Improving Clinical Trial Enrollment for Smartphone-Based AI Data Collection: A Methodological Analysis of Nudge-Based Interventions

Amil Khanzada[1]*, Faisal H. Cheema (M.D.)[2,3], Takuji Takemoto[4]

Abstract
The development of impactful medical AI technology hinges on access to high-quality training data, yet insufficient patient enrollment in clinical studies is often a limiting factor. This study examines the methodological challenges we faced in assembling a training dataset from 256,600 PCR-tested patients in clinical studies conducted in Colombia, Japan, Pakistan, and India to build a smartphone app that uses AI analysis of breathing and cough sounds to detect COVID-19. We analyzed the effectiveness and cost implications of three different data collection methods. We found that robocalls were the most efficient, enrolling 247,450 patients at USD 0.04 per call and an 8.5% success rate. Face-to-face methods achieved high enrollment rates (on average 45% to 61%) but were resource-intensive. Flyers were largely ineffective. Based on our findings and leveraging behavioral economics, we propose a framework of targeted nudge strategies to improve patient enrollment in future smartphone-based clinical studies.

JEL Classification: D91; I18; C93; C83; O33

Keywords
nudge theory — behavioral economics — clinical trial enrollment — artificial intelligence — COVID-19 — choice architecture

[1] *Doctoral Student, Headquarters of Regional Revitalization, University of Fukui, Japan*
[2] *George Mason University, Mason Enterprise, Arlington, VA, USA*
[3] *ClinRé, Washington D.C., USA*
[4] *Professor, Headquarters of Regional Revitalization, University of Fukui, Japan*
*__Corresponding author__: kad23802@u-fukui.ac.jp

Introduction

With the rapid advancements in the medical artificial intelligence (AI) field, training data collection and preparation are paramount, as small and inconsistent datasets can lead to underperforming algorithms, causing biases within certain patient groups (Whang et al., 2023). Clinical studies are the primary means to collect patient data for certain medical AI technologies.

However, insufficient patient enrollment is a widely faced problem (Zahren et al., 2021) and the most cited reason for clinical trial termination in the Clinical Trials Database (Desai, 2020), and 80% of trials globally fail to meet their recruitment timelines (Clinical Trials Arena, 2012). Patient recruitment deficiencies across clinical trials significantly impede data acquisition. These structural deficiencies were exacerbated during the COVID-19 pandemic by health system constraints and shifting behavioral responses to uncertainty, including fear and resource-seeking behaviors (Alifano et al., 2020, Baddeley, 2020). This shortage of participants creates a critical bottleneck in generating high-quality data, limiting the development and advancement of AI models in medical applications. The lack of sufficient high-quality data to train robust AI models is a limiting factor for the progress of these medical AI innovations.

Between April 2020 and March 2024, we collected cough samples from 256,600 patients in clinical studies across Colombia, Japan, Pakistan, and India. This resulted in various patented audio signal processing and deep learning algorithms for COVID-19 classification (Khanzada et al., 2021, Tang et al., 2022). Our custom-designed data collection web app and study protocols underwent various revisions as we overcame significant patient enrollment challenges due to cross-continental cultural differences and our inability to offer monetary incentives to patients or clinics.

In this manuscript, we aim to enable the success of future medical AI technology development by proposing nudge-based behavioral economic techniques to improve patient motivation for enrollment in clinical studies. This observational study hypothesizes that three patient recruitment methodologies–in particular, face-to-face, flyer-based, and

robocall approaches–will exhibit different efficacies with respect to enrollment rates, cost-efficiency, and data quality. These are not randomized controlled trials, but rather aim to understand and characterize methodological patterns in the real world to identify novel approaches for improving patient enrollment in clinical trials.

Methods

Our IRB-approved, low-risk human subject observational clinical studies are shown in Table 1.

All procedures involving human subjects were conducted in full compliance with relevant national and institutional ethical guidelines and regulations. Ethical approval was obtained from the Institutional Review Boards (IRBs) of all participating institutions. Specifically, the study protocol for data collection at IDIME S.A. in Colombia was approved, as evidenced by the Certification of Approval dated October 26, 2021. For other participating medical colleges in Pakistan, India, and the leading university in Japan, ethical approval was similarly obtained from their respective IRBs; however, due to non-disclosure agreements, the specific names of these institutions and the detailed dates or reference numbers of their ethical approvals cannot be provided in this manuscript. Informed consent was obtained from all participants before their enrollment in the study. For face-to-face recruitment methods, written informed consent was obtained. For the robocall campaign, verbal consent was obtained via the automated system. The privacy rights of all human subjects were strictly observed, and data were de-identified to the maximum extent possible in accordance with all applicable local data privacy regulations.

In our studies with medical colleges in Karachi, Pakistan and Karnataka, India, we leveraged the Jotform web-based survey platform. Patients were instructed to perform three forced coughs directly into a smartphone held by nursing staff wearing personal protective equipment during Polymerase Chain Reaction (PCR) testing, as displayed in Figure 1. A deep breath preceded each cough to maximize signal strength. Other data collected included demographics, current symptoms, comorbidities, and PCR test results (positive or negative). We used a cross-sectional study design with two-to-six-month convenience sampling and inclusion criteria for unhealthy patients screened for COVID-19 at fever clinics and general testing wards. Study protocols and inclusion criteria were standardized across studies, with translation validated by bilingual speakers (Spanish/Japanese/English), ensuring linguistic, cultural, and conceptual equivalence. All staff involved were trained before the start of the clinical research in a live debriefing session with a descriptive slide deck demonstrating each step of the protocol and a standardized verbal script for patient enrollment. Training was conducted by the global principal investigator (PI) and clinical research coordinators to ensure globally consistent implementation across sites.

We ran our following clinical study at IDIME, a major diagnostics company in Colombia, with 80 sites in 36 municipalities and 5,000 employees who conducted millions of COVID-19 tests during the pandemic. Twelve nurses collected data at five testing sites in two major cities (Bogotá and Cali). Study smartphones were used to input patient demographics, medical history, and symptoms into our custom-built progressive mobile web app, as displayed in Figure 2, mirroring the prior Jotform-based studies in Pakistan and India.

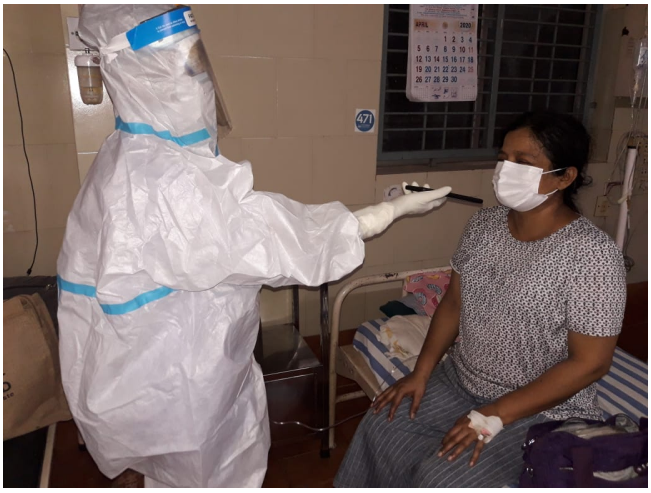

Figure 1. Nurse collecting forced cough sounds from a patient in India

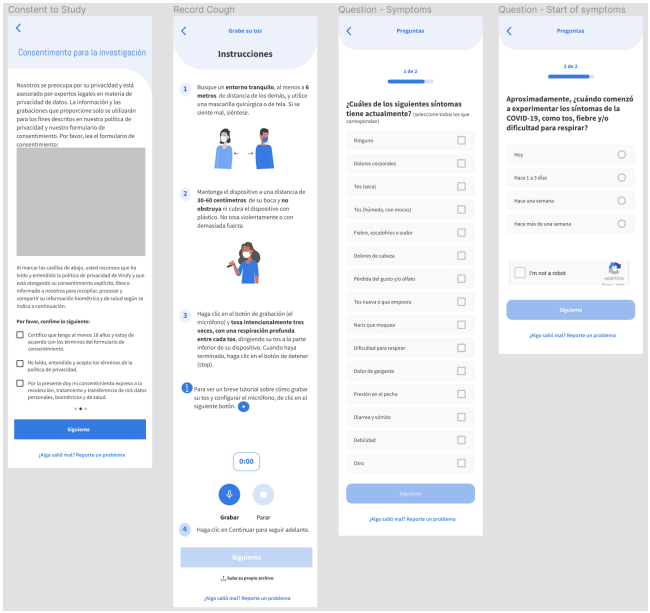

Figure 2. Screenshots of the clinical data collection app

A significant limitation was our inability to motivate patients with monetary incentives (e.g., gift cards) due to local ethical regulations, alongside frequent limits on nurse time to run our study during patient surges. This led us to work with the clinical testing site leadership in Colombia and Japan to

Country	Institution	Period	Patient Enrollment	Enrollment Rate*
Pakistan	Confidential (Medical College)	2020/4 – 2024/6	362 (nurse and doctor face-to-face)	60-90%
India	Confidential (Medical College)	2020/8 – 2020/9	97 (nurse and doctor face-to-face)	80-90%
Colombia	Confidential (IDIME Partner)	2021/8 – 2021/11	247,450 (robocall)	8.5%
Colombia	IDIME S.A. Diagnostics	2021/10 – 2023/1	6,687 (nurse face-to-face)	10-12%
Pakistan	Confidential (Medical College)	2022/7 – 2023/1	2,000 (nurse face-to-face)	30-50%
Japan	Confidential (Leading University)	2023/1 – 2024/3	4 (QR code flyers)	<1%
Total			**256,600**	

Table 1. Patient enrollment counts by institution.
*Total patient enrollment over the study period, estimated using clinic records and enrollment data.

creatively design flyers with QR codes (as shown in Figure 3) to be placed in the reception area, where patients often waited hours before taking their COVID test. We also worked with the Japanese sites to modify their test result notification email to include a link to our app. These interventions allowed patients to access the app independently and participate without direct nurse interaction.

Figure 3. Study enrollment flyers posted at Japanese and Colombian sites

Still facing limitations on patient enrollment, we partnered with a significant Colombian medical insurance institution on a novel approach to conducting a massive outbound phone-calling campaign for millions of their PCR-tested patients. An AI-powered robotic voice conversed with users about the sponsored study by mentioning the insurance company's name, obtained verbal consent, and then instructed patients to perform three forced coughs into their phones.

While we compared cost effectiveness and success rates of various methods of patient recruitment in our cross-sectional observational study, it was not designed to perform significant statistical testing of differences. Future work would require randomized controlled trials to evaluate statistical significance.

Results

Robocalling proved the most effective method for enrolling a large volume of patients, costing only USD 0.04 per call with an 8.5% success rate. To implement this approach, we overcame hurdles to achieve corporate stakeholder approvals and robust compliance with local data privacy regulations through pro bono law firm support as an academic, nonprofit research group addressing a pandemic crisis. Data quality was also a significant challenge, as PCR-tested patients of all socioeconomic and demographic groups across Colombia recorded coughs on a variety of devices, including smartphones, landlines, and feature phones, resulting in significant audio variability due to differing microphones, audio compression techniques, and cell tower quality.

This caused variations in the short-time Fourier transform (STFT) representation of the cough sounds, which resulted in linear dependencies and artifacts within the data. To overcome this challenge, we utilized a novel algorithmic technique we termed "complex clipping," which is a form of signal regularization that is analogous to ReLU activation functions commonly used in deep learning. At its core, complex clipping introduces a non-linearity in the processing of complex-valued STFT data. By mapping the complex values of the STFT to a non-negative real range, it removes linear dependencies and reduces redundancy in the STFT signal representation of coughs by setting a portion of the complex plane to zero. This results in a processed signal that retains the desired spectral characteristics while reducing noise and improving signal quality for subsequent analysis.

This new signal processing methodology enabled us to achieve high performance on this seemingly untrainable data, resulting in positive vs. negative COVID classification at 84% sensitivity, 84% specificity, and 0.93 AUC (Area Under the Curve). The primary novel aspect of this algorithm uses a complex clipping technique to generate accurate predictions even with low audio quality (Atlas et al., 2023). This new algorithm highlights that, with creative signal processing techniques, it is possible to achieve high performance even with data from low-cost high-throughput clinical trials.

Face-to-face data collection was resource-intensive despite high enrollment rates and high data quality, as nurses dedicated 5-15 minutes per patient to obtain informed con-

sent. This resulted in reduced data collection during peak times (e.g., pandemic surges) and occasional data entry errors: 4.72% of samples were discarded due to incomplete patient information or missing test results. A signal-to-noise ratio (SNR) analysis showed that audio files we collected at IDIME at 48 kHz had an average SNR of 30.92 dB, demonstrating high cough signal quality with minimal background noise. The robocall data had a significantly lower average SNR of 23.05 dB at the lower 8 kHz sampling frequency.

Discussion

Although the robocalling campaign was successful for large-scale data collection in Colombia, experiences in other contexts, especially in India and Pakistan, highlighted the methodological trade-offs of the different recruitment methods. Face-to-face recruitment by nursing personnel in India and Pakistan consistently resulted in high enrollment rates (45% to 61%) and high data quality, although labor-intensive, similar to our findings in Colombian sites. The dependency on dedicated nurse time, coupled with ethical restrictions on monetary incentives, restricted scalability during peak seasons when nursing staff were overwhelmed.

On the other hand, passive recruitment approaches, like the use of QR code flyers in Japan, had limited success, with only 4 patients enrolling over a number of months. Patients, uninformed of COVID test results for days, did not remember or prioritize returning to our app, and without financial incentives or instruction from medical staff, they lacked motivation to provide results. Relying so heavily on nursing staff made it difficult to scale the work, especially since ethical rules prevented offering financial incentives to ease the burden. This became particularly clear at the Colombian sites during the Omicron surge (late 2021–early 2022), when nurses were already stretched thin. Staff also noted that many patients were simply exhausted, as some had waited in long lines or traveled for hours by bus (as in Colombia, India, and Pakistan), while others in Japan were trying to squeeze in testing during a short break from work. In those moments, even small additional steps, like talking longer with study staff or trying to make sense of what an “AI clinical study” meant before signing the consent form, felt like too much for many people.

This illustrated the inherent limitation of depending on patient self-motivation alone, without direct intervention, particularly in heterogeneous groups with different degrees of digital literacy and related health issues. Our experiences on these continents highlighted that although certain approaches can achieve high volume (robocalls) or high quality (in-person), each is accompanied by a range of methodological limitations and drawbacks in terms of cost-efficiency, scalability, and patient engagement in real-world clinical practice.

Discussion - Leveraging Nudge Theory

To support future smartphone-based data collection clinical studies, we explore applying behavioral economics, particularly nudge theory, the study of altering behavior through minor and inexpensive interventions popularized in 2008 by Nobel Laureate Richard Thaler and Cass Sunstein (Thaler and Sunstein, 2008, Sunstein, 2018). Our efforts in various study sites taught us that the standard methods of patient recruitment were inadequate in this new context of smartphone-based AI data collection, where standard methods were not sufficient to engage patients in a high enough volume for AI model training. Our study of nudge theory inspired us to create new methodologies based on behavioral economics and patient psychology to address barriers to patient recruitment.

It is also essential to note that, as behavioral economic techniques may be quite powerful, their effects are often context-dependent and subject to inherent limitations, particularly in socially interactive settings where peer influence or environmental factors may alter outcomes (Reijula et al., 2018, Sunstein, 2018). Consequently, investigators must be highly sensitive to avoid coercing, misleading, or manipulating patients and respect patient autonomy, especially amongst underserved populations, such as in our studies (Simkulet, 2019). Our research aligns with the Belmont Report’s principles of respect, beneficence, and justice. Ethical frameworks are critical for nudge implementation, and coercion risks must be thoroughly assessed. Prior work by (VanEpps et al., 2016) also provided meaningful discussion and recommendations for applying nudge theory to clinical trial settings, such as by managing choice architecture.

Building on our findings and the principles of behavioral economics, we propose the following nudges (Table 2) to improve patient enrollment while maintaining dataset quality. While each of these nudges provides its own unique methodological strength and impact, we believe the authority-based nudge (e.g., a physician recommendation) offers the greatest opportunity for improving enrollment, especially for in-person clinical settings where there is direct patient interaction. Default enrollment strategies, such as opt-out, have the highest potential to rapidly enroll many participants, but are best conducted in situations where there is minimal risk and ethical concerns are addressed.

To situate our analysis within the broader behavioral public policy framework, the strategies in Table 2 are best understood as adjustments to the ”choice architecture” of patient enrollment, a concept popularized by Thaler and (Sunstein, 2018). These interventions make participation easier and more appealing while fully preserving patient autonomy. For example, consent form simplification reduces cognitive friction, while norm-based messaging can significantly increase the likelihood that participants actually read and understand these instructions (Bilancini et al., 2020). Additionally, utilizing a

Nudge (Category)	Description	Analysis (Pros/Cons)	Research Support
Enhancing Informed Consent Process: We found that most patients opted out of our studies during the initial step of obtaining consent.			
Recruitment Timing (Friction Reduction)	Enroll patients and obtain consent when patients are most receptive, such as while waiting for treatment. We found most patients declined if approached immediately after completing a PCR test as they rushed to head out.	**Pros:** Reduced burden on staff, increased enrollment rates. **Cons:** More steps and complexity in the protocol.	**Positive:** Waiting room surveys show higher response rates.[1]
Social Norms (Social Norms Nudge)	Frame participation as a standard practice (e.g., "50% of people participate in this research"). This must be done transparently and truthfully as a means of encouraging participation rather than an attempt to manipulate patients with false data.	**Pros:** Patients can decide to enroll with reduced cognitive burden. **Cons:** Ineffective (without deception) if participation is uncommon.	**Positive:** Physicians prescribed less nimodipine when told that most other doctors did not.[2] **Negative:** Enrollment rates did not improve amongst patients told that other peers participated.[3]
Consent Styling (Simplification / Friction Reduction)	Consent forms that are understandable and well-formatted (e.g., using 12-point font or larger, clear section titles, bullet points, and a summary of the key points in lay language) with key points summarized can accelerate patient consent. We found that many patients did not understand our initial lengthy, complex consent forms and took the safer option to opt out.	**Pros:** Improved engagement and understanding with less cognitive burden and increased trust. **Cons:** Risk of oversimplification, misinterpretation, and overstepping legal/ethical regulations.	**Positive:** A Romanian study found few clinical trial participants fully understood their consent forms.[4] Simplifying consent forms, removing overly technical jargon, and clearly outlining risks and benefits can make them easier to understand for all patients, allowing them to make a more informed choice.
Authority Figure (Authority Nudge)	Patients tend to have a higher rate of enrollment when advised by a figure of authority, like a nurse or doctor, and when a study is associated with a trusted institution. In our studies, we found that most patients enrolled when recommended by their nurse face-to-face. When an authority is used to endorse a study, care must be taken to ensure that it is not seen as an effort to coerce or force patient enrollment. The authority figure should be an information source rather than an enforcer for the study.	**Pros:** Increased study trust, enrollment likelihood, and compliance. **Cons:** Risk of coercion and study noncomprehension.	**Positive:** Patients are very likely to consent after a physician recommendation.[4,5]
Enhancing Motivation: Although many patients in our study participated due to altruism, convenience, or authority recommendation, we found a significant disincentive in enrollment was the lack of patient motivation.			
Social Incentives (Gamification)	Patients may be offered digital tokens or avatars which can be shared on social media as a reward for study participation.	**Pros:** Increased peer motivation and potential for viral spread. **Cons:** Administrative overhead, audience-specific design considerations, and enrollment bias (e.g., non-social media users).	**None**
Lottery Payment (Incentive / Gamification)	Invite participants in a raffle for prizes (e.g., gift cards, branded goods). To prevent perceptions of coercion or undue influence, reward value should not be excessive or disproportionate to the effort put forth by the patient. It is also essential to ensure the lottery system is transparent and has no perceived bias. Some may argue that offering incentives can cause an implicit bias by disproportionately targeting patients of a certain socioeconomic status, and certain IRBs and regulatory bodies do not permit it. However, it has been found that a modest incentive can be a good motivator for diverse participants.[6]	**Pros:** High motivation (gamification), cost efficiency, broader participant base (word-of-mouth effect). **Cons:** Uncertainty reduces willingness to participate, ethical and legal concerns, data bias, and administrative complexity, such as for fairness and transparency of the process.	**Positive:** A weekly lottery for $50 improved longitudinal study adherence.[7]
Sharing Research Outcomes (Feedback Nudge)	Provide patients with follow-up information about research outcomes post-publication.	**Pros:** Increased trust and awareness of study significance. **Cons:**Data privacy burden (collection of individually identifiable contact details), administrative overhead.	**None**
Prizes (Incentive-Based Nudge)	Provide small gifts or tokens for participation, e.g., branded stickers or pens.	**Pros:** Inexpensive, improved goodwill towards the study sponsor. **Cons:** Enrollment bias depending on gift type.	**None**
Choice Architecture: The below nudges aim to present a spectrum of enrollment choices in different ways to encourage enrollment.			
Default Choice (Default / Opt-Out Design)	Enroll patients by default, with an opt-out option, based on the principle of status quo bias. Although this can dramatically improve enrollment rates and the diversity of study participants by enrolling those who might not otherwise actively engage in the study, it is not without ethical issues, especially where the opt-out process is not clear, transparent, or apparent. Some patients may feel coerced into participation if they had no part in the decision-making process, which may negatively affect adherence and undermine trust.	**Pros:** Simplified process, higher enrollment, increased subject diversity. **Cons:** Potentially significant ethical and legal concerns.	**Positive:** Opt-out enrollment strategies significantly improved patient enrollment and diversity at US medical universities.[8,9]
Active Choice (Prompted Choice)	Require patients to decide whether to participate in the study before they can receive treatment. The presentation of this choice must be carefully framed to be clear, simple, and not overly coercive or burdensome. A strong call for a decision might also cause anxiety or impact the hospital's reputation.	**Pros:** All patients were made aware of the study. **Cons:** Poor framing may annoy patients, reduce hospital reputation, coerce participation, or violate ethical/legal regulations.	**Positive:** Active choice enrollment increased HIV testing among emergency department patients.[10]
Normative Choice (Norm Activation)	Patients are offered multiple "yes" options for enrollment (e.g., "yes - before test," "yes- after test," and "no"), suggesting that enrollment is a common choice.	**Pros:** May increase enrollment likelihood. **Cons:** Ethical and manipulation concerns.	**None**
Foot-in-the-Door (Commitment)	Patients are initially made a smaller request and then upsold into the full study.	**Pros:** Increased likelihood for eventual agreement. **Cons:** Risk of perceived manipulation.	**Negative:** Patients given a pre-consent survey did not improve enrollment rates.[3]

Table 2. A variety of proposed nudge interventions to enhance patient enrollment

Superscript numbers correspond to sources cited in the table: [1](Ongena and Haan, 2022); [2](Torrente et al., 2020); [3](Krutsinger et al., 2020); [4](Alexa-Stratulat et al., 2018); [5](VanEpps et al., 2016), (Halpern et al., 2021); [6](Halpern et al., 2021); [7](Husain et al., 2019); [8](Pittman et al., 2023); [9](Jia and Linder, 2022); [10](Montoy et al., 2016).

clinician's recommendation employs the authority principle (a highly effective social influence), and opt-out enrollment alters the default choice. Each of the proposed nudges was directly aimed at one of the major behavioral barriers (patient uncertainty, inertia, and cognitive overload) observed in our study.

It is also important to note that, when combined, nudges may have a higher effect, since they likely do not each uniquely affect patient participation but rather interact and synergize with one another. For instance, combining an authority figure nudge with a social norm nudge may compound trust and encourage participation more than either nudge in isolation. Future studies should experiment with a sequential nudge strategy, with a chain of different types of nudges, such as framing participation in surveys as a social norm and mentioning only thereafter the presence of a lottery-based participation award.

Conclusion

Our efforts in assembling a training dataset of 256,600 patients across three continents for developing cough-based COVID-19 diagnostic AI algorithms have revealed significant methodological limitations in current clinical trial enrollment practices for medical AI. While our large-scale robocalling campaign demonstrated the potential to rapidly acquire large datasets, its scalability and data quality remain problematic for many research contexts. Traditional face-to-face data collection, though valuable for data quality and patient engagement, proved prohibitively resource-intensive for large-scale data acquisition. Moreover, passive recruitment methods such as flyers with QR codes were largely ineffective. These findings highlight the urgent need for innovative methodological approaches to enhance clinical trial enrollment for smartphone-based AI studies, particularly in non-pandemic scenarios. Informed by behavioral economics, we propose a framework of methodological interventions, including optimized recruitment timing, strategic use of lottery incentives, default enrollment strategies, leveraging social norms and authority figures, enhanced clarity of consent forms, and research outcome reporting to participants. These methodological strategies are intended to overcome data scarcity challenges, accelerating the development of high-impact medical AI technologies. This study further illuminates the critical limitations of conventional recruitment strategies when applied to diverse patients in real-world settings, providing valuable insights for future clinical studies. Future research must focus on randomized controlled trials to rigorously assess these interventions. These trials should evaluate both individual and combined nudge strategies. Furthermore, future research must tailor nudges to diverse populations and clinical environments, while developing ethical frameworks to ensure patient autonomy and equitable access to clinical research. To ensure practical implementation, seamless integration within existing clinical workflows is needed. Finally, longitudinal studies are required to evaluate the long-term impact of these interventions on data collection, ensuring that statistical significance is appropriately considered. Prioritizing methodologically robust recruitment strategies and considering these essential factors would allow future research to unlock the potential of smartphone-based clinical trials.

Glossary:

AI technology: Artificial Intelligence, is a field of computer science that focuses on creating systems capable of performing tasks that typically require human intelligence.

Training Data: Training data, in the context of machine learning, refers to a dataset used to teach a model how to learn and make predictions. It is a collection of labeled information, such as text, images, video, or audio, that the model uses to identify patterns and relationships.

Robocall: A robocall is a phone call that uses a computerized autodialer to deliver a pre-recorded message, which, in this case, was used for patient recruitment.

Behavioral economics: Behavioral economics is a field that blends psychology and economics to understand how people make decisions in the real world, often deviating from the assumptions of traditional economic models. It explores how emotions, biases, and other psychological factors influence economic behavior.

Acknowledgments

We would like to acknowledge the following individuals and organizations for their invaluable contributions:

- Laura Gomezjurado, Tulio Juarez, Lida Gonzalez, James Molina, Jennifer Soler, Carlos Rojas, and the IDIME Diagnostics team for their superb efforts in running our operations and clinical studies in Colombia.
- Siddhi Hegde, M.D., and Shreya Sreeram, M.D., for their exemplary work designing and running our clinical studies in India.
- Nanako Isobe, Yuka Nishimura, and Keita Oda for their significant efforts in designing and supporting our clinical studies in Japan.
- ClinRé, Inc. and Ryan Yen for their skilled efforts in preparing our clinical protocols and designing our data collection apps.
- Mansoor Ahmed M.D. for his support in designing our clinical studies in Pakistan and the US.
- Kara Meister M.D., Stanford University Assistant Professor, Department Otolaryngology, Division of Pediatric Otolaryngology, and Mary L. Dunne M.D., Stanford University Distinguished Career Institute Fellow, for their kind guidance with respect to clinical study design and the medical implications of our research.

- Sneha Dharmi, Shirin Hasan, and Mahesh Karthikeyan for their copyediting support of this article.

- Les Atlas, University of Washington Electrical & Computer Engineering Professor, and Mert Pilanci, Stanford University Assistant Professor of Electrical Engineering, for their guidance on the AI components of the project.

- Jonatan Jaskilioff, Mauricio Payetta, Andres Barrantes, Yidaido Rojas, Karen Arandia, Fernando Morantes, Diana Garces, XOOR Inc., and Nuvu Inc. for their excellent efforts to create our data collection mobile application and underlying Amazon Web Services infrastructure, along with support for operations in Colombia.

- The Amazon Web Services Diagnostic Development Initiative for providing cloud credits to run our data collection infrastructure.

- Jotform for providing us with a free professional license via their Coronavirus Responder Program.

- The Virufy nonprofit organization for organizing hundreds of volunteers and dozens of partner companies to handle operational aspects of the clinical studies, including app design and AI development.

Declaration of Conflicting Interests

The first author is the founder of an organization that collaborated on the clinical studies described in this manuscript. To minimize potential bias, the second author independently reviewed the study design, data interpretation, and all manuscript drafts. The second author declares no competing interests.

Funding

This research was funded in part by the Biomedical Advanced Research and Development Authority (BARDA), a division of the Administration for Strategic Preparedness and Response (ASPR) within the U.S. Department of Health and Human Services (HHS).

References

Alexa-Stratulat, T., Neagu, M., Neagu, A.-I., Alexa, I. D., and Ioan, B. G. (2018). Consent for participating in clinical trials: Is it really informed? *Developing World Bioethics*, 18(3):299–306.

Alifano, M., Attanasi, G., Iannelli, F., Cherikh, F., and Iannelli, A. (2020). COVID-19 pandemic: a European perspective on health economic policies. *Journal of Behavioral Economics for Policy*, 4(S):35–43.

Atlas, L., Rasmussen, N., Schwock, F., and Pilanci, M. (2023). Complex clipping for improved generalization in machine learning. *arXiv preprint arXiv:2302.13527*.

Baddeley, M. (2020). Hoarding in the age of COVID-19. *Journal of Behavioral Economics for Policy*, 4(S):69–75.

Bilancini, E., Boncinelli, L., Capraro, V., Celadin, T., and Di Paolo, R. (2020). The effect of norm-based messages on reading and understanding COVID-19 pandemic response governmental rules. *Journal of Behavioral Economics for Policy*, 4(S):45–55.

Clinical Trials Arena (2012). Clinical trial delays: America's patient recruitment dilemma. Available at: https://www.clinicaltrialsarena.com/marketdata/featureclinical-trial-patient-recruitment/. Accessed: 2024-05-14.

Desai, C. J. (2020). Recruitment and retention of participants in clinical studies: Critical issues and challenges. *Perspectives in Clinical Research*, 11(2):51–53.

Halpern, S. D., Chowdhury, M., Bayes, B., Cooney, E., Hitsman, B. L., Schnoll, R. A., et al. (2021). Effectiveness and ethics of incentives for research participation: Two randomized clinical trials. *JAMA Internal Medicine*, 181(11):1479–1488.

Husain, S. A., Diaz, K. M., Schwartz, J. E., Parsons, F. E., Burg, M. M., Davidson, K. W., and Kronish, I. M. (2019). Behavioral economics implementation: Regret lottery improves mhealth patient study adherence. *Contemporary Clinical Trials Communications*, 15:100387.

Jia, J. and Linder, J. A. (2022). Default nudges in medicine—Designing the right choice. *JAMA Network Open*, 5(3):e222437.

Khanzada, A., Hegde, S., Sreeram, S., Bower, G., Wang, W., Mediratta, R. P., Meister, K. D., and Rameau, A. (2021). Challenges and opportunities in deploying COVID-19 cough AI systems. *Journal of Voice*, 35(6):811–812.

Krutsinger, D. C., O'Leary, K. L., Ellenberg, S. S., Cotner, C. E., Halpern, S. D., and Courtright, K. R. (2020). A randomized controlled trial of behavioral nudges to improve enrollment in critical care trials. *Annals of the American Thoracic Society*, 17(9):1117–1125.

Montoy, J. C. C., Dow, W. H., and Kaplan, B. C. (2016). Patient choice in opt-in, active choice, and opt-out HIV screening: Randomized clinical trial. *BMJ*, 352:h6895.

Ongena, Y. P. and Haan, M. (2022). Just you wait... and fill out this survey: Discussion of the methodological aspects of waiting room surveys. *Health Services and Outcomes Research Methodology*, 22:508–521.

Pittman, T., Bell, L., Jones, S., Brown, K., Kirchoff, K., and Flume, P. (2023). Enhancing study recruitment through implementation of an opt-out, cold contact process with consideration for autonomy, beneficence and

justice. *Journal of Clinical and Translational Science*, 7(1):e63.

Reijula, S., Kuorikoski, J., Ehrig, T., Katsikopoulos, K., and Sunder, S. (2018). Nudge, boost, or design? Limitations of behavioral policy under social interaction. *Journal of Behavioral Economics for Policy*, 2(1):5–12.

Simkulet, W. (2019). Informed consent and nudging. *Bioethics*, 33(1):169–184.

Sunstein, C. R. (2018). Misconceptions about nudges. *Journal of Behavioral Economics for Policy*, 2(1):61–67.

Tang, S., Hu, X., Atlas, L., Khanzada, A., and Pilanci, M. (2022). Hierarchical multi-modal transformer for automatic detection of COVID-19. In *Proceedings of the International Conference on Signal Processing and Machine Learning (SPML '22)*, pages 197–202.

Thaler, R. H. and Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press, New Haven.

Torrente, F., Bustin, J., Triskier, F., Ajzenman, N., Tomio, A., Mastai, R., and Boo, F. L. (2020). Effect of a social norm email feedback program on the unnecessary prescription of nimodipine in ambulatory care of older adults: A randomized clinical trial. *JAMA Network Open*, 3(12):e2027082.

VanEpps, E. M., Volpp, K. G., and Halpern, S. D. (2016). A nudge toward participation: Improving clinical trial enrollment with behavioral economics. *Science Translational Medicine*, 8(348):348fs13.

Whang, S. E., Roh, Y., Song, H., and Lee, J.-G. (2023). Data collection and quality challenges in deep learning: A data-centric AI perspective. *The VLDB Journal*, 32:791–813.

Zahren, C., Harvey, S., Weekes, L., Bradshaw, C., Butala, R., Andrews, J., and O'Callaghan, S. (2021). Clinical trials site recruitment optimisation: Guidance from Clinical Trials: Impact and Quality. *Clinical Trials*, 18(5):594–605.

Appendix 1: Clinical Protocols

Here we provide a de-identified, NDA-compliant summary of the clinical protocols, reporting all three recruitment methods, formatted to match the style of ClinicalTrials.gov for public dissemination.

Study Title: COVID-19 Screening via Smartphone Leveraging AI to Detect COVID-19 from Voice

Brief Summary

This study will evaluate whether short smartphone-recorded audio samples, such as cough, voice, and breathing sounds, can help identify adults who may require confirmatory COVID-19 testing. Participants will provide brief audio recordings and complete their routine COVID-19 test as part of standard care. Total participation time will be approximately 15-20 minutes. No treatment will be provided, and no clinical decisions will be based on audio results.

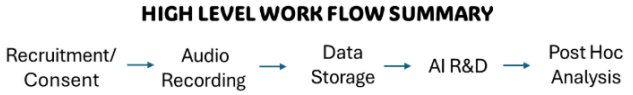

Figure 4. Summary of the workflow of patient data collection to AI analysis.

Study Design: All procedures will be conducted in accordance with the IRB-approved protocol and applicable ethical and regulatory standards.

- **Study Type:** Observational
- **Design:** Cross-sectional, multi-site
- **Time Perspective:** Single visit
- **Primary Purpose:** Screening accuracy evaluation

Eligibility Criteria

Inclusion Criteria

- Age $\geq$18 years
- Currently undergoing rt-PCR COVID-19 testing
- Able to give informed consent
- Pregnant women

Exclusion Criteria

- Inability to consent
- Vocal cord issues, tracheostomy
- Cystic fibrosis

Study Procedures

1. **Training of Clinical Staff:** All clinical staff will undergo standardized training provided by the global clinical research coordinators and the Principal Investigator (PI). Training includes informed consent procedures, safe audio collection using the mobile application, data entry workflows, and infection-control procedures.

2. **Recruitment Methods:** All recruitment approaches will emphasize voluntary participation and do not influence clinical testing or care. Participants will be recruited using three IRB-approved, non-coercive methods:

   a. **Face-to-Face Recruitment [India, Pakistan, Colombia]:**
   Patients undergoing an rt-PCR test swab test will be briefly explained the study and offered participation by their clinician at test time.

   b. **Flyers With QR Codes [Colombia, Japan]:**
   Flyers, posters, or notices displayed in approved public areas at the testing site will provide a QR code linking to study information and contact details.

   c. **Automated Outreach ("Robocalls"):**
   Automated phone calls will be made to individuals who recently undertook a PCR test once their results have been made available to obtain verbal consent and record cough and speech samples over the phone.

3. **Informed Consent & Patient Enrollment**

   a. **Face-to-Face:**

   - Staff will use only provided study-specific smartphones for data collection and audio recording, which will be sanitized before and after each use.
   - Staff will present a paper or digital consent form within the app that includes study information, participant rights, data confidentiality, and voluntary participation. A standardized consent script will be used, covering:
     - Purpose of the study
     - Procedures and expected time commitment
     - Potential risks and minimal-risk nature of participation
     - No direct clinical benefit
     - Privacy protections and data handling
     - Voluntary participation and right to withdraw
     - Participants will have an opportunity to ask questions before providing written informed consent
   - Enrollment and Study Identification: Upon consent, each participant will be provided an anonymous, unique study identification number, in order that they may request at a future date to withdraw from the study and have their data deleted.
   - Demographic and Clinical Data Collection: Using the study mobile app, research staff will collect standardized demographic and clinical information, including:
     - Site location (pre-filled)
     - Age, sex, and basic demographic details
     - COVID-19 vaccination status
     - Self-reported respiratory symptoms (e.g., cough, fever, shortness of breath)
     - Relevant medical conditions (e.g., asthma, COPD, cardiovascular disease)
     - PCR test results

   b. **Flyers with QR Codes (reducing staff burden):**

   - The study team will implement participant-initiated recruitment methods to supplement standard face-to-face outreach.
   - Informational flyers containing QR codes linked to the mobile study app will be placed in designated reception and waiting areas where individuals wait prior to routine COVID-19 testing.
   - Potential participants will be able to scan the QR code to access study information and initiate participation independently, without requiring direct staff interaction.
   - At selected international sites, routine COVID-19 test-result notification emails will be modified, subject to local IRB approval, to include a link directing recipients to the study information page and mobile application.
   - Research staff will monitor the placement and availability of recruitment materials and will ensure that all procedures remain compliant with ethical and regulatory requirements.

   c. **Automated Outreach (Robocalls):**

   - An automated calling system will be deployed to deliver a pre-recorded message approved by the IRB.
   - The call will provide general information about the study purpose and recipients will be asked whether they authorize the recording of their voice and cough for research purposes.
   - If authorization is granted, the system will guide the individual through a standardized set of instructions for completing the audio recordings.
   - All audio data collected through this automated system will be linked to a study ID and stored securely according to the study's data management plan.
   - Recipients who decline authorization at any point or hang up, will not have their recordings collected.
   - All automated call scripts, branching logic, and messages will be reviewed and approved by the IRB prior to implementation.

Compliance and Monitoring:

- Study personnel will ensure that the automated system functions consistently with the IRB-approved script and logic pathway.
- Any system errors, participant complaints, or deviations will be documented and reported according to the study's deviation management procedures.

4. **Audio Recording:** Participants will provide short cough, voice, and breathing recordings using a smartphone app or a phone call (for robocalls). No medical intervention or treatment will be provided.

5. **Clinical Testing:** The participant will undergo rt-PCR testing performed by clinical personnel as part of standard care. Study procedures will not alter or delay routine clinical testing. After the participant exits, staff will sanitize the testing area and study smartphones in accordance with site infection-control procedures.

6. **Data Reporting and Database Entry:** Once clinical test results are available, study personnel will enter the rt-PCR results into the secure study database and associate them with the participant's unique study identifier.

7. **Data Security and Confidentiality:** All data, including audio recordings and clinical test results, will be securely stored using encrypted transmission and storage systems on Amazon Web Services (AWS) with access only to authorized study staff. Audio data will be analyzed only after being fully de-identified.

8. **AI Data Analysis:** The anonymous audio samples paired with corresponding rt-PCR results will be made available to the AI R&D team for algorithm development and evaluation.

9. **Quality Control:** All study procedures, consent documentation, and data entry will be periodically reviewed by the PI or designated monitors to ensure protocol adherence and data integrity.

10. **Primary Outcome Measures:**

    - Screening Sensitivity: Ability of the audio-based system to identify individuals who test positive via standard COVID-19 testing.
    - Negative Predictive Value (NPV): Probability that a negative audio-based screen corresponds to a negative clinical COVID-19 test.
    - Area Under the ROC Curve (AUROC): Overall accuracy compared with standard diagnostic testing.

11. **Study Locations:** Multiple clinical testing sites across Colombia, Japan, India, and Pakistan.

12. **Data Sharing Statement:** Only anonymized, aggregate data may be shared in scientific publications or presentations.

13. **Sponsor / Collaborators:** Qualified investigators and institutional partners (details withheld per NDA).