



OPEN ACCESS

EDITED BY

Bassey Ebenso,
University of Leeds, United Kingdom

REVIEWED BY

Eve Namisango,
African Palliative Care
Association, Uganda
Kehinde Sharafadeen Okunade,
University of Lagos, Nigeria

*CORRESPONDENCE

Delivette Castor
dc2022@cumc.columbia.edu

SPECIALTY SECTION

This article was submitted to
Implementation Science,
a section of the journal
Frontiers in Health Services

RECEIVED 21 July 2022

ACCEPTED 23 August 2022

PUBLISHED 12 September 2022

CITATION

Castor D, Saidu R, Boa R, Mbatani N,
Mutsvangwa TEM, Moodley J, Denny L
and Kuhn L (2022) Assessment of the
implementation context in preparation
for a clinical study of
machine-learning algorithms to
automate the classification of digital
cervical images for cervical cancer
screening in resource-constrained
settings. *Front. Health Serv.* 2:1000150.
doi: 10.3389/frhs.2022.1000150

COPYRIGHT

© 2022 Castor, Saidu, Boa, Mbatani,
Mutsvangwa, Moodley, Denny and
Kuhn. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Assessment of the implementation context in preparation for a clinical study of machine-learning algorithms to automate the classification of digital cervical images for cervical cancer screening in resource-constrained settings

Delivette Castor^{1,2*}, Rakiya Saidu^{3,4}, Rosalind Boa^{3,4},
Nomonde Mbatani^{3,4}, Tinashe E. M. Mutsvangwa⁵,
Jennifer Moodley^{4,6}, Lynette Denny^{3,4} and Louise Kuhn^{2,7}

¹Division of Infectious Diseases, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, United States, ²Department of Epidemiology, Mailman School of Public Health, Columbia University Irving Medical Center, New York, NY, United States, ³Department of Obstetrics and Gynaecology, University of Cape Town, Cape Town, South Africa, ⁴Groote Schuur Hospital and South African Medical Research Council, Gynaecology Cancer Research Centre, University of Cape Town, Cape Town, South Africa, ⁵Division of Biomedical Engineering, Department of Human Biology, University of Cape Town, Cape Town, South Africa, ⁶Women's Health Research Unit, School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa, ⁷Gertrude H. Sergievsky Center, Vagelos College of Physicians and Surgeons, Columbia University Irving Medical Center, New York, NY, United States

Introduction: We assessed the implementation context and image quality in preparation for a clinical study evaluating the effectiveness of automated visual assessment devices within cervical cancer screening of women living without and with HIV.

Methods: We developed a semi-structured questionnaire based on three Consolidated Framework for Implementation Research (CFIR) domains; intervention characteristics, inner setting, and process, in Cape Town, South Africa. Between December 1, 2020, and August 6, 2021, we evaluated two devices: MobileODT handheld colposcope; and a commercially-available cell phone (Samsung A21ST). Colposcopists visually inspected cervical images for technical adequacy. Descriptive analyses were tabulated for quantitative variables, and narrative responses were summarized in the text.

Results: Two colposcopists described the devices as easy to operate, without data loss. The clinical workspace and gynecological workflow were modified to incorporate devices and manage images. Providers believed either device would likely perform better than cytology under most circumstances unless the squamocolumnar junction (SCJ) were not visible, in which case cytology was expected to be better. Image quality ($N = 75$) from the MobileODT device and cell phone was comparable in terms of achieving good focus (81% vs. 84%),

obtaining visibility of the squamous columnar junction (88% vs. 97%), avoiding occlusion (79% vs. 87%), and detection of lesion and range of lesion includes the upper limit (63% vs. 53%) but differed in taking photographs free of glare (100% vs. 24%).

Conclusion: Novel application of the CFIR early in the conduct of the clinical study, including assessment of image quality, highlight real-world factors about intervention characteristics, inner clinical setting, and workflow process that may affect both the clinical study findings and ultimate pace of translating to clinical practice. The application and augmentation of the CFIR in this study context highlighted adaptations needed for the framework to better measure factors relevant to implementing digital interventions.

KEYWORDS

digital cervical-cancer screening implementation assessment cervical cancer, implementation science, CFIR, Automated Visual Evaluation, cervical cancer, CFIR framework, Automated Visual Evaluation of the cervix

Introduction

Global disparities in cervical cancer incidence and mortality persist. Women living in sub-Saharan Africa bear a disproportionate burden due to the failure of current cervical cancer screening programs (1). Screen-and-treat (SAT) is a strategy for cervical cancer screening that has been shown to improve outcomes and is currently recommended by the World Health Organization for low- and middle-income country (LMIC) settings (2). Recently developed smartphone technologies that utilize machine-learning algorithms to automate the classification of digital cervical images are an encouraging method with the SAT approach to address the achievement gap in LMICs (3–7). Preliminary studies suggest that Automated Visual Evaluation (AVE) devices could detect cervical cancer precursor lesions in some populations (3, 5, 7). Private companies have developed proprietary devices to obtain images and run automated classifiers to improve screening (8, 9). Other groups are using non-proprietary cell phones to take cervical photographs to integrate automated classifiers through apps or other interfaces eventually. Requirements for successful use of AVE devices highlight attributes of the device and the algorithm that affect cervical image quality, but not the clinical setting (3). Pragmatic trialists advocate for measuring these operational attributes that influence study outcomes and ultimately impact the translation of scientific innovation to practice settings, but this is rarely done (10–12). Implementation science frameworks typically used for “real-world” delivery can be applied to assess and inform implementation issues during the clinical study stage and strengthen inference about internal and external validity (13). In preparation for a clinical study to evaluate the performance characteristics of AVE methods utilizing images collected from both a proprietary mobile device

and a commercially-available smartphone for cervical cancer screening among women living with HIV and women not infected with HIV, we assessed the implementation context and image quality using an implementation framework.

Materials and methods

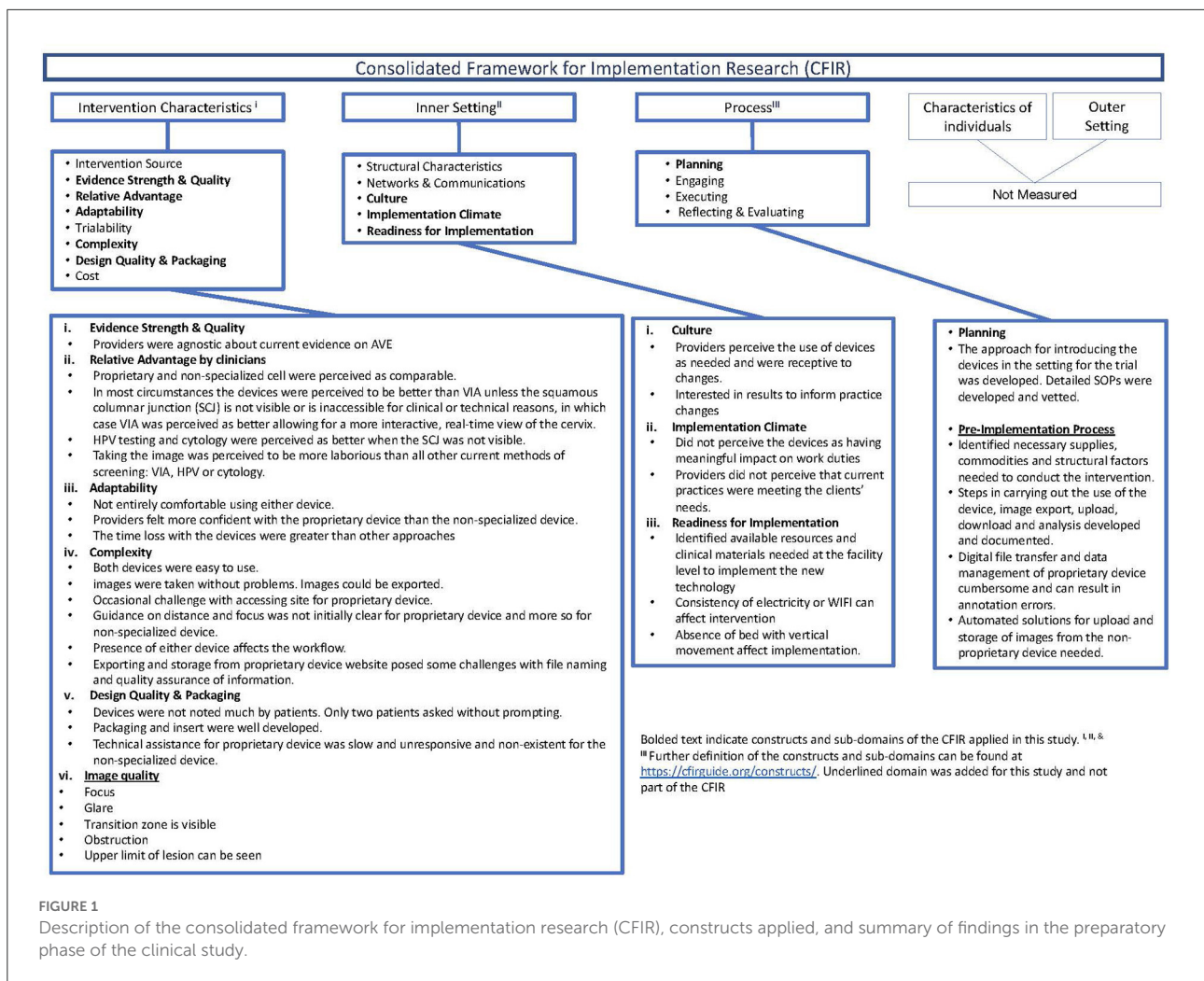
Our team, which included experienced colposcopists, conducted a mixed-methods study utilizing the Consolidated Framework for Implementation Research (CFIR) to examine two devices: MobileODT hand-held colposcope and the Samsung A21ST, a commercially-available cell phone. The MobileODT device is an FDA-approved hand-held colposcope that captures and stores digital photographs of the cervix and can interface with a cloud-based machine-learning algorithm to generate an automated diagnosis based on the cervical image (8, 9).

Study setting

The study was implemented at a free-standing clinical research site, constructed out of repurposed shipping containers, on the compound of Khayelitsha Site B Primary Health Care Clinic (KPHC), a large public clinic serving a disadvantaged community on the outskirts of Cape Town, South Africa.

Consolidated framework for implementation research

The CFIR is an implementation science framework for assessing barriers and facilitators to the successful



implementation of novel interventions through five constructs: (1) Intervention characteristics, (2) Inner setting, (3) Process, (4) Outer Setting, and (5) Characteristics of the individuals (14). Typically, the CFIR is used during the implementation phase after established effectiveness. We applied the CFIR during the preparatory phase of a clinical study and utilized selected sub-domains of three of the constructs (Figure 1). Further, we modified the CFIR to include pre-trial readiness activities, including the intervention's commodities and data management aspects. We applied four of the eight sub-domains and added a sub-domain on digital image quality within the intervention characteristics construct; adaptability, complexity, perceived strength and quality of evidence base, and relative advantage of the mobile devices compared to the other screening modalities, including visual inspection with acetic acid (VIA), cytology, and human papillomavirus (HPV) testing. We utilized the sub-domains of readiness for implementation, climate, and culture within the inner setting to assess adaptations needed

to incorporate the device into the clinical space and workflow. Last, we used the planning sub-domain and added a pre-implementation sub-domain to the process construct to identify the commodities, including data transfer and management capabilities needed to acquire, manage and analyze the digital image data required to effectively screen for cervical cancer.

We developed an eighty-five-item semi-structured questionnaire. A study team member administered the semi-structured interview to two clinicians preparing to use the MobileODT and Samsung A21ST devices at the clinical research site. Responses were coded as either yes/no or using Likert scales. Providers elaborated with qualitative responses on any quantitative measures.

In addition, colposcopists visually inspected the quality of all the cervical images taken in the context of site preparation and assessed them for technical adequacy. Image quality was scored for five key attributes; (1) focus, (2) glare, (3) visibility

of the squamous columnar junction (SCJ), (4) occlusion, and (5) detection of lesion and range of lesion, including the upper limit. Composite image quality scores ranged from a minimum of 0 to a maximum score of 5. Frequency counts and percentages were tabulated for quantitative categorical variables, and narrative responses were summarized in the text.

Results

Between December 1, 2020, and August 6, 2021, two providers screened 75 women at the research site in Khayelitsha, South Africa. Results from this evaluation using the CFIR are shown in [Figure 1](#).

Intervention characteristics

Providers' perceptions of the relative advantage of the devices compared to standard practice interventions (i.e., VIA, HPV testing, cytology) were similar before and after use. Providers thought that either machine would outperform VIA or cytology in general. However, if lesions are present on the endocervix or the squamous columnar junction (SCJ) is not visible or inaccessible for clinical or technical reasons. Cytology was perceived as better, and VIA was perceived as comparable or better, allowing for a more interactive real-time view of the cervix. Providers agreed that HPV testing was expected to be better when the SCJ was not visible but had a divergence of opinion when comparing the perceived relative advantage of the mobile devices to HPV testing under other conditions, such as when anogenital lesions are present.

Both devices had a learning curve to be used effectively but were easy to activate, take, edit, and upload images without data loss. Documentation of the images, including correct labeling before image capture and upload and storage after image capture, required additional attention. To improve usability, the MobileODT device has some design modifications; a stand, built-in light source, autofocus feature, and handsfree activation. MobileODT automated image upload to a storage platform but required some input to ensure appropriate patient linkage before image collection. The cell phone required experimentation with external light sources and flash features. Providers used a gimbal and the "open-camera" app as strategies for a stand-and-hands-free activation for the cell phone; however, neither option was more accessible than holding the phone in their hand. A gimbal is a tool that uses motors and intelligent sensors to stabilize the camera, but the type that was purchased did not allow retention of raw digital images. Labeling, uploading, and storing the cell phone images required providers to complete additional manual steps that were not needed with the MobileODT device.

Image quality

[Figure 2](#) shows individual and summarized image quality assessment scores for MobileODT ($N = 75$) and the cell phone ($N = 38$). MobileODT and the cell phone performance in terms of achieving good focus (81% vs. 84%), obtaining visibility of the SCJ (88% vs. 97%), avoiding occlusion (79% vs. 87%), and detection of the range of the lesion to include the upper limit (63% vs. 53%) was comparable. However, the MobileODT device outperformed the cell phone's function in taking photographs free of glare (100% vs. 24%). MobileODT and cell phone cumulative scores across the five image quality metrics were; 33% vs. 13% scored a five, 48% vs. 37% scored a four, 15% vs. 34% scored a three, 0.04% vs. 13% scored a two, and 0% vs. 0.03% scored a one, respectively.

Inner setting

The space available at the clinical research site in the exam room posed minor obstacles for using either device. We rearranged the space to create an adequate area for the device and staff to operate comfortably. Infrastructure constraints, including frequent and regular loss of electricity and WIFI, adversely affected uploading the images from both devices. A gynecological bed with vertical movement could better accommodate image capture from both devices. The clinical workflow was modified for pre-collection labeling with both devices. The cell phone required manual upload and storage.

Process

Both devices performed reasonably well—they charged quickly, the MobileODT device linked to the website and cloud, operated with minimal issues, and posed no significant logistical challenges to the providers in taking reasonable images. About 10% of the MobileODT device used were delayed and needed a restart. MobileODT automated data upload, whereas the cell phone required the provider to upload the images in a cumbersome and time-intensive way manually. We used the "open camera" app, and subsequently used a gimbal for their "always on" feature on the cell phone because cell phone images were more affected by variation in lighting. The built-in flash, flashlight feature of the cell phone, or the gimbal's always on could not adequately mirror the consistent light source available on MobileODT. Providers modified the use of the device (e.g., take the proprietary device off the stand used to stabilize) because of anatomical variation in some patients. The MobileODT device automatically uploaded images to secure storage, but there were constraints to the proprietor's cloud. We did not test the automated classification feature and return of results of the MobileODT device in this evaluation phase.

Discussion

Our novel application of the CFIR framework during the preparatory phase of a clinical study evaluating the classification of cervical images taken with two AVE devices revealed many domains where operational issues can affect the conduct and outcome of the clinical research. Implementation science frameworks developed for systematic evaluation in real-world settings can be utilized to assess the pragmatic aspects of experimental designs that can influence study outcomes through a theory-based approach (13). We modified the CFIR construct on intervention characteristics to incorporate data transfer and management practices for digital images and machine learning interventions. Also, the process construct was modified to include pre-implementation materials and processes, including electrical, internet, and other digital infrastructure needed for implementation. The changes and use of the CFIR will strengthen the clinical study and could accelerate translation to public health practice.

There were several limitations and strengths to this study. Study limitations included few providers and sites in this preliminary evaluation. While providers gleaned some patient feedback, we did not directly assess patient perspectives in this evaluation. We evaluated two devices simultaneously. Using a less costly, commercially-available, non-specialized cell phone may overcome some economic barriers to broader reach and better coverage in health systems of LMICs.

Conversely, the proprietary device successfully addressed some challenges with taking and storing high-quality digital images. The device evaluation was conducted while clinical services were adapted to mitigate COVID-19 risk. The unique period resulted in frequent consultations within the clinical team and rapid identification and sharing of best practices, which heightened the organization's readiness for practice change. The evaluation was conducted at a research site on a public primary health care clinic compound in a deeply-impooverished, densely-populated urban community in Cape Town. The cervical cancer screening program setting is representative of other resource-limited locations where this approach is likely to be implemented if found to be effective. The assessment likely identified implementation challenges that must be confronted in other public care, resource-constrained environments.

We applied the CFIR to examine implementation factors at this early stage of the clinical study. We selected the CFIR because it is an implementation and scaling framework we intend to apply throughout the research and development process. Should the device show clinical effectiveness, we will use the CFIR in its classical form during the introduction and scale-up phase of implementing a health intervention. In this evaluation, we utilized three of the five domains of the framework. We modified two domains used to adequately measure the program, facility, and provider-level preparation needed to support the digital devices. Should the devices prove

effective, the two additional domains not included in this evaluation, outer setting and individual characteristics, would also need to be evaluated.

The context domain includes an assessment of efforts to support policy changes and the broader implementation context that becomes relevant once the intervention is implemented in a real-world setting. For instance, the device should be incorporated into national guidelines. Implementation plans will need to be modified or developed. The digital device should be integrated into national and local health information systems. Human resource considerations should be extended to include skilled bioinformaticians, engineers, and developers to sustain the evolving system. Costing studies to inform implementation and funding advocacy will also be necessary (15). Funding for maintaining hardware, stable electricity, the digital infrastructure, and costs for essential accessories and acquiring, managing, storing, and analyzing the data for machine learning (16). Other important global factors to address in implementing machine learning digital devices will involve complex ethics, legality, and social issues that should be anticipated and addressed (15, 17, 18).

Embedded within the broader context are other vital factors that we believe are critical for sustainable scale-up, but we could not fully assess in this study; the evolving algorithm, device, provider, and patient that would be further examined in the context of scaling that we could not adequately assess. Critically important issues relating to the algorithm and device include data quality improvements, starting with reliable data availability with validated clinical outcomes. We may reduce the likelihood of false classification due to limited data with annotated digital image data linked with histological results and other clinical, anatomical, or biological factors that may confound the digital method, sampling a higher fraction of the population. Improvements in the classifier may increase advocacy and demand for the digital method among various cervical cancer stakeholders in the health system. Also, important aspects of the machine learning device are algorithmic transparency, safeguards to minimize the harmful effects of bias, regulatory guidance, liability, and accountability. At the provider level, capacity strengthening through training on the use of the device, Machine learning, and AI, as well as alternations to the workflow, task-shifting, and sharing to improve program efficiency while addressing program quality (19–21). Individual characteristics important for scaling include assessments of end-user values and preferences. User-specific issues that will influence scaling include the use of the device concerning patient's privacy, data security, knowledge and perception of the relative advantage of the device compared to routine methods for cervical cancer screening, and patient's time (14). Further, considerations about equity in linkage and access to gynecologic care sustained patient safety with the application of the algorithm and use of the device with reinforcement

through clinical expertise; and the machine learning device's impact on the patient-physician relationship (15, 17, 21).

Experience from other machine learning interventions in Africa suggests that scaling up machine learning digital methods can be supported with data demonstrating that digital aid's effectiveness, precision, and efficiency improvements as the training data expands (22). We endeavor to support scale-up at this earliest stage. We identified common issues in using any device for image capture that should be addressed as part of site preparation; space reconfiguration, labeling, data storage, and adapting to anatomic variation in patients. We also identified device-specific issues, such as an internal light source, stability, and connectivity, that would need to be addressed. The systematic image quality assessment we undertook provided insight into quality metrics that would be important to include in quality control. We were able to gauge the clinicians' confidence, comfort, and enthusiasm about the investigated intervention. The evaluation also provided benchmarks to identify and address emerging issues at the site level that may influence study findings.

Conclusion

The CFIR and image quality assessment provide a comprehensive approach to monitoring and evaluation, and quality improvement practices with this screening tool should effectiveness be proven. Although automated diagnosis from digital cervical images is a promising approach, several operational considerations need careful attention to ensure that this potential screening modality can operate at a high standard (4).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The study was approved by the University of Cape Town #665/2020 Human Research Ethics Committee

References

1. Denny L, de Sanjose S, Mutebi M, Anderson BO, Kim J, Jeronimo J, et al. Interventions to close the divide for women with breast and cervical cancer between low-income and middle-income countries and high-income countries. *Lancet*. (2017) 389:861–70. doi: 10.1016/S0140-6736(16)31795-0
2. Kuhn L, Denny L. The time is now to implement HPV testing for primary screening in low resource settings. *Prev Med*. (2017) 98:42–4. doi: 10.1016/j.ypmed.2016.12.030

and Columbia University #AAAT0187 Institutional Review Board. The patients/participants provided their written informed consent to participate in this study.

Author contributions

DC, LK, RS, LD, RB, and JM conceptualized and designed the study. DC, LK, and RS managed and analyzed the data. LK and LD acquired funding for this study. DC prepared the figures and tables and the first draft of the manuscript. All authors reviewed, edited, and commented on the final draft. All authors contributed to the article and approved the submitted version.

Funding

This study was funded in part by the National Cancer Institute (R01CA254576 and R01CA250011).

Acknowledgments

The authors thank the participants and staff of the study at Khayelitsha Site B Primary Health Care Clinic (KPHC).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

3. Desai KT, Befano B, Xue Z, Kelly H, Campos NG, Egem D, et al. The development of "automated visual evaluation" for cervical cancer screening: the promise and challenges in adapting deep-learning for clinical testing: interdisciplinary principles of automated visual evaluation in cervical screening. *Int J Cancer*. (2022) 150:741–52. doi: 10.1002/ijc.33879
4. Hu L, Bell D, Antani S, Xue Z, Yu K, Horning MP, et al. An observational study of deep learning and automated evaluation of cervical images for cancer screening. *J Natl Cancer Inst*. (2019) 111:923–32. doi: 10.1093/jnci/djy225

5. Bae JK, Roh HJ, You JS, Kim K, Ahn Y, Askaruly S, et al. Quantitative screening of cervical cancers for low-resource settings: pilot study of smartphone-based endoscopic visual inspection after acetic acid using machine learning techniques. *JMIR Mhealth Uhealth*. (2020) 8:e16467. doi: 10.2196/16467
6. Catarino R, Vassilakos P, Scaringella S, Undurraga-Malinverno M, Meyer-Hamme U, Ricard-Gauthier D, et al. Smartphone use for cervical cancer screening in low-resource countries: a pilot study conducted in Madagascar. *PLoS ONE*. (2015) 10:e0134309. doi: 10.1371/journal.pone.0134309
7. Tran PL, Benski C, Viviano M, Petignat P, Combescure C, Jinoro J, et al. Performance of smartphone-based digital images for cervical cancer screening in a low-resource context. *Int J Technol Assess Health Care*. (2018) 34:337–42. doi: 10.1017/S0266462318000260
8. Mink J, Peterson C. MobileODT: a case study of a novel approach to an mHealth-based model of sustainable impact. *Mhealth*. (2016) 2:12. doi: 10.21037/mhealth.2016.03.10
9. Peterson CW, Rose D, Mink J, Levitz D. Real-time monitoring and evaluation of a visual-based cervical cancer screening program using a decision support job aid. *Diagnostics*. (2016) 6:20–8. doi: 10.3390/diagnostics6020020
10. Nieuwenhuis JB, Irving E, Oude Rengerink K, Lloyd E, Goetz I, Grobbee DE, et al. Pragmatic trial design elements showed a different impact on trial interpretation and feasibility than explanatory elements. *J Clin Epidemiol*. (2016) 77:95–100. doi: 10.1016/j.jclinepi.2016.04.010
11. Worsley SD, Oude Rengerink K, Irving E, Lejeune S, Mol K, Collier S, et al. Series: pragmatic trials and real world evidence: paper 2. Setting, sites, and investigator selection. *J Clin Epidemiol*. (2017) 88:14–20. doi: 10.1016/j.jclinepi.2017.05.003
12. Zuidgeest MGP, Goetz I, Groenwold RHH, Irving E, van Thiel G, Grobbee DE, et al. Series: pragmatic trials and real world evidence: paper 1. Introduction. *J Clin Epidemiol*. (2017) 88:7–13. doi: 10.1016/j.jclinepi.2016.12.023
13. Rudd BN, Davis M, Beidas RS. Integrating implementation science in clinical research to maximize public health impact: a call for the reporting and alignment of implementation strategy use with implementation outcomes in clinical research. *Implement Sci*. (2020) 15:103. doi: 10.1186/s13012-020-01060-5
14. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci*. (2009) 4:50. doi: 10.1186/1748-5908-4-50
15. USAID. Foundation TR. Foundation BMG. *Artificial Intelligence in Global Health: Defining a Collective Path Forward*. Washington, DC: USAID (2019).
16. Hosny A, Aerts H. Artificial intelligence for global health. *Science*. (2019) 366:955–6. doi: 10.1126/science.aay5189
17. The Lancet. Artificial intelligence in global health: a brave new world. *Lancet*. (2019) 393:1478. doi: 10.1016/S0140-6736(19)30814-1
18. Wiegand T, Krishnamurthy R, Kuglitsch M, Lee N, Pujari S, Salathe M, et al. WHO and ITU establish benchmarking process for artificial intelligence in health. *Lancet*. (2019) 394:9–11. doi: 10.1016/S0140-6736(19)30762-7
19. Cartolovni A, Tomicic A, Lazic Mosler E. Ethical, legal, and social considerations of AI-based medical decision-support tools: a scoping review. *Int J Med Inform*. (2022) 161:104738. doi: 10.1016/j.ijmedinf.2022.104738
20. Paul AK, Schaefer M. Safeguards for the use of artificial intelligence and machine learning in global health. *Bull World Health Organ*. (2020) 98:282–4. doi: 10.2471/BLT.19.237099
21. Mehta MC, Katz IT, Jha AK. Transforming global health with AI. *N Engl J Med*. (2020) 382:791–3. doi: 10.1056/NEJMp1912079
22. Ling LJ. [Management of several special periodontal problems]. *Zhonghua Ya Yi Xue Hui Za Zhi*. (1986) 5:36–8.