IRIS

IMPLEMENTATION RESEARCH AND INNOVATION SUPPORT

April 2023
IRIS Learning Brief 3

How might chatbots support reproductive health?

Findings from three new studies

IRIS

Background

A chatbot is a computer program that simulates human conversations.

Chatbots vary in their complexity from simple programmes where users engage through a pre-determined decision-tree of questions and answers to sophisticated software that uses natural language processing and artificial intelligence to understand free text questions and generate new responses.

Chatbots are already widely used in the customer service industry and are increasingly used in health and education. However, their use in the field of sexual and reproductive health (SRH) is relatively new.

iven their application in other fields, chatbots could improve SRH by:

- providing rapid and responsive information and counselling in conversational format
- offering links to high-quality products and services

- serving as a companion on a reproductive health journey
- developing user agency, self-efficacy, and new skills
- supporting service delivery by automating some tasks e.g. collecting routine data

There is a small (around 40 papers), but expanding, evidence base within healthcare on chatbot design, user experience, and the outcomes of chatbot use. This learning brief summarises the key findings of three studies that applied different methods to summarise what is currently known about chatbots that are used in the field of SRH.

We have used this evidence to offer recommendations on future innovation and development of chatbots to improve SRH.

Methods

We used three complementary methods to bring together information on chatbots for SRH:

A systematic review on chatbots for contraception

This review aimed to systematically collate and interpret evidence to determine whether and how chatbots improve contraceptive knowledge, attitudes and behaviors. The secondary aim of the review was to identify and summarize best practice recommendations for chatbot development to improve contraceptive outcomes including the cost effectiveness of chatbots where evidence is available.

To complete this review, we systematically searched for English-language, peer-reviewed, original research, and grey literature published between January 2010 and September 2022. We identified papers that reported on key outcomes including improved contraceptive knowledge, access to contraceptive information and use of services, contraceptive uptake, contraceptive continuation, and communication or negotiation skills for contraceptive use.

Chatbot landscaping and stakeholder interviews

The purpose of this study was to identify and describe existing SRH chatbots, regardless whether they had been evaluated, and share key insights and best practices from chatbot developers, evaluators, and implementers. We identified these chatbots through publications and implementer networks. We interviewed stakeholders who had been involved in developing or evaluating chatbots about the challenges, opportunities, and decisions made during chatbot development.

Realist review of how and why chatbots might improve SRH

The purpose of this study was to identify assumptions made about the potential value of chatbots for SRH and collate the evidence to support or challenge each assumption.

We started with evidence-based guidance on best practice in SRH services which informed a framework for how chatbots might support this best practice. We then tested these assumptions by identifying evidence to support and challenge each assumption. We used this validated framework to generate evidence-informed statements about how chatbots might support SRH services and new thinking on future developments for chatbots in SRH.

What we found:

Systematic review on chatbots for contraception: The Evidence

There is limited rigorous evidence on whether chatbots improve contraceptive knowledge, attitudes, and behaviours. This review identified 8 original research papers and 7 gray literature papers. Seven of the chatbots evaluated were operating in high income countries, UK (2), USA (4) and Japan (1). Eight chatbots were operating in low- or middle-income countries, Kenya (3), South Africa (1), Uganda (1), Nigeria (1), Bangladesh (1) and India (1). One chatbot had international reach.

The evidence on whether chatbots support contraceptive use is conflicting. One RCT suggests no effect on intention to use contraception (Maeda et al., 2020), a small cohort study suggests increased uptake of contraception in adolescent girls (Chernick et al., 2021) and a development report suggests no impact on access to services (PSI, 2019). There is poor quality evidence to suggest increased contraceptive knowledge from interacting with chatbot content, as measured by in-chatbot quizzes (Handforth & Bertteman, 2018).

The chatbots included in this review showed limited user reach and engagement that varied from over 100,000 users in 5-months (Wang et al., 2022), to under 10,000 users in 6-months (Winskell, 2022).

There is high variability in user feedback on the acceptability, convenience, and accessibility of chatbots. Chatbots to improve contraceptive outcomes were perceived as acceptable, convenient, and private by some users. However, other users experienced chatbots as incompetent, impersonal, unsympathetic, reported technical difficulties and found interaction burdensome (Maeda et al., 2020; Nadarzynski et al., 2018, Brannock, 2019; Rahman et al., 2021; Chernick et al., 2021). Many users appreciated the anonymity of talking to a chatbot (Brannock, 2019; Nadarzynski et al., 2021; Rahman et al., 2021) but other users had concerns about the privacy of the data captured by the chatbot (Nadarzynski et al., 2021).

There was no evidence on the costs or cost-effectiveness of chatbots for impacting contraceptive knowledge, attitudes, and behaviours.

Systematic review on chatbots for contraception: Best practices in designing and developing chatbots

From the systematic review we identified best practice recommendations. We included a recommendation where it met two criteria:

1. it came from more than one of the studies identified on chatbots for SRH and

2. it was consistent with the broader literature on chatbots.

Recommendations for chatbot design:

Chatbots should be designed and developed with input from their target audience (Bonnevie et al., 2021; Handforth & Bertermann, 2018; Winskell, 2022; Work&Co, 2022). Users need to be at the centre of chatbot design and development work. Developers need to identify and respond to different user needs, interests, and realities, including whether a chatbot is a channel with which users can engage. For example, one study found that a chatbot that had been developed for a primarily female audience was being used almost exclusively (93%) by self-reported male users (Wang et al., 2022). The authors attributed low female engagement to the high levels of stigma and shame experienced by young women in this context around engaging with SRH content.

A chatbot's user interface should be intuitive and easy to use. To encourage engagement with a chatbot, it should be developed on commonly used devices and platforms. To avoid drop-offs, the user experience needs to be made frictionless and seamless, and users should not be required to go through any superfluous steps (Handforth & Bertermann, 2018; Winskell, 2022).

User privacy should be ensured and visible while engaging with a chatbot. Despite the anonymity offered by chatbots, concerns with data privacy may prevent people from engaging with chatbots. Being transparent about how the captured data will be used will assuage these concerns and increase trust in the chatbot (Handforth & Bertermann, 2018; Nadarzynski et al., 2021).

The persona of a chatbot should reflect the type of relationship desired between chatbot and user. For example, "peer persona" chatbots that offer informal conversations using slang and emojis and that build relationships may focus on shifting shared social norms, whereas "healthcare professional persona" chatbots might be more formal and offer trusted health information about contraceptives. (Bonnevie et al., 2021; Handforth & Bertermann, 2018; Lok Woo et al., 2020; PSI, 2022; Work&Co, 2022).

Chatbot designers should recognize that using a chatbot for contraceptive information is a new behaviour and provide additional support for how to engage with a chatbot. When adopting any new health technology, users need instructions and guidance on how to use chatbots and why to use them (Nadarzynski et al., 2021; Winskell, 2022).

Deployment:

Where possible, chatbot developers should associate chatbots with a reputable and well-known entity (organization, influencer) to build user trust. Evidence suggested that users experienced chatbots as trustworthy when the chatbots seemed to give accurate answers to their questions and when the chatbots were associated with a reputable organization (Brannock, 2019; Winskell, 2022).

Chatbots should link users to human support when needed. Users, particularly young people exploring sensitive and important topics, may need more support than an automated chatbot can offer. For example, if a young person reported an abusive sexual relationship then further intervention may be needed to protect them. Solutions could include human moderation of chats, links to human support triggered by key words, or information about local organisations that might help (Handforth & Bertermann, 2018; Wang et al., 2022).

The chatbot landscaping and stakeholder interviews.

There is a growing number of chatbots that support SRH. We identified and described 35 chatbots globally that aimed to improve SRH. About two thirds (21) of the chatbots identified operate in low-or-middle income countries (LMICs) with 3 operating across multiple LMICs, 11 in Asia and 7 in Africa. Most (19) aimed to engage young people under 25 years. To collect stakeholder feedback, we interviewed 11 chatbot developers and one evaluator to collate their insights.

Most chatbots were designed to be transactional, rather than relational. Most chatbots were designed to do a specific task, such as provide information on contraceptive methods, quickly and effectively. However, a small number aimed to develop an ongoing relationship or a conversation to influence behaviour over time.

Most chatbots in SRH were designed for an audience within one country, with a small number aiming to reach narrowly defined populations with tailored content and language. There were few chatbots working across multiple countries with small changes to enable localisation.

Most chatbots were developed for young people and covered a broad range of SRH topics. The majority of the chatbots signposted to SRH services for users who required additional services. For example, the chatbot might provide information on contraception and then signpost to online and offline locations where contraception could be obtained. The information offered by chatbots was more likely to cover many topics with limited detail rather than a single topic in depth.

Current SRH chatbots are limited in their ability to mimic natural human conversation. SRH chatbots allow users to engage through limited free text and/or choose from a menu of questions. Chatbots for SRH are mostly driven by simple decision-tree algorithms rather than machine learning and this means that the responses that they generate seem more formulaic or "less human". For example, menu-based chatbots can answer a limited number of questions and do not interpret questions asked as free text. Investment in highly sophisticated chatbots in other fields is not present in SRH.

The most common SRH chatbot persona was a knowledgeable, relatable young woman. We found no chatbot personas that were male, but some were non-human, gender neutral robots.

Chatbot developer and evaluator perspectives

Stakeholders agreed that the primary purpose of most chatbots was to serve as an information source, providing accurate and digestible information. Chatbots can shortcut some of the screening work that is needed when searching for SRH information through a web browser, which can produce content that is inaccurate, inappropriate, or irrelevant.

There is a tension between localisation and reach. There are trade-offs between chatbots that are developed to feel familiar to specific populations and those designed to reach large audiences. The varied and changing preferences of potential users means that it is difficult to please everyone, for example, humour and popular culture references may appeal to some users while others may be put off by the lack of seriousness. Language is an important way to build relationships. SRH chatbots should employ positive, colloquial, friendly, and professional language, with attention to consistency. This tension may change as more sophisticated chatbots are able to present the same information in different formats and voices.

Chatbot developers and evaluators in some contexts have struggled to identify safe, non-judgemental, appropriate referral options where chatbots can signpost. Where chatbots provide information and support self-management their value is limited when they cannot link to non-judgemental and reliable sources of additional products and services.

The realist review of how and why chatbots might improve SRH

Using the Centers for Disease Control (CDC) guidance on service delivery for SRH we identified five opportunities for chatbots to contribute to SRH care and we searched for evidence to support or refute these assumptions. The CDC is an internationally recognized science-based and data-driven organization that has a high level of consistency with WHO, UK and European recommendations making these findings widely applicable.

1. Seeking SRH information and services is often stigmatised and chatbots could contribute to the provision of anonymous and non-judgemental services

2. Making SRH decisions requires large amounts of complex information and chatbots could deliver complex information in a conversational format and in digestible 'chunks'.

3. SRH decisions are often made in collaboration with others and chatbots could model safe and appropriate conversations. Conversations started with chatbots may be shared within online and offline conversations with friends and family. This may be when people forward or cut and paste chatbot conversations or when they discuss them with family or friends.

4. Some SRH decisions are time sensitive and chatbots could provide 24/7 responses.

5. People use SRH products over many years, often with limited support from health care professionals. Chatbots could provide timely and continuous support across a woman's reproductive journey.

We found evidence to support all of these assumptions with some qualifications and additions:

Evidence suggests that people are more likely to disclose sensitive information to chatbots that are perceived to be anonymous and that engaging with a chatbot may be less embarrassing, less stigmatizing, and more private than other SRH services. However, for those facing the significant stigma or who lack of access to a private digital device, using a chatbot may still be difficult.

There is evidence to suggest that the conversational structure that chatbots use, delivers information that is digestible, engaging, and accessible and requires less work than use of a search engine.

Chatbots may integrate with online and offline social networks, as chatbots mimic discussions with family and friends and model possible responses within these conversations. Chatbot users may share content with family or friends as part of online or offline conversations, for example sharing messages from chatbots in social media interactions or discussing chatbot content within face-to-face conversations.

The constant availability of chatbots means that they can provide information and signpost when it is needed and may encourage users to return to the chatbot to learn additional information and work towards goals.

There is evidence to suggest that chatbots should not be developed as stand-alone interventions but should be embedded in local systems of service delivery so that people can move smoothly from the chatbot to traditional services and back again as they need different types of support. Chatbots developed for SRH should acknowledge and build on their capacity to support social interactions online and offline, for example by generating sharable content and supporting conversations among several people.

There is a lack of service chatbots in the SRH space whereas they are widely used in other areas of health care, particularly mental health care in high income countries where they deliver automated counselling and support.

Recommendations for future chatbot innovation and research

The overarching finding of this body of work is that while there has been growing chatbot development in the field of SRH, there is limited evidence of what approaches work, why, for whom, and in what context. This presents an opportunity for increased investment in generating evidence around chatbot approaches, innovation, and outcomes. We present a few key recommendations for future chatbot innovation and research here.

Conduct research to identify the marginal benefits, costs, and potential limitations of chatbot interventions to improve SRH outcomes. It is necessary to generate evidence of whether and how chatbots can effectively impact SRH knowledge, attitudes, behaviors, and other priority outcomes. These fundamental questions have not yet been satisfactorily answered, with most outcomes to date focused on basic user engagement metrics. There are early, promising indications that chatbots can engage with users in dynamic ways that have an impact advantage over other technologies and channels, but more evidence is needed including who is left out. It is also imperative to explore the cost-effectiveness of chatbots as decision-makers need to make informed trade-offs in how they invest limited programmatic dollars.

Disseminate best practices and tools for SRH chatbot development, user engagement, and evaluation. The current field is fragmented, with many SRH organizations building their own bespoke chatbots without the benefit of learning about how and why other chatbots have succeeded or failed. Building on the evidence of what works proposed in the first recommendation, it would be helpful to establish an open-source repository of evidence-based and theory-informed workflows, code, theories of change, persona descriptions, and other supportive materials for SRH chatbot development that could be adopted and adapted in various contexts to enhance impact. Additionally, it will be important to share best practices for evaluating SRH chatbots that can maintain user privacy and safety but that can also generate meaningful insights that can improve the chatbot experience.

Develop and test chatbots for SRH that use machine learning and/or voice technology. Chatbot technology is rapidly evolving and it is important to consider the implications and opportunities for chatbots to offer more personalised "human" answers to complex questions. Additionally, investing in large language voice models in non-English languages can begin to address some of the access, financial, literacy, and language barriers inherent in traditional chatbots. The intersection of AI, NLP, and voice technology may expand who can engage with chatbots, thereby increasing equity. More sophisticated chatbots may also adapt content and tone of voice for different audiences or in response to user preference, thereby addressing some of the current tensions between localization and reach.

Develop and test the ability of chatbots to provide SRH support over the user's reproductive health journey. While many chatbots currently provide "session-based" engagement with users, the technology has the potential to provide longitudinal person-centered care. Additionally, integrating improved data insights could support a more predictive and proactive engagement with users than the more reactive approach used by most chatbots. This could allow chatbots to better accompany and support users as their needs and priorities change. This would require further development of features that support longer term relationships between humans and chatbots.

Develop and test chatbots that anticipate use within wider social networks. Chatbots have traditionally been developed for a "one-on-one" user experience. While this is a useful interface for delivering key information and tailoring to one person's reproductive journey, there may be value in allowing multiple users (a woman's/girl's partner, friends, or other influencers) into the same chat to facilitate a moderated session on partner communication, social norms around SRH, and other topics that might benefit from the inclusion of more than one user.

Improve the linkages between chatbots and the local ecosystem of SRH services. While chatbots can be a useful tool in a reproductive journey, they can't deliver services or intervene when there is an immediate crisis. It is important to embed chatbots in local service delivery ecosystems so that people can move seamlessly between chatbot and human help as required. This will require careful thought about how privacy and safety is ensured and how data is shared.

Test the ability of chatbots to deliver priority SRH outcomes at scale. The evidence to date suggests that chatbots have been relatively limited in their reach compared to other mass and social media channels. Additional research is needed to understand the implications, impact, and cost of deploying chatbots at scale.

References

Bonnevie, E., Lloyd, T D., Rosenberg, S D., Williams, K., Goldbarg, J. & Smyser, J (2021). Layla's Got You: Developing a tailored contraception chatbot for Black and Hispanic young women. *Health Education Journal*, 80(4), 413–424. https://doi.org/https://dx.doi.org/10.1177/0017896920981122

Brannock, M. (n.d.). *Investigating Attitudes and Preferences Towards a Chatbot Pregnancy Guide Within Facebook's Social Media Platform Amongst Pregnant Women in Kenya*. Master's thesis, Duke University. Retrieved from https://hdl.handle.net/10161/18868.

Chernick, L. S., Stockwell, M. S., Gonzalez, A., Mitchell, J., Ehrhardt, A., Bakken, S., Westhoff, C. L., Santelli, J., & Dayan, P. S. (2021). A User-Informed, Theory-Based Pregnancy Prevention Intervention for Adolescents in the Emergency Department: A Prospective Cohort Study. *The Journal of Adolescent Health : Official Publication of the Society for Adolescent Medicine*, 68(4), 705–712. https://doi.org/10.1016/J.JADOHEALTH.2020.07.020

Handforth, C., & Bertermann, K. (2018). *How Girl Effect built a chatbot.* https://doi.org/10.13140/RG.2.2.26155.90404

Lok Woo, W., Gao, B., Rafi Omar Al-Nima, R. & Ling, W. K. (2020). Development of Conversational Artificial Intelligence for Pandemic Healthcare Query Support. *International Journal of Automation, Artificial Intelligence and Machine Learning*, 1(1), 54–79. https://researchlakejournals.com/index.php/AAIML/article/view/35

Maeda, E., Miyata, A., Boivin, J., Nomura, K., Kumazawa, Y., Shirasawa, H., Saito, H., & Terada, Y. (2020). Promoting fertility awareness and preconception health using a chatbot: A randomized controlled trial. *Human Reproduction*, 35(SUPPL 1). http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emexb&NEWS=N&AN=637628961

Nadarzynski, T., Puentes, V., Pawlak, I., Mendes, T., Montgomery, I., Bayley, J., & Ridge, D. (2021). Barriers and facilitators to engagement with artificial intelligence (AI)-based chatbots for sexual and reproductive health advice: a qualitative analysis. *Sexual Health*, 18(5), 385–393. https://doi.org/https://dx.doi.org/10.1071/SH21123

Nadarzynski, T., Bayley, J., Llewellyn, C., Graham, C., & Kidsley, S. The acceptability of artificial intelligence (AI)-led chatbot, an automated advice system for sexual health. HIV Medicine. 2018;19(2); S109. Doi: 10.1177/2055207619871808

Population Services International (PSI). (2022). *Pilot Story: Will access to sex-positive and reproductive health information through a chatbot lead to increased contraceptive use amongst Kenyan youth? - Frontier Technologies Hub - Medium*. https://medium.com/frontier-technologies-hub/pilot-story-will-access-to-sex-positive-and-reproductive-health-information-through-a-chatbot-d41738947d0c

Rahman, R., Rahman, M. R., & Tripto, N. I. (2021). Adolescentbot: Understanding opportunities for chatbots in combating adolescent sexual and reproductive health problems in bangladesh. C*onference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3411764.3445694

Wang, H., Gupta, S., Singhal, A., Muttreja, P., Singh, S., Sharma, P., & Piterova, A. (2022). An Artificial Intelligence Chatbot for Young People's Sexual and Reproductive Health in India (SnehAI): Instrumental Case Study. *Journal of Medical Internet Research*, 24(1), e29969. https://doi.org/https://dx.doi.org/10.2196/29969

Winskell, A. (Accessed 2022). *Building chatbots for health content? This is for you.* : PSI. https://www.psi.org/2021/12/building-chatbots-for-health-content-this-is-for-you/

Work&Co. (Accessed 2022). *Planned Parenthood Case Study* | Work & Co. https://work.co/clients/planned-parenthood/

About IRIS

The Implementation Research and Innovation Support (IRIS) for family planning project is an innovative facility supporting the Bill & Melinda Gates Foundation to generate and share actionable evidence on effective, efficient, scalable supply and demand interventions in family planning. IRIS provides technical assistance to support research, evaluation and learning dissemination associated with the foundation's family planning projects.

For more information, please visit our webpage