**Introductory information**

Health data gathering in India’s multistakeholder setting with public participants, medical institutions, community health workers, and nonprofit healthcare providers is manual, time-consuming, and low-fidelity, resulting in limited success tapping into the promise of big data ([Chawla, 2023](https://pubmed.ncbi.nlm.nih.gov/37621782/); [Das, 2021](https://pmc.ncbi.nlm.nih.gov/articles/PMC8725124/)). Despite well-intentioned state-wide program development, peer reviewed research and our year-long engagement including 60+ stakeholder interviews in Jalgaon (Maharashtra) indicate this is a result of lack of data and transparency regarding local maternal health concerns including high-risk pregnancies, home deliveries, past failures in systemic support, and future requirements for public health improvement. Since 2023, we have firsthand knowledge of the limitations faced by district-wide nonprofits we work with like [Aadhar Bahuudeshiya Sanstha](https://aadharbsanstha.org/) (25+ years in Jalgaon having served 40,000+ women and children) in mobilizing public health resources from the state; limitations they notably do not face in combating child trafficking due to their collaboration with law enforcement via Just Rights for Children (JRC), a civil society network, who have created a central repository of thousands of documented child trafficking cases with a consortium of 250 nonprofits nationally. Our partner, SimPPL, is engaging with JRC informally.

**Proposal:** We envision a JRC-like effort to create a standardized, central repository of health metrics starting small, with metrics we gather from conversational and medical report data from ~5000 consenting users in a 2 year pilot that leverages our pilot-tested, free digital literacy platform offering medically-verified clinical information over WhatsApp, called **Sakhi**. Our novel three-point platform, Sakhi, incubated with MIT and now UNICEF, offers an opportunity to reinvent multistakeholder data gathering for nonprofits running MCH programs, and is informed by medical professionals affiliated with the Indian Council of Medical Research partnering with us so we can ensure successful outcomes for NGOs and the state. Our technology already exists, and is field-tested in both India and Bangladesh–instead, these funds are specifically requested to launch an experiment to encourage users that trust Sakhi to upload medical records obtained at antenatal care checkups and other medical visits, and additionally employ conversational data from users to present a real-time, longitudinal picture of their health concerns to both NGOs and state agencies, in a transparent and evidence-based effort towards effective program delivery. **We hypothesize that linking multilingual and multimodal conversational AI data through consumer messaging platforms can source localized concerns impeding program compliance and improve maternal health support.** We test our hypothesis in digitally forward cities like Jalgaon that nevertheless suffer from a lack of adequate measurement and reporting of local health concerns; resulting in lack of experienced medical professionals and limitations in medical supplies available.   
**Partners:** [Sakhi](https://sakhi-health.com) engages with partners at the nonprofit [SimPPL](https://simppl.org/) (Dr. Swapneel Mehta, cofounder), rebuilding digital trust, with experience deploying multilingual tools in Dhaka and Chittagong for [menstrual health literacy](https://sakhi.simppl.org)[[1]](#footnote-0). We partner with the [Indian Council of Medical Research](https://www.icmr.gov.in/) (Dr. Mona Duggal, lead MCH), Aadhar Bahuudeshiya Sanstha (Dr. Bharti Patil, cofounder), and Sakhi has been incubated with UNICEF as a result of our successful pilots with 200+ families in Jalgaon, for maternal health literacy. Potential partners if we determine a necessary expansion to Nashik will include Dr. Praveen Gedam, Divisional Commissioner of Nashik, formerly an NIH grant co-applicant with Dr. Mehta; and current co-grantee with Dr. Duggal. Sakhi is advised by Prof. Nina Mazar, formerly co-founder of the World Bank’s Behavioral Insights unit.  
**Solution:** Our solution for health data collection is [Sakhi](https://sakhi-health.com), a digital health literacy platform, through which we expect users–expectant mothers–to upload medical information with consent for insights to be parsed and utilised to inform NGO and state MCH programs. The solution is three-point and caters to multiple stakeholders jointly: the first, an AI-powered multilingual chatbot for expectant mothers, integrated with existing maternal health programs to improve access to accurate health information across diverse linguistic groups (Hindi, Marathi, local dialects). Second, Sakhi creates localized knowledge bases of information through input from ASHA workers who address incoming queries from expectant mothers through a mobile-based responsive web application (no installs needed) to generate, translate, edit, and send answers to mothers’ queries using GenAI, in real-time. This enables the expansion of the knowledge base for the chatbot, and simplifies their workload by reducing the time to action on maternal health concerns by up to 7x. Third, nonprofits, who are provided extensive, detailed, and duly anonymized regional health concerns sourced from ASHA workers as well as conversational traces analyzed using bespoke natural language processing pipelines for local data that our team has extensive experience developing.

**Next Steps:** Sakhi is provided as a free product for end users that are expectant mothers, to upload medical (labs, pathology, testing) reports in addition to engaging in conversations with Sakhi, starting with ~1000 families in Jalgaon. **We will launch an 18-month IRB-backed experiment** **with an RCT** randomized at either the block level or urban community level among several local population groups, specifically for expectant mothers. Key metrics to identify include a subset of the indicators for high-risk pregnancies (self-reported, validated with medical reports), symptom-to-diagnosis correlation, anemia prevalence, recurring health issues such as various body aches, discharge, bleeding, complications, sepsis; medical topics and themes identified from conversations (human validated), geographical distribution of concerns, migration patterns for healthcare service access, and nutritional inadequacies are an initial set of identified concerns, though the final set will be population-specific. We will expand coverage to ~4000 mothers in the same region (depending on early outcomes including dialect-based linguistic accuracy of our model, and compliance) or nearby cities where we have ongoing projects. Aadhar Sanstha also runs 6 local OPD clinics allowing us to increase compliance through appropriate interventions in case of low initial compliance with uploading medical information for participating mothers. We conduct monthly stakeholder impact assessments; evaluating the platform's data and insights with ICMR, nonprofit leadership, and frontline workers, to assess the impact on their decision-making. Beyond traditional chatbots, we will utilise text-to-text, text-to-voice, and voice-to-voice input/output options (see last slide in [this deck](https://bit.ly/sakhi-demo)) for multimodal data gathering from low (digital) literacy users. **6 months are reserved for post-assessment of the experiment**; compiling of reports, convening policymakers, and publication.

**Impact:** By combining technological innovation with deep community engagement and strategic partnerships, our team is well placed to drive state-wide impact through our unique platform. We will:

1. Leverage existing NGO networks and government partnerships to rapidly expand our user base.
2. Implement a tiered B2B pricing model to ensure financial sustainability while maintaining free accessibility for mothers.
3. Collaborate with academic institutions and research organizations to publish evidence-based studies on the platform's impact, further validating its effectiveness.
4. Engage with policymakers and health officials to advocate for the integration of AI-powered solutions in national maternal health strategies.

**Budget Narrative**

The proposed budget of $150,000 is strategically allocated to support our innovative maternal health data gathering initiative in India. The largest portion, $65,250 for Other Expenses, will primarily fund the staffing costs development and maintenance of data collection and conversational analysis, including server costs and software licenses. This is crucial for enabling the multilingual chat and data analysis capabilities central to our project. Personnel costs for partners are in the $20,000 allocated for Sub-contracts, which will support local NGO partners like Aadhar Bahuudeshiya Sanstha in implementing on-the-ground programs and data collection efforts. The $13,000 for Capital Assets/Equipment will provide necessary hardware for ASHA workers and NGO staff to access and utilize the Sakhi platform effectively. Travel expenses of $13,750 will facilitate field visits and stakeholder meetings essential for project coordination and impact assessment. $20,000 for Supplies will cover materials needed for community engagement and health education activities. The $18,000 in Indirect Costs will support administrative overhead.

1. this was the 2023 variant of the current Sakhi maternal health platform developed at SimPPL for menstrual health. [↑](#footnote-ref-0)