing Privacy and Progress: A Review of Privacy Challenges, Systemic Oversight, and Patient Perceptions in AI-Driven Healthcare." *Applied Sciences* 14 (2): 675. https://doi.org/10.3390/app14020675.

World Health Organization. 2024a. "Ethics and Governance of Artificial Intelligence for Health. Guidance on Large Multi-Modal Models." Geneva, Switzerland: World Health Organization.

———. 2024b. "The Role of Artificial Intelligence in Sexual and Reproductive Health and Rights." Geneva, Switzerland: World Health Organization. https://www.who.int/publications-detail-redirect/9789240090705.

Yaqoob, Ibrar, Ibrahim Abaker Targio Hashem, Abdullah Gani, Salimah Mokhtar, Ejaz Ahmed, Nor Badrul Anuar, and Athanasios V. Vasilakos. 2016. "Big Data: From Beginning to Future." *International Journal of Information Management* 36 (6, Part B): 1231–47. https://doi.org/10.1016/j.ijinfomgt.2016.07.009.

Zegeye, Betregiorgis, Nicholas Kofi Adjei, Dina Idriss-Wheeler, and Sanni Yaya. 2022. "Individual and Community-Level Determinants of Knowledge of Ovulatory Cycle among Women of Reproductive Age in 29 African Countries: A Multilevel Analysis." *BMC Women's Health* 22 (1): 394. https://doi.org/10.1186/s12905-022-01984-8.