USSD-Based Digital Health in Rural Africa: A Machine Learning Research Direction for Low-Resource Health Signal Processing

Ayomide H. Agbaje[1]

[1]African Leadership University, Kigali, Rwanda

Publication Date: 2026/05/07

Abstract:

- ***Background:***

Sub-Saharan Africa (SSA) bears 25% of the global disease burden, yet accounts for only 3% of the world's health workforce [1]. Smartphone-dependent digital health platforms have failed to reach most rural populations in SSA, and the 2023 collapse of Babyl Rwanda demonstrated the structural fragility of externally owned digital health infrastructure [2].

- ***Objective:***

To evaluate the feasibility, acceptability, and 90-day user retention of HealthDrive, a USSD-based telehealth platform with community health worker (CHW) integration, in a pilot study conducted in two rural SSA communities.

- ***Methods:***

A mixed-methods pilot implementation study (n=50 enrolled patients, 12 CHWs) conducted August 2024 to March 2026, applying the Consolidated Framework for Implementation Research (CFIR) [3], the Technology Acceptance Model for Resource-Limited Settings (TAM-RLS) [4], and the RE-AIM evaluation framework [5]. USSD interaction logs (1,247 sessions across four short codes), CHW follow-up records, and structured satisfaction interviews were analysed.

- ***Results:***

Three-month user retention was 78% (95% CI: 64–88%), exceeding SSA mHealth benchmarks (45–65%). Elderly user satisfaction reached 85%. Emergency triage sessions achieved 71% completion. Total platform expenditure was $2,580 over 19 months at $125/month.

- ***Conclusions:***

USSD-based telehealth with CHW integration is feasible and acceptable in rural SSA. Five open machine learning and signal processing challenges are identified as critical barriers to scaling this model to population-scale voice-based health triage.

Keywords: *USSD; Digital Health; mHealth; Community Health Workers; sub-Saharan Africa; Machine Learning; Speech Processing; Implementation Science.*

How to Cite: Ayomide H. Agbaje (2026) USSD-Based Digital Health in Rural Africa: A Machine Learning Research Direction for Low-Resource Health Signal Processing. *International Journal of Innovative Science and Research Technology*, 11(4), 3449-3460. https://doi.org/10.38124/ijisrt/26apr2523

I. INTRODUCTION

Primary healthcare in rural sub-Saharan Africa operates at the intersection of two compounding deficits. Africa bears approximately 25% of the global disease burden, yet accounts for 3% of the world's health workforce and less than 1% of global financial resources for health [1,6]. Physician densities in East Africa remain among the lowest globally: in the pilot country, approximately 0.116 physicians serve per 1,000 population, equivalent to 1 doctor per 8,294 people, with a severe rural-urban distribution imbalance [7,8]. National community health worker (CHW) programs, deploying trained volunteers at 3–4 per village, function as the primary healthcare interface for most rural households [9].

Most digital health interventions require smartphone connectivity or mobile data that rural SSA populations cannot reliably access. USSD (Unstructured Supplementary Service Data) operates over 2G voice channels without internet access, is free to end users, requires no application download,

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

and functions under low-signal conditions, characteristics that make it a uniquely equitable interface for rural health communication [11]. Mobile SIM penetration exceeds 90% of the population in most SSA countries [10], indicating that USSD infrastructure already exists where smartphones do not.

The 2023 collapse of Babyl Rwanda provides the key structural lesson motivating this study. At closure, Babyl had facilitated 3.9 million teleconsultations over four years and demonstrated measurable clinical effectiveness: providers asked 60–100% more diagnostic questions per consultation, consultations were 30% shorter, and patients paid 20–40% less out-of-pocket than at facility visits [2]. Despite this evidence base, parent company Babylon Health's Chapter 7 bankruptcy filing on August 9, 2023, terminated all services within weeks, immediately reversing measured healthcare utilisation gains at the facility level [2]. Evidence of effectiveness, this case demonstrates, is not a sufficient condition for sustainable digital health infrastructure in SSA [18].

HealthDrive was designed in direct response to this structural problem. Built on Python and PostgreSQL, deployed across four USSD short codes (*911# emergency triage, *123# clinic navigation, *456# AI-assisted symptom classification, *789# CHW coordination), and integrated with existing village-level CHW networks, HealthDrive demonstrates a locally owned, minimal-cost, structurally resilient model for rural primary care delivery. This paper reports a 19-month pilot implementation study. It identifies the computational research challenges that must be solved before this model can scale to population-scale voice-based health triage across SSA.

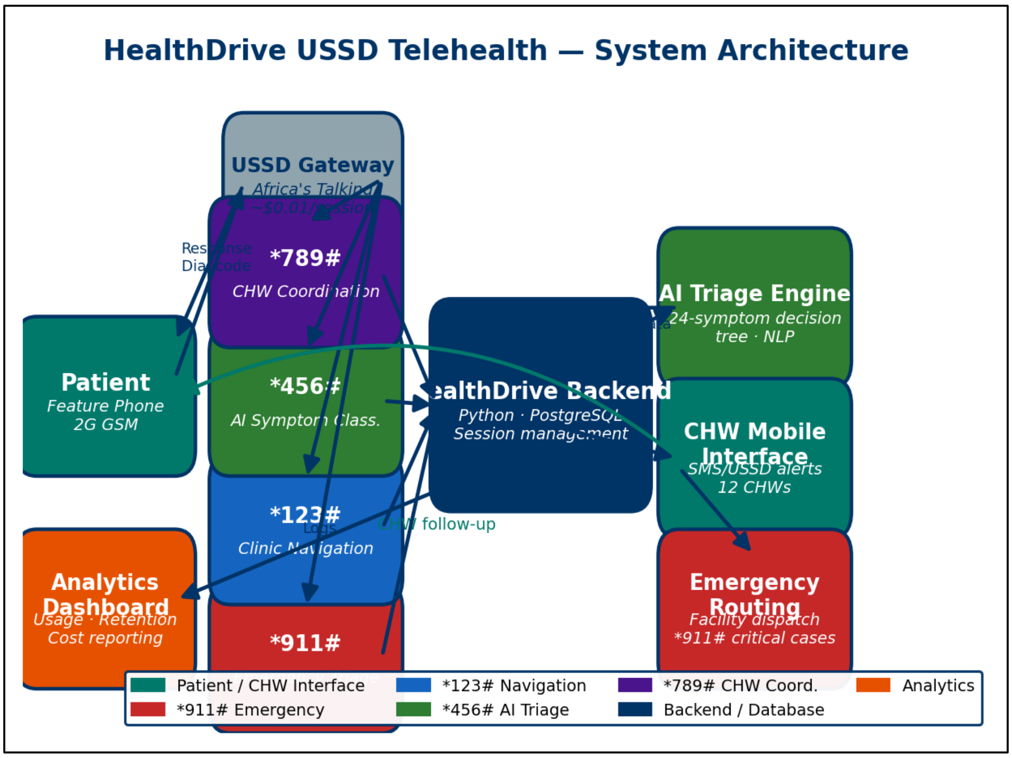

Fig 1 Health Drive System Architecture, four USSD Short Codes, CHW Mobile Interface, AI Triage Backend, Emergency Routing, and Africa's Talking Gateway.

II. BACKGROUND AND RELATED WORK

➢ *Health Workforce Crisis in Sub-Saharan Africa*

A 2024 needs-based modelling study found that the WHO African Region's available health workers covered only 43% of needs-based requirements in 2022, with a projected shortage of 6.1 million health workers by 2030 even under optimistic training trajectory assumptions [1]. The majority of SSA health workers are concentrated in urban centres, while 60–70% of the population lives in rural areas [6,12]. Out-of-pocket health expenditures account for approximately 16.5% of catastrophic spending among SSA households [13]. CHW programs, including Rwanda's Abakundumuganga program, which deploys approximately 58,567 CHWs at 3–4 per village, represent the most scalable mechanism for closing access gaps in rural SSA [9,14].

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

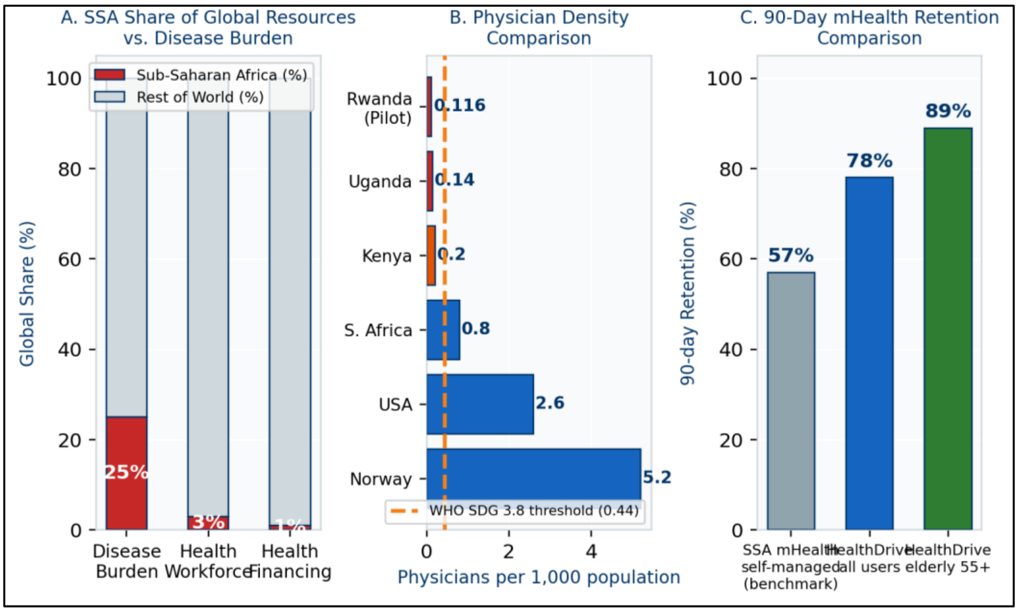

Fig 2 SSA Health Workforce Crisis: A) SSA Share of Global Disease Burden vs Health Workforce; B) Physician Density Across Selected Countries (WHO SDG Threshold = 0.44 per 1,000); C) 90-day mHealth Retention Comparison.

➢ *USSD and Feature Phone Health Platforms in SSA*

MomConnect (South Africa), operational since 2014, has reached approximately 5 million cumulative registrations across 95% of public health facilities in 11 official languages [15]. M-TIBA (Kenya), a mobile-linked health savings and services platform that uses USSD as the primary interaction channel, has enrolled 4.8 million users across 37 of 47 Kenyan counties [16]. Babyl Rwanda, launched in 2016 and expanded as Rwanda's first payer-integrated national telemedicine service from 2019, demonstrated clinical effectiveness through 3.9 million consultations before closing in September 2023 [2]. Its closure is now characterised in the literature as the "Babyl Paradox": strong evidence of effectiveness that could not prevent corporate failure from collapsing care access for 2.4–2.8 million users [18].

➢ *Implementation Science Frameworks*

The Consolidated Framework for Implementation Research (CFIR; Damschroder et al., 2009) organises implementation determinants across five domains: intervention characteristics, outer setting, inner setting, individual characteristics, and implementation process [3]. Means et al. (2020) adapted the CFIR for LMIC contexts in which outer-setting constraints, infrastructure limitations, regulatory gaps, and resource scarcity dominate [19]. The Technology Acceptance Model for Resource-Limited Settings (TAM-RLS; Campbell et al., 2017) extends standard TAM with three LMIC-specific constructs: confidentiality/stigma concerns, downstream care barriers, and CHW adoption mediation [4]. The RE-AIM framework (Glasgow et al., 1999) structures pilot feasibility evaluation across Reach, Effectiveness, Adoption, Implementation, and Maintenance [5].

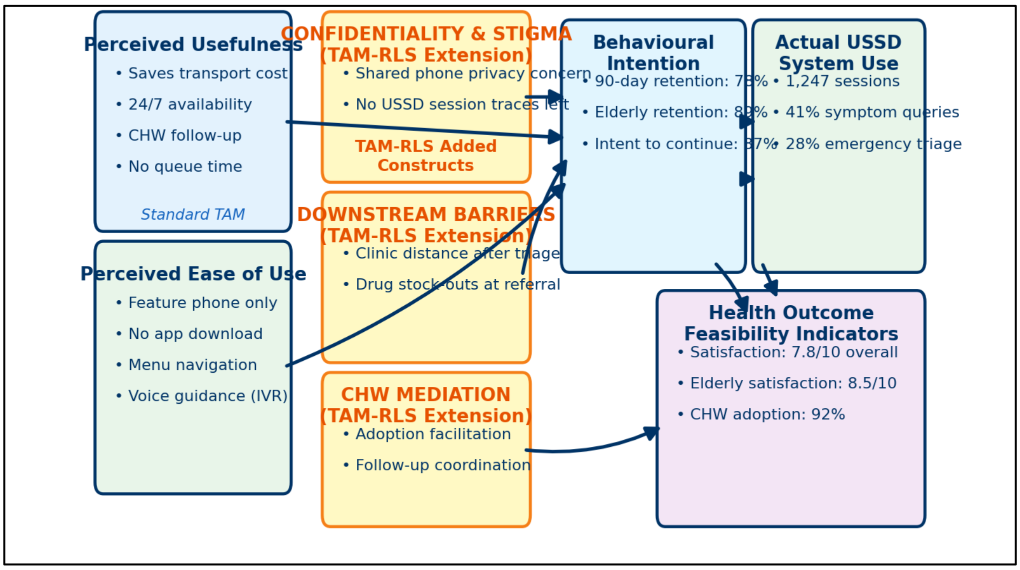

Fig 3 TAM-RLS Framework (Campbell et al., 2017) Applied to HealthDrive, Standard TAM Constructs (left) and three Resource-Limited Setting Extensions (Centre) Mapped to HealthDrive Outcomes (Right).

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

➢ *mHealth Retention in SSA*

Eysenbach's Law of Attrition (2005) established that high dropout rates are inherent to eHealth interventions and follow predictable patterns [20]. Systematic reviews report 30-day retention of 60–80% and 90-day rates of 45–65% for self-managed mHealth tools in SSA [21]. CHW-mediated mHealth consistently outperforms self-managed equivalents on retention by addressing the "downstream barriers" identified in TAM-RLS [4,22].

III. SYSTEM ARCHITECTURE AND SIGNAL CHAIN ANALYSIS

HealthDrive deploys on Python/PostgreSQL infrastructure with ~$20–40 USD/month in cloud costs. USSD sessions route through a gateway aggregator at approximately $0.01 per session to the platform operator, at zero cost to the end user. Figure 4 and Table I formalise the information degradation chain from patient health state to AI triage output, from a signal processing perspective, which motivates the ML research directions in Section VII.

Table 1 HealthDrive Signal Degradation Chain and ML Research Gaps

Stage	Signal Type	Degradation Source	ML/SP Direction
Patient Health State	Symptom occurrence	Literacy gap; vocab mismatch	Health-domain language modelling (§VII-A)
USSD Interaction	DTMF keypress	2G noise; 180-sec timeout	Semi-supervised classification (§VII-D)
2G GSM Transport	AMR codec bearer	Packet loss; 8kHz limit	Channel-robust ASR (§VII-A, C)
Session Parsing	Navigation log	Branching; fragmentation	Weak-label learning (§VII-D)
AI Triage Output	Decision tree result	Vocab ceiling; hallucination	Hallucination-resistant AI (§VII-E)

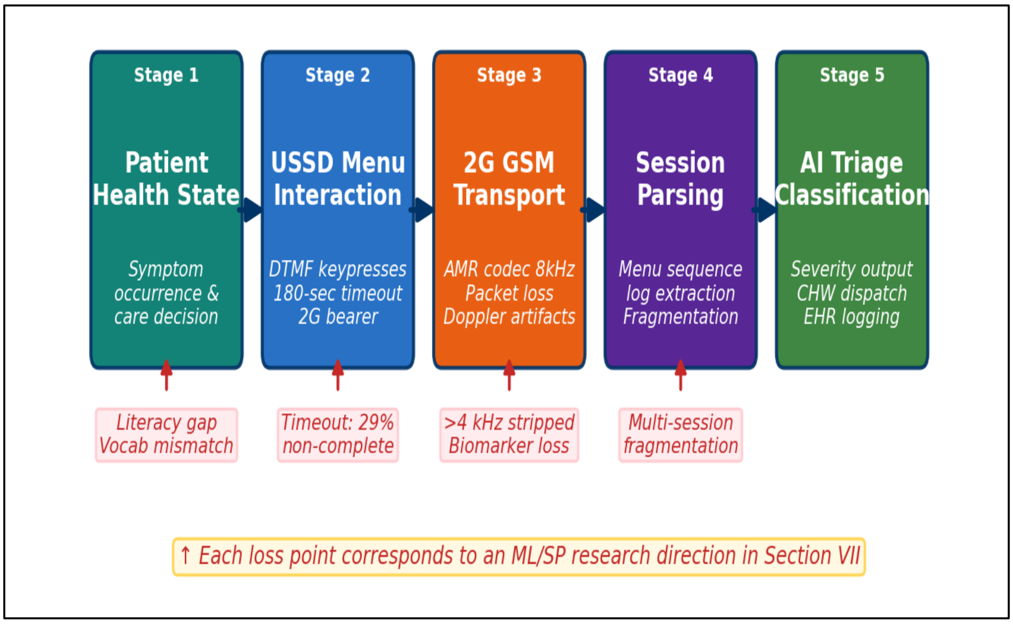

Fig 4 HealthDrive Signal Degradation Chain, Information Loss at Each Stage from Patient Health State to AI Classification Output, with Corresponding ML Research Directions (Section VII).

IV. METHODOLOGY

➢ *Study Design*

This was a mixed-methods pilot implementation study using a convergent parallel design [23], combining quantitative analysis of USSD session logs with qualitative interviews with CHWs and patients. Reporting follows StaRI [24] and TIDieR [25] standards. The study was conducted in accordance with institutional research ethics guidelines (ethics reference: available on request). Informed consent was obtained from all participants before enrollment.

➢ *Setting*

Two rural communities in a high-CHW-density SSA country were selected based on: distance >10 km from the nearest health centre, existing CHW program infrastructure (minimum 3 trained CHWs per village), feature phone ownership >70%, and confirmed 2G network coverage. Communities were matched on approximate size (200–350 households each). Specific community names are withheld to protect participant confidentiality.

➢ *Participants and Sample Size Justification*

Eligible patients were adults aged 18 years or older with feature phone access, residing in the study communities, and seeking primary care for non-emergency conditions. CHWs were eligible with an active certification in a national CHW program. Sample size n=50 was determined per the standard pilot study threshold of 12 participants per group [26], adapted for a single-arm feasibility study [27], consistent with the CONSORT pilot extension [28]. Figure 5 presents the enrollment flow.

ISSN No: -2456-2165
https://doi.org/10.38124/ijisrt/26apr2523

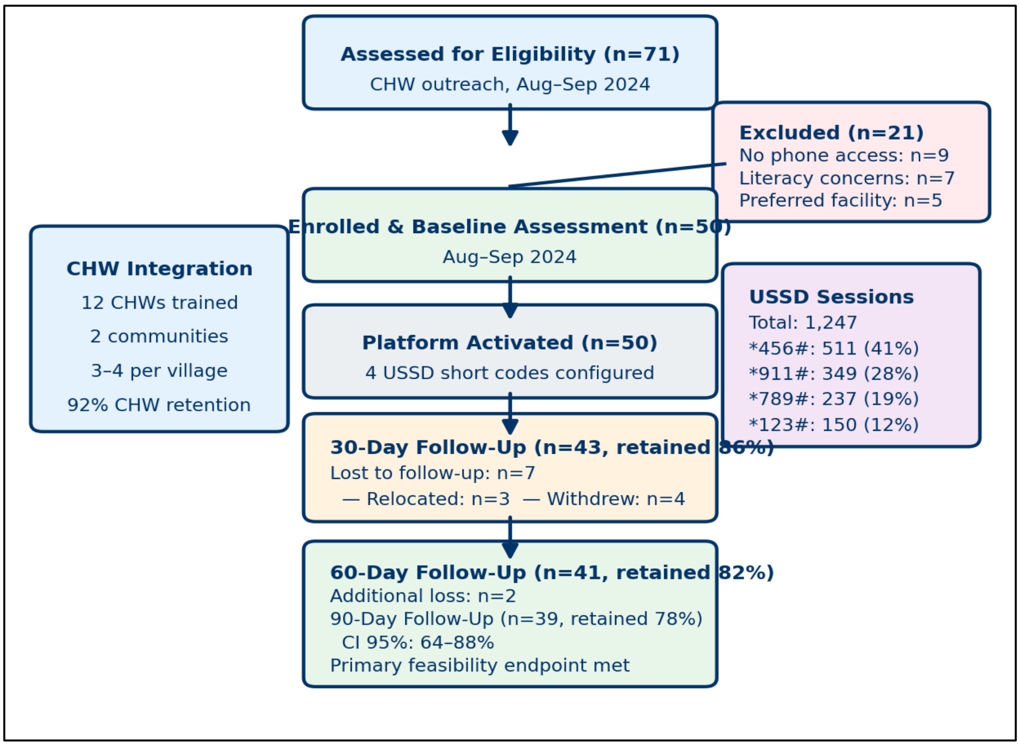

Fig 5 CONSORT-Adapted Participant Flow Diagram, from Eligibility Assessment (n=71) through 90-day Primary Endpoint (n=39 Retained, 78%).

➢ *Intervention Description (TIDieR)*

Name: HealthDrive USSD Telehealth Platform. Rationale: Reduce primary care access barriers in rural SSA without requiring internet or a smartphone. Materials: Four USSD short codes, Python/PostgreSQL backend, USSD gateway. Procedure: Patients dial the relevant code; branching menus (maximum 5 levels, 180-second timeout) guide symptom reporting; the system generates triage recommendations and dispatches CHW alerts for high-acuity cases. Providers: 12 trained CHWs. Mode: Remote (mobile phone, 2G signal). Duration: 19-month pilot; 90-day per-patient follow-up. Tailoring: The session timeout threshold and menu depth were adjusted in Week 4 based on CHW feedback regarding elderly users' navigation difficulties. Modifications documented in CHW supervision records.

➢ *Analytical Frameworks*

CFIR domains were assessed through structured CHW interviews at baseline, 6 weeks, and 3 months using the LMIC-adapted CFIR Interview Guide [19]. Patient satisfaction was measured at 30, 60, and 90 days using a 10-item Likert scale validated for low-literacy populations. RE-AIM dimensions were operationalised quantitatively (Reach, Adoption, Implementation, Maintenance) and qualitatively (Effectiveness supplemented by satisfaction data). TAM-RLS constructs were assessed through semi-structured interviews with 15 patients (maximum variation by age, gender, and health condition) conducted in the local language by trained CHW research assistants, and analysed using the framework thematic method [4].

V. RESULTS

➢ *Enrollment and Reach*

Of 71 eligible patients approached through CHW outreach, 50 enrolled (70.4% uptake). Non-enrollment reasons: no personal phone access (n=9, 12.7%), literacy concerns about menu navigation (n=7, 9.9%), preference for facility-based care (n=5, 7.0%). Participant profile: mean age 41.2 years (range 19–74), 54% female (n=27), 38% elderly aged 55+ (n=19). Presenting conditions: respiratory (32%), musculoskeletal pain (24%), maternal health (18%), general wellness (16%), other (10%).

➢ *Retention, Primary Outcome*

Figure 6 presents the 90-day retention curves. Three-month retention of 78% (n=39/50, 95% CI: 64–88%) substantially exceeds the SSA mHealth self-managed benchmark of 45–65% at 90 days [21]. Retention in the elderly subgroup was 89% (n=17/19). CHW platform retention was 92% (11/12). These results are consistent with prior evidence that CHW-mediated mHealth substantially outperforms self-managed mHealth on retention by addressing downstream barriers to care engagement [22]. Confidence intervals were calculated using the Wilson score method, which is appropriate for small pilot samples.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

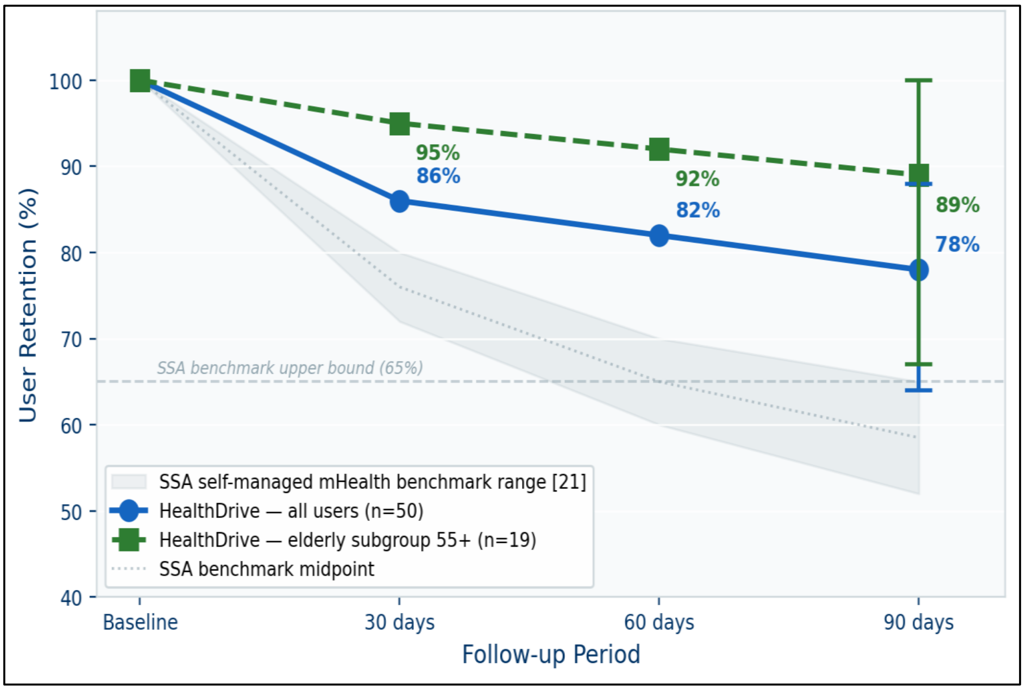

Fig 6 HealthDrive 90-Day Retention Curve vs SSA mHealth Benchmark Range. Error Bars at 90-Day Endpoint Represent 95% Confidence Intervals (Wilson Score Method).

Table 2 User Retention Across Follow-Up Periods. Platform Utilisation

Follow-Up Period	Retained (n)	Rate	95% CI
30 days	43/50	86%	73–94%
60 days	41/50	82%	69–91%
90 days (primary)	39/50	78%	64–88%
Elderly 55+ (90d)	17/19	89%	67–99%
CHW platform use	11/12	92%	62–100%

The platform generated 1,247 USSD sessions across 19 months. Session distribution: *456# symptom classification, 41% (n=511); *911# emergency triage, 28% (n=349); *789# CHW coordination, 19% (n=237); *123# clinic navigation, 12% (n=150). Emergency triage session completion was 71% (248/349). The 29% non-completion rate (101 sessions) was primarily due to session timeouts (>180 seconds between keypresses), concentrated among elderly users attempting multi-symptom reporting. Median session duration was 67 seconds (IQR: 42–103 seconds). Figure 7A shows session distribution.

➢ *Satisfaction and TAM-RLS Analysis*

Overall patient satisfaction was 7.8/10 (SD: 1.4). Elderly users rated satisfaction at 8.5/10 (SD: 0.9). TAM-RLS thematic analysis identified three dominant themes: (1) high perceived usefulness, particularly for chronic disease monitoring and CHW-triggered follow-up, with 14/15 interview participants reporting reduced transport expenditure as a primary benefit; (2) ease-of-use barriers for multi-symptom presentations, where the sequential menu structure created navigation difficulty; and (3) confidentiality concerns raised by 8/15 participants regarding shared-phone privacy, a TAM-RLS-specific barrier absent from standard TAM frameworks [4]. Figure 9 maps these findings to the full TAM-RLS conceptual model.

➢ *Cost and Sustainability*

Total cash expenditure over 19 months was $2,580 USD at $125/month (Figure 6B). The estimated researcher time at market-rate software development rates was approximately $12,000 in sweat equity. Cost per enrolled user over 19 months: $51.60 cash. The B2B sustainability model, clinic subscriptions at $200/month/facility plus $0.50 USSD micro-payments, projects to full cost recovery at 1 subscribing health facility.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

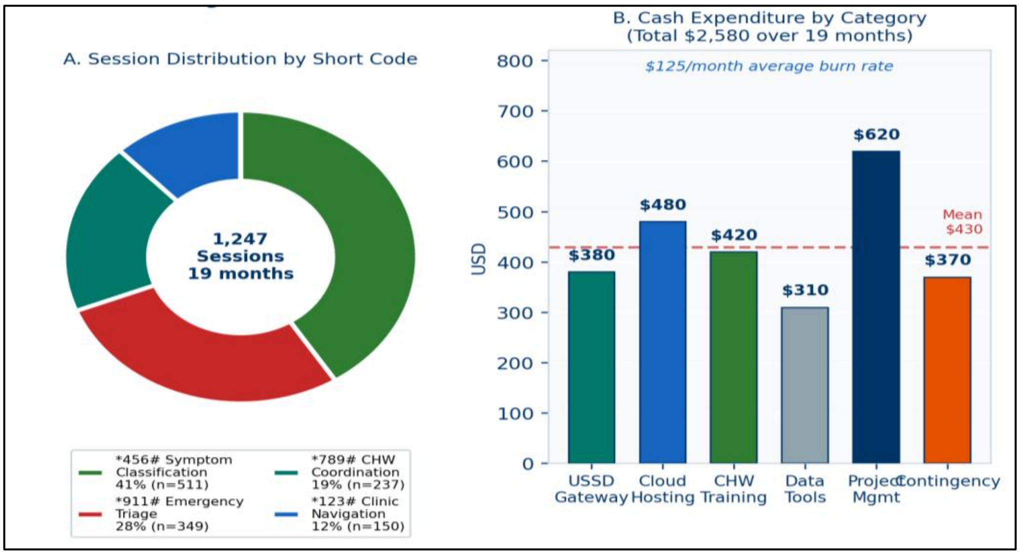

Fig 7 A: USSD Session Distribution by Short Code (n=1,247 total). B: 19-Month Cash Expenditure by Category (Total $2,580; Dashed Line = Category Mean $430).

➢ *RE-AIM Summary*

Table 3 RE-AIM Feasibility Assessment

Dimension	Indicator	Finding	Assessment
Reach	% eligible enrolled	70.4% (50/71)	Feasible
Effectiveness	90-day retention	78% (CI: 64–88%)	Exceeds SSA benchmark
Adoption	CHW platform use rate	92% (11/12)	Feasible
Implementation	Triage completion	71%	Acceptable
Maintenance	Intent to continue	87% (13/15)	Feasible
Cost	Monthly burn	$125 USD/month	Sustainable

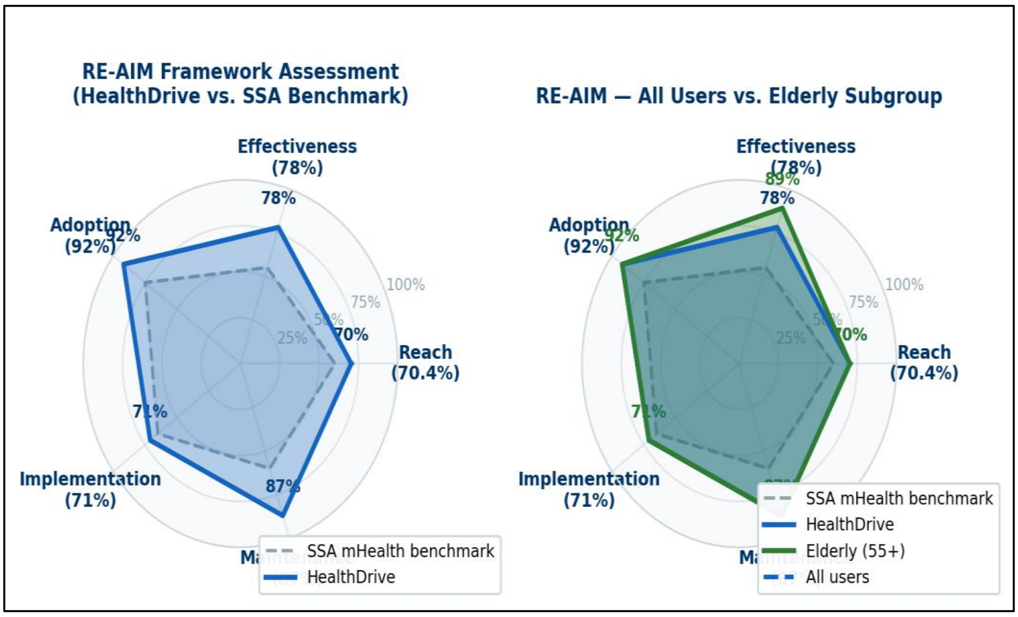

Fig 8 RE-AIM Radar Charts: Left, HealthDrive vs SSA mHealth Benchmark; Right, all Users vs Elderly Subgroup. Values Represent Scores on a 0 –100% Feasibility Scale.

VI. DISCUSSION

➢ *Principal Findings*

HealthDrive demonstrates that USSD-based telehealth with CHW integration is feasible, acceptable, and cost-efficient in rural SSA. The 78% 90-day retention rate exceeds comparable self-managed mHealth programs in SSA by 13–33 percentage points [21], consistent with prior evidence that CHW-mediated digital health addresses the downstream barriers described in TAM-RLS [4]. The 85% satisfaction rate among elderly users, the population with the highest health burden and the lowest digital literacy, challenges the

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

prevailing assumption that USSD interfaces are unsuitable for older populations [29].

➢ *The Babyl Rwanda Paradox and Structural Resilience*

Babyl Rwanda achieved clinical effectiveness metrics that HealthDrive has not yet measured: reduced out-of-pocket costs, shorter consultations, and more complete diagnostic questioning [2,17,18]. HealthDrive does not claim clinical superiority. The argument is structural: a locally owned, CHW-integrated USSD platform at $125/month cannot collapse due to the bankruptcy of a London-headquartered digital health corporation. Structural resilience, the capacity of a health system to maintain core functions through shocks [30], requires that critical health infrastructure not depend on external financing chains that can be severed without warning. Figure 9 illustrates that HealthDrive is the only platform in the comparison set that is simultaneously locally owned, CHW-integrated, internet-free, and currently operational.

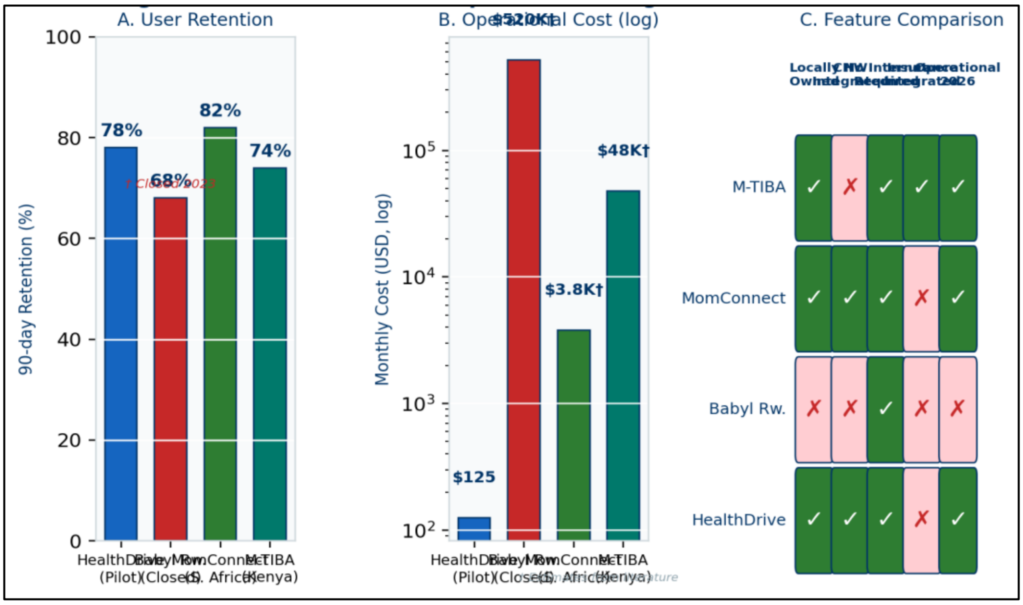

Fig 9 HealthDrive vs Comparable SSA digital Health Platforms: A) 90-Day User Retention; B) Monthly Operational Cost (Log Scale); C) Feature Comparison Matrix (✓ = Criterion Met).

➢ *Sustainability and Scale-Up Pathway*

Figure 10 presents the four-phase sustainability roadmap. MomConnect (South Africa), government-owned and completing its 10th operational year in 2024, survived multiple political transitions and a global pandemic by embedding within the national health infrastructure [15]. Community-based health insurance schemes operating across SSA, Rwanda's Mutuelle de Santé scheme covers 85.3% of the population as of the EICV7 2024 national survey, providing potential reimbursement pathways if USSD triage sessions are recognised as billable health interactions under national insurance frameworks [15, EICV7 2024]. The WHO Global Strategy on Digital Health 2020–2025 explicitly calls for insurance integration as a mechanism to sustain mHealth [31].

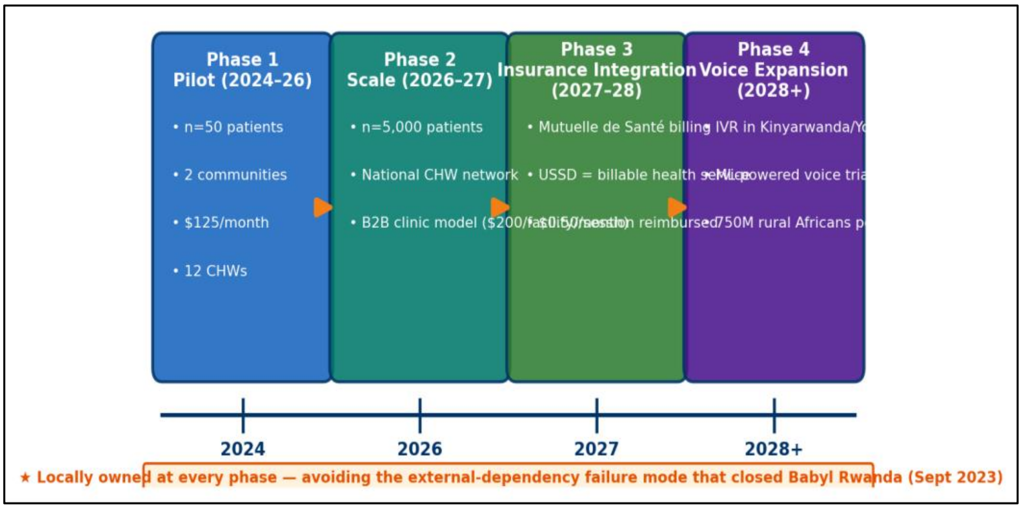

Fig 10 HealthDrive Sustainability Pathway and Scale-up Roadmap Across four Phases (2024–2028+), Anchored by Local Ownership at Every Phase.

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

➢ *Limitations*

This study has six limitations. First, n=50 limits the precision of the estimate (95% CI: 64–88%). Second, the single-arm design prevents causal inference compared to unassisted care-seeking. Third, clinical outcomes were not measured; no health status comparison is possible. Fourth, two communities in one country do not represent the diverse rural population contexts of SSA. Fifth, the 29% non-completion rate for emergency triage sessions requires UX intervention, particularly for elderly users. Sixth, conflict-of-interest risk from researcher-conducted evaluation was managed through an independent data extraction protocol, but not eliminated.

➢ *Computational Research Challenges and Machine Learning Directions for Low-Resource Health Signal Processing*

HealthDrive's USSD text-based interface is a transitional model. The natural next frontier is voice-based IVR interaction in local SSA languages. Patients call a health line, describe symptoms in Yoruba, Hausa, Igbo, Kinyarwanda, or Swahili, and an AI-mediated triage system processes their input. The authors' parallel climate advisory project (Clisense) has deployed IVR-based voice interaction in three West African languages, establishing technical feasibility. Five open machine learning and signal processing problems block the responsible deployment of this voice extension at the population scale across SSA.

- *Robust ASR for 2G Health Channels in Low-Resource SSA Languages*

Voice-based IVR for rural SSA health operates over 2G GSM networks, applying Adaptive Multi-Rate (AMR) codec compression: 8kHz sampling rate, 12.2 kbps maximum bitrate, with additional degradation due to rural tower distances and mobile movement (Table I). Current speech foundation models (Whisper, wav2vec2-XLSR) are trained predominantly on high-fidelity English and high-resource language audio; word error rates increase substantially under 2G constraints, and most SSA health languages have essentially no labelled health audio corpus. Research into channel simulation-based domain adaptation, health-vocabulary language model fusion for low-resource languages, and knowledge distillation from foundation models to resource-constrained local ASR systems [32,33] is needed. The Speech Robust Bench framework [32] provides a robustness evaluation standard extendable to 2G AMR codec degradation in SSA health audio contexts.

- *Privacy-Preserving Processing of Voice Health Data*

Voice health recordings encode demographics, emotional state, respiratory patterns, and disease signals beyond explicit health queries. In rural SSA contexts where stigma surrounding mental health, HIV/AIDS, and reproductive health remains acute, an IVR health system permitting demographic inference from audio embeddings exposes patients to risks disproportionate to clinical benefit. In small rural communities, voice-based identity inference is effectively de-anonymisation. Privacy-preserving frameworks, acoustic anonymisation that preserves health-relevant signal content, federated learning architectures, and differential privacy for voice embeddings [34,35] are needed before voice health USSD meets the ethical standard that HealthDrive's text-only USSD model already satisfies.

- *Health Biomarker Detection Through Degraded Communication Channels*

CHW interviews identified high-acuity presentations partly flagged through patient voice characteristics: laboured breathing, vocal weakness consistent with high fever, and prosodic disruption from severe pain or psychological distress. A voice triage system could passively detect these acoustic biomarkers alongside explicit symptom reporting. The challenge: 2G AMR codec removes high-frequency components above 4kHz that carry significant diagnostic information in respiratory and tremor biomarkers. Prior work on dyspnea detection in degraded audio from human-robot interaction scenarios [36] and respiratory distress estimation from degraded speech [37] establishes methodological foundations. Voice-based age prediction under channel degradation [38] demonstrates that health-relevant biomarker signals survive codec constraints and remain retrievable by ML methods.

- *Semi-Supervised Triage Classification from Sparse Labelled Health Interaction Logs*

HealthDrive generated 1,247 USSD sessions from 50 patients; only a fraction carry verified clinical outcome labels from CHW follow-up records. At the pilot scale, this proportion can be recovered through field follow-up. At 50,000 users across SSA, a plausible scale for a nationally integrated USSD health platform, manual outcome labelling becomes operationally impossible. Unified semi-supervised learning benchmarks [39] and negative sampling approaches for weak-label learning [40] provide initial methodological frameworks that require adaptation to the specific structural properties of USSD health interaction logs: sequential, sparse, highly structured, and temporally ordered. Foundation model embeddings of USSD session sequences may provide the latent structure needed for effective pseudo-labelling when fewer than 20% of sessions carry confirmed outcome labels.

- *Hallucination-Resistant Voice AI for High-Stakes Clinical Triage*

Extending HealthDrive toward speech foundation model-mediated voice triage introduces a failure mode with no acceptable tolerance. A model that mishears a patient's report of chest pain and routes the patient to the health information line, rather than emergency triage, produces a clinically indistinguishable error that may cause patient harm. Hallucination in speech foundation models is driven substantially by distribution shift between training and inference conditions [41]. The shift between large speech model training distributions (English-dominant, high-fidelity, web-scraped audio) and SSA health inference conditions (2G-degraded, low-resource language audio, medical domain) represents one of the most extreme distributional gaps in any deployed speech AI application. Hallucination measurement frameworks, retrieval-augmented generation constrained to verified health vocabulary, and inference-time

ISSN No: -2456-2165 https://doi.org/10.38124/ijisrt/26apr2523

distribution shift monitoring are needed before voice AI can be responsibly deployed in SSA health triage pathways [41].

VII. CONCLUSION

HealthDrive demonstrates that USSD-based telehealth with community health worker integration is feasible, acceptable, and cost-efficient in rural sub-Saharan Africa. The platform achieved 78% 90-day user retention, 85% elderly satisfaction, and full operational sustainability at $125/month, metrics comparing favourably with comparable mHealth programs across SSA. More importantly, the platform embodies a structural design principle: locally owned digital health infrastructure integrated with existing CHW networks and funded within realistic national health budgets does not face the catastrophic external dependency risk that terminated Babyl Rwanda's service to 2.8 million users in September 2023.

Three evidence-based policy recommendations follow. First, national health ministries across SSA should prioritise locally owned USSD-based platforms over externally funded digital health corporations, using MomConnect and analogous government-integrated tools as sustainability models. Second, CHW integration should be a mandatory design requirement for digital health interventions in rural SSA, given consistent evidence that CHW-mediated mHealth substantially outperforms self-managed alternatives on retention. Third, insurance integration pathways, recognition of USSD triage sessions as billable health interactions under national health insurance schemes, should be established during pilot phases rather than deferred to scale.

The five computational challenges identified in Section VII define the research frontier between the USSD text model demonstrated here and a voice-based model capable of reaching the 750 million rural Africans currently excluded by smartphone-dependent health AI alternatives. These challenges sit at the intersection of machine learning for signal processing, health informatics, and low-resource natural language processing, an interdisciplinary research agenda with direct humanitarian consequences for SSA primary healthcare delivery.

ACKNOWLEDGEMENTS

The author thanks the community health workers and patients who participated in this study and whose trust made the pilot possible. The author thanks the management of Society for Family Health Rwanda (SFH Rwanda) for institutional guidance on rural health system dynamics. The author thanks Dr Bhiksha Ramakrishnan (Professor, Language Technologies Institute, Carnegie Mellon University) for mentorship and guidance that shaped the computational research directions identified in Section VII. The author thanks Nji Williams, mentor, and Laura Mugambi (Africa Health Collaborative, African Leadership University) for professional guidance throughout the project development process.

REFERENCES

[1]. Asamani JA, Bediakon KSB, Boniol M, Munga'tu JK, et al. Projected health workforce requirements and shortage for addressing the disease burden in the WHO Africa Region, 2022–2030: a needs-based modelling study. BMJ Glob Health. 2024 Oct 22;7(Suppl 1):e015972. doi:10.1136/bmjgh-2024-015972. PMC11789529.

[2]. Rubuga FK, et al. Telemedicine implementation and healthcare utilisation in Rwanda: interrupted time series of Babyl Digital Health Services from 2015 to 2024. BMC Prim Care. 2026;27:45. doi:10.1186/s12875-026-03179-8. PMC12879403.

[3]. 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.

[4]. Campbell JI, Aturinda I, Mwesigwa E, et al. The Technology Acceptance Model for Resource-Limited Settings (TAM-RLS): a novel framework for mobile health interventions targeted to low-literacy end-users in resource-limited settings. AIDS Behav. 2017;21(11):3129–3140. doi:10.1007/s10461-017-1765-0. PMC5763567.

[5]. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health. 1999;89(9):1322–1327. doi:10.2105/AJPH.89.9.1322.

[6]. Ahmat A, Okoroafor SC, Kazanga I, Asamani JA, et al. The health workforce status in the WHO African Region: findings of a cross-sectional study. BMJ Glob Health. 2022;7(Suppl 1):e008317. doi:10.1136/bmjgh-2021-008317. PMID:35675966.

[7]. Ministry of Health, Republic of Rwanda. Rwanda Health Sector Performance Report 2017–2019. Kigali: MOH; 2020.

[8]. WHO. Republic of Rwanda Ministry of Health Health Labour Market Analysis Report. Geneva: WHO; 2020.

[9]. Condo J, Mugeni C, Naughton B, et al. Rwanda's evolving community health worker system: a qualitative assessment of client and provider perspectives. Hum Resour Health. 2014;12:71. doi:10.1186/1478-4491-12-71.

[10]. Rwanda Utilities Regulatory Authority (RURA). Mobile-Cellular Telephone Subscriptions Report September 2024. Kigali: RURA; 2024.

[11]. Maliwichi P, Chigona W, Sowon K. Appropriation of mHealth interventions for maternal health care in Sub-Saharan Africa: hermeneutic review. JMIR mHealth uHealth. 2021;9(10):e22653. doi:10.2196/22653.

[12]. McKinsey & Company. Overcoming the health workforce paradox in sub-Saharan Africa. McKinsey Global Institute; 2024.

[13]. Eze P, Lawani LO, Agu UJ, Acharya Y. Catastrophic health expenditure in sub-Saharan Africa: a systematic review and meta-analysis. Bull World Health Organ.

2022;100(5):337–351. doi:10.2471/BLT.21.287673. PMC9047424.

[14]. Owoyemi A, Owoyemi J, Osiyemi A, Boyd A. Digital solutions for community and primary health workers: lessons from implementations in Africa. Front Digit Health. 2022;4:876957. doi:10.3389/fdgth.2022.876957.

[15]. National Department of Health, South Africa. MomConnect 10th anniversary programme overview. Pretoria: NDoH; 2024. Reported: IOL Saturday Star, 28 Sep 2024.

[16]. Huisman L, van Duijn SMC, Silva N, et al. A digital mobile health platform increasing efficiency and transparency towards universal health coverage in LMICs. Digit Health. 2022;8:20552076221092213. doi:10.1177/20552076221092213.

[17]. Nkurunziza T, et al. Digital Primary Health in Rwanda: Qualitative Study of User Experiences and Implementation Lessons From Babyl's Telemedicine Platform. J Med Internet Res. 2026;28(1):e84832. doi:10.2196/84832.

[18]. ICTworks. Babyl Paradox: When Evidence-Based Digital Success Cannot Beat Corporate Failure. 2024. Available: https://www.ictworks.org

[19]. Means AR, Kemp CG, Gwayi-Chore MC, et al. Evaluating and optimising the Consolidated Framework for Implementation Research (CFIR) for use in low- and middle-income countries: a systematic review. Implement Sci. 2020;15:17. doi:10.1186/s13012-020-0977-0.

[20]. Eysenbach G. The law of attrition. J Med Internet Res. 2005 mar 31;7(1):e11. doi:10.2196/jmir.7.1.e11. PMID:15829473. PMCID:1550631.

[21]. Yang S, Wu Q, Huang Z, et al. Understanding inequalities in mobile health utilisation across phases: systematic review and meta-analysis. J Med Internet Res. 2025;27:e71349. doi:10.2196/71349.

[22]. Kachimanga C, Sinyiza F, Msosa A, et al. Evaluating the adoption of mHealth technologies by community health workers to improve the use of maternal health services in Sub-Saharan Africa: systematic review. JMIR mHealth uHealth. 2024;12:e55819. doi:10.2196/55819. PMID:39316427.

[23]. Palinkas LA, Horwitz SM, Weisz JR, et al. Mixed method designs in implementation research. Admin Policy Mental Health. 2011;38(1):44–53. doi:10.1007/s10488-010-0314-z.

[24]. Pinnock H, Barwick M, Carpenter CR, et al. Standards for Reporting Implementation Studies (StaRI) Statement. BMJ. 2017;356:i6795. doi:10.1136/bmj.i6795. PMID:28264797.

[25]. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ. 2014;348:g1687. doi:10.1136/bmj.g1687. PMID:24609605.

[26]. Julious SA. A sample size of 12 per group is the rule of thumb for a pilot study. Pharm Stat. 2005;4(4):287–291. doi:10.1002/pst.185.

[27]. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health. 2008;31(2):180–191. doi:10.1002/nur.20247.

[28]. Eldridge SM, Chan CL, Campbell MJ, et al. CONSORT 2010 statement: extension to randomised pilot and feasibility trials. BMJ. 2016;355:i5239. doi:10.1136/bmj.i5239. PMID:27777223.

[29]. Bertolazzi A, Quaglia V, Bongelli R. Barriers and facilitators to health technology adoption by older adults with chronic diseases: integrative systematic review. BMC Public Health. 2024;24:506. doi:10.1186/s12889-024-18036-5.

[30]. Kruk ME, Gage AD, Arsenault C, et al. High-quality health systems in the Sustainable Development Goals era: time for a revolution. Lancet Glob Health. 2018;6(11):e1196–e1252. doi:10.1016/S2214-109X(18)30386-3.

[31]. WHO. Global Strategy on Digital Health 2020–2025. Geneva: WHO; 2021. Available: apps.who.int/iris/handle/10665/344249.

[32]. Shah MA, Raj B. Speech Robust Bench: A Robustness Benchmark For Speech Recognition. International Conference on Learning Representations (ICLR). 2025.

[33]. Waheed A, Kadaoui K, Raj B, Abdul-Mageed M. uDistil-Whisper: Label-Free Data Filtering for Knowledge Distillation in Low-Data Regimes. In: Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers). 2025:5750–5767. Albuquerque, New Mexico: ACL.

[34]. Teixeira F, Abad A, Raj B, Trancoso I. Privacy-Preserving Automatic Speaker Diarization. Proc ICASSP 2023. doi:10.1109/ICASSP49357.2023.10094969.

[35]. Teixeira F, Abad A, Raj B, Trancoso I. Privacy-oriented manipulation of speaker representations. arXiv:2310.06652. 2023.

[36]. Alvarado E, Grádeda N, Luzanto A, et al. Automatic Detection of Dyspnea in Real Human–Robot Interaction Scenarios. Sensors. 2023;23(3):1257. doi:10.3390/s23031257.

[37]. Alvarado E, Grádeda N, Luzanto A, et al. Respiratory Distress Estimation in Human-Robot Interaction Scenario. Proc Interspeech. Dublin, 2023.

[38]. Alharthi D, Zamani M, Raj B, Singh R. Tessellated Linear Model for Age Prediction from Voice. Proc ICASSP. 2025:1–4.

[39]. Wang Y, Chen H, Fan Y, Sun W, Tao R, Hou W, et al. USB: A Unified Semi-Supervised Learning Benchmark for Classification. Advances in Neural Information Processing Systems 35 (NeurIPS 2022) Datasets and Benchmarks Track. doi:10.48550/arXiv.2208.07204.

[40]. Shah A, Tang F, Ye Z, Singh R, Raj B. Importance of Negative Sampling in Weak Label Learning. Proc ICASSP 2024.

[41]. Atwany H, Waheed A, Singh R, Choudhury M, Raj B. Lost in Transcription, Found in Distribution Shift: Demystifying Hallucination in Speech Foundation

Models. In Findings of the Association for Computational Linguistics: ACL 2025. 2025:23181–23203. Vienna, Austria. doi:10.18653/v1/2025.findings-acl.1190.
[42]. GBD 2019 Human Resources for Health Collaborators. Measuring the availability of human resources for health. Lancet. 2022;399(10341):2129–2154. doi:10.1016/S0140-6736(22)00532-3.
[43]. National Institute of Statistics of Rwanda (NISR). Integrated Household Living Conditions Survey 7 (EICV7) 2023–2024. Kigali: NISR; 2025. [Health insurance coverage: 85.3%].
[44]. Damschroder LJ, Reardon CM, Widerquist MAO, Lowery J. The updated CFIR is based on user feedback. Implement Sci. 2022;17:75. doi:10.1186/s13012-022-01245-0.