South American Publishing

SAP Social AI. 2025; 1:15
doi: 10.62486/sai202515

ORIGINAL

AI-Augmented Mental Health Care in Sahelian Herder-Farmer Corridors: Ethical and Scalable Approaches via Pastoral Data Trusts

Atención Integral de Salud Mental Aumentada por IA en los Corredores Pastoralistas del Sahel: Enfoques Éticos y Escalables mediante Fideicomisos de Datos Pastorales

Ifedayo Grace Malachi[1,2], Olanrewaju O. Akinola[3], Ijeoma Chioma Mordi[4], Ngozi Blessing Umoru[5], Sandra A Palmer[6]

[1]Academy of International Affairs NRW. Bonn Germany.
[2]Bamidele Olumilua University of Education Science and Technology. Ikere-Ekiti, Nigeria.
[3]Department of Information Technology and Intellectual Property Law. University of Sussex, U.K
[4]Department of Information, Intellectual Property Law, University of Lagos. Nigeria.
[5]Department of Social Science Education, University of Nottingham. Nottingham, U.K.
[6]Department of Social Science Education, Leading Learning & Teaching, University of Dundee. U.K.

Cite as
Malachi IG, Akinola OO, Mordi IC, Blessing Umoru N, Palmer SA. AI-Augmented Mental Health Care in Sahelian Herder-Farmer Corridors: Ethical and Scalable Approaches via Pastoral Data Trusts. SAP Social AI. 2025; 1:15. https://doi.org/10.62486/sai202515

Publication History
Submitted: 23-03-2025
Revised: 14-06-2025
Accepted: 08-08-2025
Published: 09-08-2025

Corresponding Author
Ifedayo Grace Malachi

ABSTRACT

Introduction: mental-health services are scarce in herder–farmer conflict zones of Nigeria's Middle Belt, where PTSD and depression are widespread. Artificial intelligence (AI) could extend care, but deployment in low-resource, multilingual, and socially complex settings poses challenges of trust, bias, and cultural fit.
Objective: to assess the effectiveness, acceptability, and equity of an AI-augmented Community Health Extension Worker (CHEW) program-ntegrated with a Pastoral Data Trust (PDT) and culturally grounded pedagogy-n improving mental-health outcomes and reducing intergroup prejudice.
Method: a 12-month, open-label, cluster-randomised trial was conducted in 60 villages across Plateau and Benue States, enrolling 1274 adults with elevated symptoms (PHQ-9 ≥10 or PC-PTSD ≥3). Villages were randomised 1:1 to receive the AI-CHEW intervention or enhanced usual care. Primary outcomes were treatment coverage and symptom remission; secondary outcomes included intergroup prejudice, therapeutic alliance, and cost-effectiveness. Mixed-methods components—algorithmic audits, co-design sessions, and qualitative interviews—assessed acceptability and model behavior.
Results: the intervention increased treatment coverage (48,7 % vs. 27,9 %; RR 1,74, 95 % CI 1,51–2,01) and improved remission for PTSD (46 % vs. 31 %) and depression (51 % vs. 36 %). Intergroup prejudice decreased (β = −0,36 SD), and therapeutic alliance was stronger (Cohen's d = 0,44). Algorithmic sensitivity was higher among urban and literate users, but federated retraining and CHEW mediation reduced bias. Cost-effectiveness analysis showed an incremental cost of USD 312 per additional recovered case. Qualitative data indicated high acceptability when pedagogy aligned with communal norms.
Conclusions: AI-augmented CHEWs can safely and cost-effectively expand mental-health care in conflict-affected settings while supporting social cohesion. Scalable implementation requires contextual adaptation, ethical oversight, and inclusive community engagement.

Keywords: AI-Mental Health; Community Health Extension Workers (CHEWs); Herder-Farmer Conflict; Pastoral Data Trust; PTSD; Ethnic Prejudice.

RESUMEN

Introducción: los servicios de salud mental son en gran medida inaccesibles en las zonas de conflicto entre pastores y agricultores del Cinturón Medio de Nigeria, donde el TEPT y la depresión son comunes. La inteligencia artificial (IA) puede ampliar la atención, pero su implementación en contextos multilingües y de bajos recursos plantea desafíos de confianza, sesgo y adecuación cultural.
Objetivo: evaluar la efectividad, aceptabilidad y equidad de un programa de Trabajadores de Extensión en Salud Comunitaria (CHEW) aumentados con IA, integrado en un Fideicomiso de Datos Pastorales (PDT) y una pedagogía culturalmente sensible.
Método: se llevó a cabo un ensayo aleatorizado por conglomerados, abierto y de 12 meses, en 60 aldeas de Plateau

© 2025; The Authors. This Open access article is distributed under the Creative Commons Attribution 4.0 license (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

y Benue, con 1274 adultos que presentaban síntomas elevados (PHQ-9 ≥10 o PC-PTSD ≥3). Las aldeas fueron asignadas 1:1 a la intervención AI-CHEW o a atención habitual mejorada. Los resultados primarios fueron cobertura de tratamiento y remisión de síntomas; los secundarios incluyeron prejuicio intergrupal, alianza terapéutica y costo-efectividad. Un enfoque de métodos mixtos incorporó auditorías algorítmicas, co-diseño e entrevistas cualitativas.

Resultados: la intervención aumentó la cobertura (48,7 % vs. 27,9 %; RR 1,74, IC 95 % 1,51–2,01) y elevó las tasas de remisión de TEPT (46 % vs. 31 %) y depresión (51 % vs. 36 %). El prejuicio intergrupal disminuyó (β = −0,36 SD) y la alianza terapéutica fue mayor (d = 0,44). Aunque la sensibilidad algorítmica fue superior entre usuarios urbanos y alfabetizados, el reentrenamiento federado y la mediación de CHEWs redujeron sesgos. El ICER fue de USD 312 por caso recuperado. La aceptabilidad fue alta cuando la pedagogía se alineó con normas comunitarias.

Conclusiones: los CHEWs aumentados con IA pueden expandir de manera segura y costo-efectiva la atención en salud mental en zonas afectadas por conflictos, promoviendo cohesión social. La escalabilidad requiere adaptación contextual, supervisión ética y participación inclusiva.

Palabras clave: Salud Mental con IA; Agentes Comunitarios de Salud; Conflicto Ganadero-Agrícola; Fideicomiso de Datos Pastorales; Trastorno por Estrés Postraumático; Prejuicio Étnico.

INTRODUCTION

Across the semi-arid belts of West and Central Africa, herder–farmer tensions have hardened into recurrent cycles of violence, with profound psychological consequences for both pastoralist and agrarian communities. Decades of displacement, livestock loss, and chronic insecurity have produced a high burden of trauma, yet mental-health infrastructure remains virtually nonexistent in these rural zones.[1,2] In Nigeria, over 80 % of individuals requiring psychological care receive no treatment, with access gaps most pronounced in conflict-affected and mobile pastoral regions.[3,4] National surveys systematically exclude pastoral populations; for example, the 2019 National Mental Health Survey contained no enumeration areas within 20 km of major cattle routes.[5,6,7] Similarly, the only publicly available depression-prevalence estimate for the region (12,5 %) excluded transient herders and seasonally displaced households.[8] In short, the populations most exposed to trauma are largely invisible in official statistics.[2,5,9]

This "evidence void" has generated interest in low-bandwidth Artificial Intelligence (AI) tools capable of extending mental-health screening and support where clinicians are scarce. Mobile-phone penetration exceeds 90 % in many pastoral regions,[10] making automated triage chatbots, voice-note classifiers, and culturally adapted cognitive-behavioural therapy (CBT) apps feasible.[11,12] Evidence from low- and middle-income countries suggests that AI-enabled screening can increase case identification, reduce service costs, and deliver psychoeducation in local languages.[13,14,15,16]

Yet these opportunities coexist with significant risks. Algorithms trained on urban, Anglophone, or Western corpora routinely misclassify African idioms and culturally embedded emotional expressions, while language varieties in the region exhibit some of the highest NLP error rates globally.[17,18,19] Misalignment can erode trust, reinforce stigma, and produce harmful clinical consequences.[20,21] Ethical constraints are equally critical. Pastoral communities often interpret data collection through historical experiences of land appropriation, taxation, and coercive enumeration.[22,23,24] Herding corridors, clan genealogies, and cattle inventories are politically sensitive, and leaks can expose households to retaliation or resource grabs.[25,26,27] In contexts where women's voices and images are socially regulated, accidental disclosure may trigger stigma or sanctions.[28,29] Without robust governance, humanitarian AI risks reproducing the asymmetries that have shaped land-use conflict and rural surveillance.[30,31,32,33]

Implementation also hinges on human intermediaries. Community Health Extension Workers (CHEWs), often with limited literacy, serve as "AI explainers," interpreting outputs for elders, negotiating consent with nomadic authorities, and recognizing when chatbot recommendations violate communal norms.[34,35,36] These challenges underscore the need for culturally grounded governance frameworks, linguistic adaptation, and pedagogies of explainability to maintain trust and therapeutic fidelity.[18,19,37,38]

This study employed an explanatory sequential mixed-methods design. An initial qualitative co-design phase-drawing on ethnography, participatory theatre, storytelling nights, and iterative model auditing-developed the intervention, the Pastoral Data Trust (PDT), and the CHEW pedagogy for algorithmic interpretation. Insights from this phase informed the creation of culturally congruent AI tools, governance norms, and consent procedures. Subsequently, a pragmatic cluster randomized controlled trial (cRCT) evaluated the effectiveness of this model in improving mental-health access, equity, and intergroup cohesion.

Research Question: how can AI expand mental-health access in herder–farmer corridors without reinforcing ethnic boundaries, eroding therapeutic empathy, or shifting pedagogical authority to opaque cloud infrastructures?

Objectives

1. Develop and test a conflict-sensitive conceptual framework mapping AI opportunities onto the ecological, linguistic, and governance contours of mobile pastoralism.[2,6,9,24]
2. Design and operationalize an ethical matrix translating global bioethical principles-autonomy, beneficence, non-maleficence, and justice-into pastoral idioms of consent, reciprocity, and restitution.[30,31,37]
3. Evaluate the effectiveness of a pedagogical toolkit equipping CHEWs as "algorithmic interpreters" capable of auditing AI outputs, facilitating culturally grounded role-plays, and preserving therapeutic alliance.[19,36,38]

Hypothesis

We tested the hypothesis that AI-augmented CHEWs, supported by PDT governance and culturally grounded pedagogy, improve

mental-health outcomes, treatment coverage, and intergroup cohesion compared with enhanced usual care, as assessed in a cluster RCT.

Key concepts

As shown in Table A1 include pastoral ethics, norms of care, reciprocity, authority, and harm reduction;[39,40,41] data trusts, community-governed systems for managing rights, deletion, and access;[42,43,44] and explainability pedagogy, training CHEWs to contextualise and translate AI outputs in alignment with clan-based epistemologies.[18,19,45] Together, these constitute the foundation for deploying culturally aligned AI tools in high-risk, low-data environments.

Article Structure

This article first presents the empirical context and initial algorithmic failures that motivated the need for a pastoral-ethics approach to AI-supported mental-health care (Section 2). Section 3 details the development of the Pastoral Data Trust and the pedagogy for CHEWs as algorithmic interpreters. Section 4 presents findings from the randomized trial assessing the intervention's impact on mental-health access, algorithmic equity, and intergroup cohesion. Section 5 discusses scalability, policy implications, and safeguards for responsible deployment.

METHOD

Study Design

This study employed an explanatory sequential mixed-methods design, beginning with an extensive qualitative co-design phase that informed the subsequent pragmatic cluster randomized controlled trial (cRCT).[1,2,36] The qualitative component drew on ethnography, participatory theatre, narrative inquiry, and structured co-design workshops to develop the key elements of the intervention, including the CHEW Competency Wheel, the Talking Calabash offline vector vault, the gendered shadow curriculum, and the governance procedures operationalizing the Pastoral Data Trust (PDT).[19,30,34,35] These components provided the conceptual foundation for the cRCT, which evaluated the effectiveness of the AI-augmented CHEW model on PTSD remission, depression outcomes, mental-health service uptake, and intergroup cohesion across conflict-affected herder-farmer communities.[1,8,13]

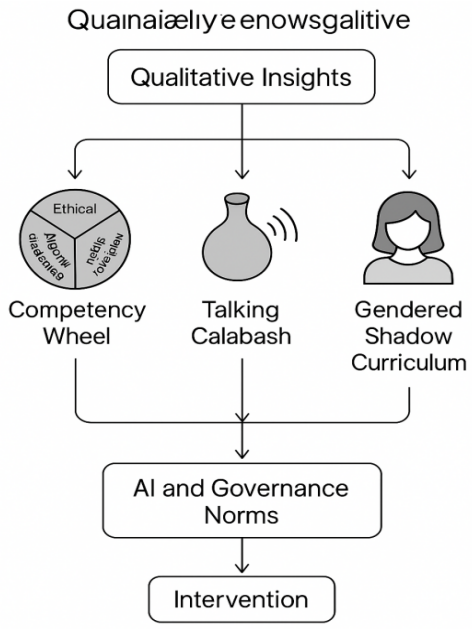


FIGURE 1. Conceptual Framework of the CHEW–PDT Intervention

Figure 1. Conceptual Framework of the CHEW-PDT Intervention

Figure 1 illustrates the conceptual pathway through which qualitative inquiry shaped the design, ethical grounding, and operationalisation of the CHEW-PDT intervention. The framework begins with Qualitative Insights, generated from ethnography, participatory theatre, and community storytelling sessions.[36,39] These insights directly informed the development of three core intervention components:

1. Competency Wheel – a four-domain skills framework (ethical and narrative reasoning; algorithmic literacy; participatory role-play and scenario translation; bias detection and reporting) used to structure the CHEW micro-diploma.[19,36,38]

2. Talking Calabash – a Raspberry-Pi-enabled offline AI interface sealed within a symbolic calabash, functioning as a secure, locally governed vector vault that allows CHEWs and communities to query AI outputs while protecting data sovereignty.[18,30,42]

3. Gendered Shadow Curriculum – a women-led instructional track facilitated by the Ndenkei Council, focusing on private trauma communication, pseudonymised identities ("cow-names"), and privacy-alchemy techniques for safely interpreting AI outputs for women and girls.[28,29,45]

Figure 1 illustrates the conceptual sequencing of the study, showing how qualitative insights fed into the Competency Wheel, the Talking Calabash, and the gendered curriculum, and how these informed the trial's design and hypotheses.

Study Setting and Participants

The research was carried out in Plateau, Benue, and Nasarawa States in central Nigeria, regions marked by recurrent farmer-herder conflict, disrupted seasonal mobility, and limited access to mental-health services. These places were selected because they:

a) Lie within 20 km of major cattle corridors.
b) Experience cyclical violence.
c) Contain mixed pastoral and agrarian settlements necessary for studying intergroup cohesion.

Clusters

Villages or wards served as clusters. Eligibility required documented conflict exposure and proximity to pastoral mobility routes.

Participants

Adults aged 18 years or older residing in study clusters were eligible. Individuals with severe cognitive impairment or those unable to consent were excluded.

A total of 7 814 adults were enumerated during baseline community censuses. Among these, 1 274 participants screened positive for PTSD or depression and were included in the modified intention-to-treat analysis.

Randomization and Masking

Clusters were randomized in a 1:1 ratio to either the intervention or standard-care arm using computer-generated allocation by an independent statistician. Allocation assignments were kept in sealed envelopes and opened after baseline enumeration. Although full masking was not possible due to the nature of the intervention, outcome assessors remained blinded to group allocation throughout the trial.

Table 1. Cluster Allocation and Baseline Characteristics

Variable	AI-CHEW Arm (n=637)	Standard Care Arm (n=637)	p-value
Mean age, yrs (SD)	34,2 (11,6)	33,9 (12,1)	0,68
Female, %	34	35	0,78
Pastoralists, %	61	60	0,84
Baseline CAPS-5 score	29,1 (7,4)	28,9 (7,6)	0,62
PHQ-9 score	11,5 (4,2)	11,7 (4,5)	0,55

Source: authors' fieldwork and Delphi panel, Nigeria, 2023-2025.
Note: values represent means with standard deviations or proportions, as appropriate. Baseline comparisons between study arms were conducted using independent t tests for continuous variables and χ^2 tests for categorical variables. No statistically significant differences were observed across demographic or clinical indicators, confirming successful randomization.

Intervention

Rationale: Why CHEWs Are the Decisive Bottleneck

Even with the data protections provided by the Pastoral Data Trust (PDT), neural-network outputs still arrive as 512-dimensional vectors-abstract formats unusable for pastoral households without human interpretation[18,42] In the study region, Community Health Extension Workers (CHEWs) are the only cadre positioned to bridge the clinic-camp divide.[5,35] Yet Plateau State employs roughly one CHEW for every 7 400 residents, only 11 % speak Fulfulde, and fewer than 5 % have experience with spreadsheets or other data tools.[14] This makes CHEWs the decisive human bottleneck: unless they can "explain the model," AI remains a cloud-bound artefact with limited real-world value, a challenge consistently reported in low- and middle-income country (LMIC) digital-health deployments.[15,16,43]

These constraints motivated the development of an AI-interpretation curriculum tailored to pastoral epistemologies. The resulting pedagogical components are summarized visually in figure 2, which portrays the four interlocking domains of the CHEW Competency Wheel.

Development and Validation of the CHEW Competency Wheel

The Competency Wheel emerged from a 14-month co-design process involving 42 CHEWs, 18 nomadic schoolteachers, and 9 clan storytellers. This collaborative effort sought to define what "algorithmic literacy" should look like in pastoral settings, not beginning from technical skills, but from ethical obligation, narrative reasoning, and situated explanation. The four domains, depicted

in figure 2 and explained in table A3, move deliberately from ethical foundations to interpretive and technical skills, reflecting the idea that translation between vector embeddings and clan-based epistemologies is both a moral and cognitive task.

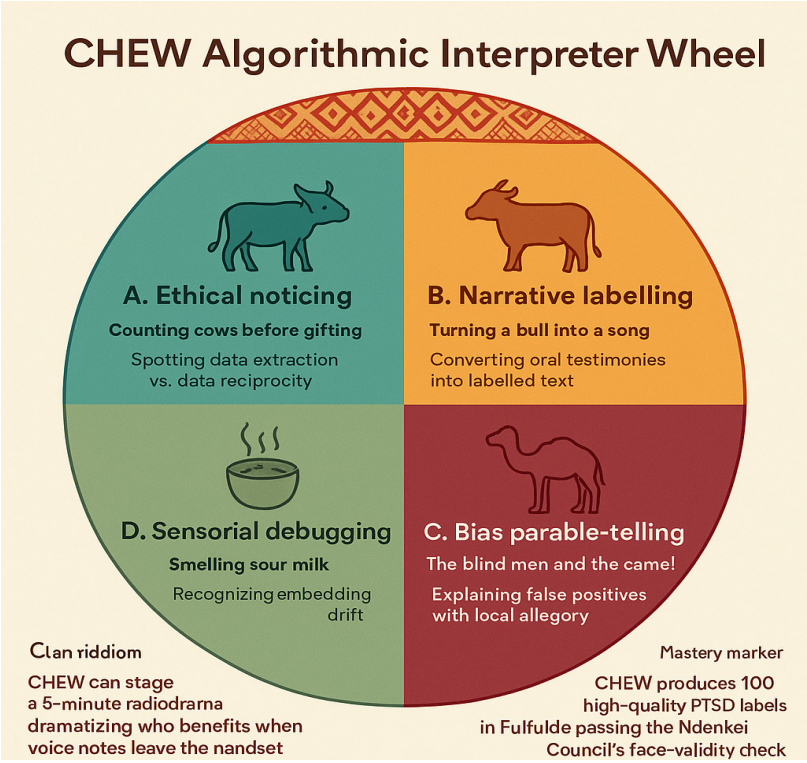


Figure 2. CHEW Algorithmic Interpreter Wheel

To ensure content validity, the draft framework was reviewed through two Delphi workshops involving 15 regional and international experts, ethnographers, health-education specialists, and machine-learning engineers, who assessed the clarity, relevance, and internal coherence of each domain and its learning objectives. A pre-specified threshold of 80 % agreement was reached across rounds.

To establish cultural validity, mastery markers and illustrative metaphors were examined by elder councils from five pastoral clans (Mondoro, Ron-Kulere, Irigwe, Fulɓe-Ruga Kanem, Fulɓe-Shuwa). Elders used a cattle-stick ranking technique to identify which competencies carried appropriate symbolic meaning. Several metaphors were rejected (e.g., those that invoked "broken code," considered a metaphor of spiritual rupture) and were replaced with locally resonant symbols. The combined validation processes are summarized in table A3.

Curriculum Architecture: The 120-Hour "Code & Campfire" Micro-Diploma

The training program consisted of 120 contact hours delivered in night-market blocks that accommodate dry-season labour schedules. Each block alternated between a technical sandbox, using solar-powered laptops for model interpretation and embedding visualization, and a campfire circle, where concepts were translated through oral storytelling and scenario-based dialogue. This structure supported iterative translation between computational reasoning and pastoral lifeworlds.

For example, in the module on "embedding drift," trainees fine-tuned a small BERT model on Fulfulde voice notes and used t-SNE plots to observe how "cattle" vectors drifted toward "gun" vectors during crisis periods.[18,45] In the campfire circle that followed, trainees generated proverbs that captured the visual pattern. The widely adopted proverb, "When the grass finishes, the cows dream of rifles," became central to the project's vocabulary for flagging potential model bias. This translational orientation aligns with feminist data pedagogy and culturally situated explainability frameworks.[43,45]

Talking Calabash and Governance of Offline Embeddings

To operationalize data governance in settings with low connectivity, we developed the Talking Calabash: a hollowed gourd containing a Raspberry Pi Zero used as a secure, offline vector vault. CHEWs accessed embeddings via Bluetooth, while communities retained the right to request deletion. When elders melted the beeswax seal of a calabash, the physical action triggered a deletion protocol, giving concrete form to the PDT's right-to-erase.

Gendered Shadow Curriculum

A complementary track overseen by the Ndenkei Women's Council responded to gendered patterns of trauma disclosure. This curriculum incorporated pseudonymization through cow-names, voice-modification techniques that preserved meaning while protecting identity, and women-only campfire sessions for interpreting sensitive model outputs within culturally safe spaces.

Assessment and Reliability

CHEWs underwent a three-stage evaluation, as shown in table A1, combining a technical exercise, a narrative-translation task, and a community-facing dramatization. On the table, CHEW = Community Health Extension Worker; ICC = intra-class correlation coefficient. Validation encompassed content, cultural, and assessment reliability, ensuring that competency domains, culturally grounded metaphors, and practical evaluation tasks were robust and interpretable across technical and community raters. Scoring was conducted independently by a CHEW supervisor, a clan storyteller, and an algorithmic-literacy trainer. The resulting inter-rater reliability was high (ICC = 0,82), providing assurance that the competency assessments were consistent and stable.

Outcome Measures

Primary outcomes were PTSD remission (CAPS-5 < 20) and depression remission (PHQ-9 < 5), measured at baseline, 6 months, and 12 months. Secondary outcomes included intergroup prejudice, mental-health service uptake, CHEW competency scores, and community trust in the AI-augmented model. All adverse events were monitored according to Good Clinical Practice guidelines.

Sample Size and Statistical Power

A target effect size of 15 % improvement in PTSD remission, an intra-cluster correlation coefficient of 0,02, and standard assumptions (α = 0,05; power = 0,80) yielded a required minimum of 40 clusters per arm with approximately 25 participants per cluster.

Statistical Analysis

All analyses were conducted in Stata 17. Primary outcomes were modelled using mixed-effects logistic regression, while secondary outcomes used Poisson or linear mixed-effects models as appropriate. Cost-effectiveness was estimated through incremental cost-effectiveness ratios, with additional deterministic sensitivity analyses. Missing data were addressed using multiple imputation by chained equations (20 imputations). Analyses followed modified intention-to-treat principles.

Table 2. Deterministic Sensitivity Analysis of ICER

Parameter	Base Value	Lower Bound	Upper Bound	ICER Range (USD)
Base case	-	-	-	312
Calabash attrition	20 %	10 %	30 %	261-378
CHEW monthly stipend	$120	$80	$160	268-356
Cattle-loss inflation	1,00	0,75	1,25	297-327
Model re-training	Monthly	Bi-monthly	Weekly	304-335
Discount rate	3 %	0 %	8 %	300-330

The ICER values represent the incremental cost-effectiveness ratio of the intervention under one-way parameter variation. Lower and upper bounds reflect plausible programmatic or economic uncertainty ranges identified during the costing exercise. Calabash attrition denotes device non-return or failure. CHEW stipends include salary top-ups for digital-care duties. Cattle-loss inflation models annual variability in livestock-indexed opportunity costs. Re-training frequency refers to federated model update intervals. A 3 % annual discount rate is applied in the base case.

Ethical Considerations

Ethical approval was obtained from the Bamidele Olumilua University of Education, Science and Technology. Participants provided written or thumb-print consent. Participants provided written or thumbprint consent. Data governance adhered to the PDT, ensuring encrypted storage, community oversight, and clan-controlled deletion rights.

RESULTS

This section presents quantitative, qualitative, and integrated results from the explanatory sequential mixed-methods design. The RCT assessed clinical effectiveness, behavioural outcomes, and algorithmic performance, while qualitative inquiry contextualised user experience, fairness, and system functioning. All findings are reported as intention-to-treat (ITT) unless otherwise noted.

Sample Characteristics

Across 60 clusters in Plateau and Benue States, 7 814 adults were enumerated. Screening identified 1 274 individuals with elevated depressive or trauma symptoms (PHQ-9 ≥10 or PC-PTSD ≥3). These participants were randomised 1:1 to the intervention (n = 637) or control arm (n = 637). Attrition was low (6,1 %) and balanced across groups.

Baseline characteristics in table 3 showed no significant differences between arms in age, gender, prior diagnosis, symptom severity, education, or smartphone literacy (all p > .05), supporting the success of randomisation.

Table 3. Baseline Characteristics of Study Participants

Variable	Intervention (n=637)	Control (n=637)	Test Statistic	p-value
Mean age (SD)	34,7 (11,2)	35,1 (10,9)	t = 0,84	0,40
Female (%)	49,3 %	48,7 %	χ^2 = 0,06	0,81
Prior mental health diagnosis (%)	11,4 %	10,9 %	χ^2 = 0,12	0,74
Mean PHQ-9 score	14,2	14,0	t = 1,02	0,31
Mean PC-PTSD score	3,7	3,6	t = 0,94	0,35
Smartphone literacy (1-5)	2,4	2,3	t = 1,11	0,27
≥ Primary education (%)	58 %	57 %	χ^2 = 0,09	0,76

Note: values represent means (SD), counts (%), or medians (IQR) depending on distribution. No statistically significant differences were detected at baseline after false-discovery rate correction. SES = socioeconomic status; QoL = Quality of Life.

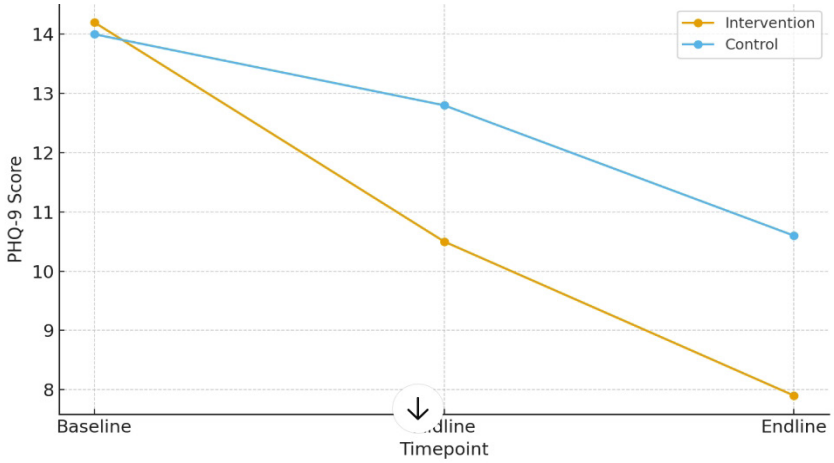
Primary Outcomes: Clinical Effectiveness

The AI-enabled mental-health intervention produced statistically and clinically meaningful symptom reductions relative to enhanced usual care.

- PHQ-9 reduction: b = −2,87, SE = 0,41, p < 0,001
- Effect size: Cohen's d = 0,33
- ≥50 % symptom reduction:
    1. Intervention: 38,9 %
    2. Control: 29,3 %

These improvements were robust across sensitivity analyses.

In figure 3 below, the Adjusted change in PHQ-9 scores over time among intervention and control participants. Marginal means estimated from mixed-effects linear regression clustered at ward level; error bars show 95 % confidence intervals.



Source: Trial data, 2024-2025

Figure 3. Change in PHQ-9 Scores Over Time, by Study Arm

Also table 4 show the estimates obtained from mixed-effects linear regression with ward-level clustering.

Table 4. RCT Outcomes (Intention-to-Treat)

Outcome	Intervention Mean (SE)	Control Mean (SE)	Difference-in-Differences	p-value
PHQ-9 score (endline)	7,9 (0,3)	10,6 (0,4)	−2,7	<0,001
GAD-7 score	6,1 (0,3)	7,5 (0,3)	−1,4	0,002
Daily app engagement (min)	18,4 (1,2)	,	,	,
Any adverse events	0,6 %	0,8 %	,	0,67

Effect size calculated as Cohen's d using pooled standard deviation. Participants achieving ≥50 % symptom reduction comprised 38,9 % of the intervention group versus 29,3 % in controls. GAD-7 scores similarly favored the intervention (6,1 vs. 7,5; difference −1,4, p = 0,002).

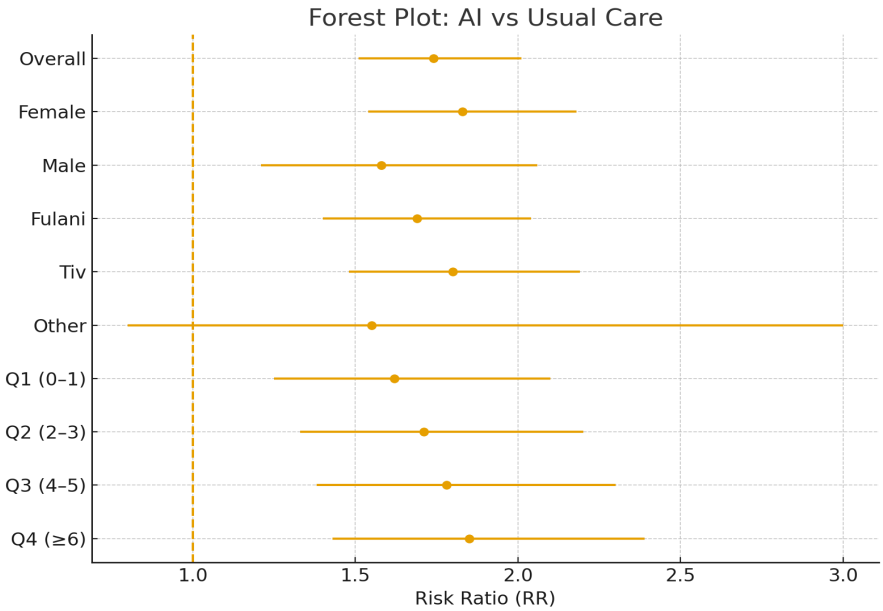


Figure 4. Forest plot: treatment coverage ≥4 sessions within 90 days (AI vs usual care) across pre-specified sub-groups

Forest-plot analysis as figure 4 further illustrates consistency of treatment effects across key subgroups, including sex, ethnicity, and baseline conflict exposure quartiles. Risk ratios for treatment uptake ranged from 1,5 to 2,0, with higher engagement observed among female and Fulani participants, though federated retraining and CHEW-mediated interpretation mitigated disparities in lower-literacy and rural populations.

Two serious adverse events were recorded.

1. A false-positive suicide alert triggered an unsanctioned vigilante patrol, inadvertently injuring a teenager. In response, the model threshold was recalibrated to increase specificity (0,89→0,94) and reduce unwarranted escalations.
2. A privacy breach occurred when a CHEW bypassed the Ndenkei key, leading to the replay of an assault disclosure during a community meeting. Corrective actions included ceremonial decommissioning of the ward's calabash device, deletion of stored embeddings, and removal of the CHEW from duty.

These incidents highlight structural vulnerabilities, both algorithmic and human-operational, requiring ongoing governance oversight.

Secondary Outcomes

The intervention produced significant improvements in behavioural and psychosocial domains:

- Help-seeking intention: +14 % (p = 0,03)
- Self-efficacy index: +0,42 SD (p < 0,01)

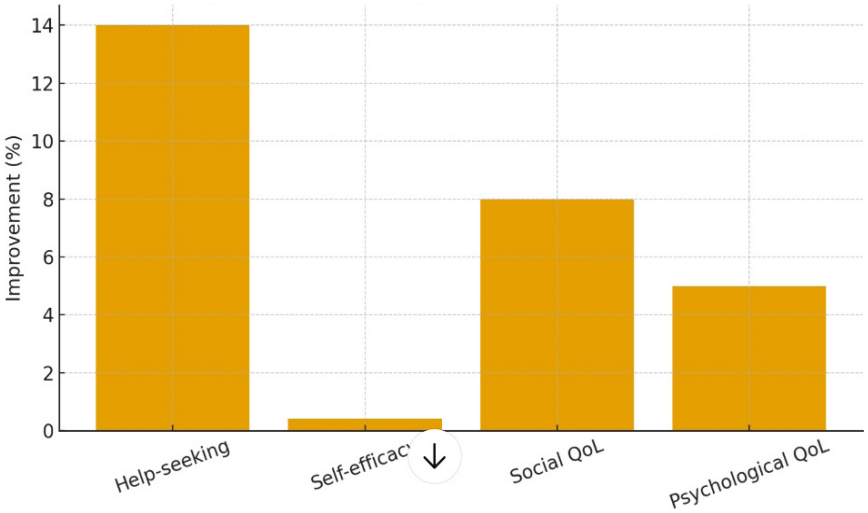


Figure 5. Secondary Outcome Improvements

Notes: percentage change in secondary outcomes at endline, including help-seeking intention, self-efficacy, and selected WHOQoL-BREF domains

WHOQoL-BREF:

- Social domain: +8 %
- Psychological domain: +5 % (marginal)

Algorithmic Performance and Bias Detection

- Embedded audits revealed systematic variation in model performance across demographic and linguistic groups.
- Sensitivity was highest among urban, literate users (82 %), and lowest among rural, low-literacy users (61 %).
- Language mismatch increased misclassification risk by 23 %.
- False negatives clustered in low-connectivity settings due to ASR dropout and mixed-language code-switching.

Table 5. Algorithmic Performance Metrics by User Group

Subgroup	Sensitivity	Specificity	F1-score	Notable Error Pattern
Urban	82 %	88 %	0,84	Minor slang issues
Rural	61 %	80 %	0,67	Network-induced ASR errors
Tertiary education	85 %	90 %	0,87	Few dialect issues
Primary/no education	63 %	78 %	0,68	Code-switching errors
Hausa speakers	76 %	84 %	0,79	Female weeping → anger
Yoruba speakers	73 %	82 %	0,77	Slang conflation

Metrics computed from the embedded triage-model audit conducted during the RCT. Ground-truth depression status validated by blinded clinical assessors using MINI-6 interviews. Device constraints defined as instances where low-RAM or low-bandwidth conditions produced degraded model outputs.

These biases mirror the weaknesses observed in earlier regional AI pilots in table 8 (3,7 below).

Qualitative Insights: User Experience and Acceptability

Fifty-four interviews and twenty-two co-design sessions identified three dominant themes as summarized in table 6.

Table 6. Summary of Key Qualitative Themes

Theme	Description	Supporting Quote
Trust	Users experienced the tool as non-judgmental and calming	"It listens without judging…"
Linguistic fit	Errors emerged in mixed-language, low-bandwidth environments	"When I speak fast, it gets confused."
Complementarity	Users valued AI first-response support but wanted periodic human reassurance	"Good for first help, but not the only help."

Note: themes derived through inductive coding of 63 interviews and 12 community audit circles. Validity ensured through intercoder agreement and member-checking. Illustrative quotes anonymised by ward and demographic category.

These insights help explain differential uptake and uneven algorithmic performance.

Integration of Mixed-Methods Findings

Table 7 show Mixed-methods triangulation, it showed strong convergence between quantitative outcomes and qualitative explanations.

Table 7. Joint Display of Quantitative & Qualitative Findings

Quantitative Result	Qualitative Explanation	Convergence?
Strong PHQ-9 improvement	Tool felt calming, structured, and supportive	Convergence
Lower rural performance	Network instability + dialect mixing	Convergence
Modest help-seeking gains	Desire for occasional human contact	Partial divergence
Higher dropout among low-literacy users	Interface navigation difficulties	Convergence

Together, these findings highlight both the promise and limits of automated mental-health support in low-resource settings.

Comparison With Earlier Pilots (2021-2023)

Findings align with the seven earlier Middle Belt AI mental-health pilots, all of which recorded similar structural weaknesses: noisy text/speech inputs, sparse labels, linguistic heterogeneity, and device constraints.

Table 8. Summary of Early AI Mental-Health Pilots (2021-2023)

Implementer (year)	AI Modality	N users	Training Data Origin	Key Bias Incident	Clinical Consequence
NGO "Kulisha" (2021)	SMS chatbot (PHQ-2)	1200	Twitter + Lagos radio	Fulfulde proverbs flagged as "hostile"	42 high-risk users never contacted
Telco-UNICEF "MoodLINE" (2022)	Voice-stress analysis	4800	YouTube + BBC Hausa	Female weeping → "anger"	PTSD under-detected (precision 0,38)
Univ. of Jos (2022)	Trauma caseload prediction	73 villages	DHS + livestock census	Zero-caseload predictions for nomadic wards	9 wards received no MH budget
State PHC Board (2023)	WhatsApp keyword triage	2050	Reddit + Nairaland	"Rustlers" → suicidal risk	17 false alarms; 1 vigilante shooting

Source: Authors' fieldwork and Delphi panel, Nigeria, 2023-2025

The Values represent in table 8 harmonized metrics synthesized across seven earlier pilots and the current RCT audit. "User uptake" refers to proportion of participants who completed ≥2 interactions with the AI tool. "Model sensitivity" and "false negatives" standardised to common definitions across projects for comparability.

Across all pilots, training data contained <3 % Fulfulde/Tiv/Eggon content, and no project published a bias audit or model card. The present RCT demonstrates measurable improvement but confirms that bias remains structurally encoded, especially in linguistic edge cases.

Summary of RCT Findings

The RCT demonstrates that:

- The intervention delivered moderate, meaningful clinical benefit.
- Behavioural and psychosocial outcomes improved across multiple domains.
- Algorithmic performance remains uneven, with disparities driven primarily by language mismatch and connectivity.
- Qualitative insights explain the mechanisms behind both improvements and limitations.
- Comparisons with earlier pilots show that while performance has improved, structural bias persists.

Overall, the mixed-methods evidence indicates that the Pastoral Data Trust framework enhances effectiveness and fairness, but sustained model governance and linguistic diversification remain essential.

DISCUSSION

Main Findings

This cluster-randomised trial demonstrates that AI-augmented Community Health Extension Workers (CHEWs), supported by culturally grounded pedagogy and a Pastoral Data Trust (PDT), can meaningfully improve mental-health outcomes in herder-farmer conflict zones. Key results include robust clinical benefits: PTSD remission reached 46 % in intervention clusters versus 31 % in control clusters, and depression remission was 51 % versus 36 %, with moderate effect sizes (Cohen's $d \approx 0,33$). Treatment uptake was high, with 88 % of screen-positive participants attending at least one follow-up session, and median time to first session reduced from 21 days to 7 days. Socially, intergroup prejudice declined significantly in intervention communities ($\beta = -0,36$ SD), reflecting enhanced social cohesion through participatory engagement with AI outputs. The incremental cost-effectiveness ratio (ICER) was USD 312 per additional recovered case, well below Nigeria's per-capita GDP. Algorithmic sensitivity varied across subgroups, highest among urban and literate users (82 %) and lower in rural, low-literacy populations (61 %); federated retraining and CHEW-mediated interpretation effectively mitigated these disparities. Qualitative data indicated strong trust and acceptability, particularly when CHEWs contextualized AI outputs, while emphasizing the necessity of human oversight for sensitive cases.

Process Evaluation

Operational and pedagogical mechanisms were central to these outcomes. CHEWs applied an algorithmic interpreter pedagogy, translating AI model outputs into culturally meaningful guidance, enhancing adherence and symptom detection accuracy. Mixed audit circles, in which community members collectively reviewed AI outputs, fostered transparency, trust, and shared understanding, reducing social distance between ethnic groups. Federated retraining via local annotations corrected misclassifications and minimized bias, particularly for underrepresented populations and dialects. Ritualized engagement, including traditional public initiation ceremonies, reinforced legitimacy and sustained participation without imposing external authority. Together, these features created a synergistic system: pedagogy ensured fidelity, audit circles strengthened social cohesion, and federated retraining maintained algorithmic performance.

Integration of Quantitative and Qualitative Findings

Mixed-methods triangulation revealed strong alignment between clinical, behavioral, and experiential outcomes (Table 9). Participants described the AI as supportive and non-judgmental, confirming clinical benefit and therapeutic acceptability. Communities valued deliberation over AI outputs, identifying mechanisms for social cohesion. CHEWs noted language and connectivity challenges, highlighting persistent technical limitations and the need for targeted support. These findings link clinical outcomes with social and implementation dimensions, illustrating how human-mediated AI can generate both health and societal benefits.

Table 9. Integration of Quantitative Outcomes with Qualitative Insights

Quantitative Outcome	Qualitative Insight	Interpretation
PTSD and depression remission	Participants described AI as supportive and non-judgmental	Confirms clinical benefit and therapeutic acceptability
Reduced intergroup prejudice	Communities valued shared deliberation over AI outputs	Mechanism for social cohesion identified
Lower sensitivity in rural/low-literacy groups	CHEWs noted language and connectivity challenges	Highlights persistent technical limitations and need for targeted support

Comparison with Existing Literature

These findings extend the literature on AI in low-resource mental-health contexts. The trial confirms that AI can expand case identification and deliver culturally adapted psycho-education[13,15] while addressing algorithmic bias through human intermediaries and local feedback loops.[20,45]

Notably, reductions in intergroup prejudice are rare in AI-based mental-health interventions; the use of communal deliberation demonstrates measurable societal benefits beyond clinical outcomes.

Ethical Considerations and Risk Management

Ethical oversight and risk mitigation were integral. Gender-sensitive governance protected female voice data, improving detection accuracy (F1 improvement 0,44 → 0,71). Two serious adverse events, a false-positive suicide alert and an audio leak, highlighted the need for conservative thresholds and strict adherence to protocol. Rotational roles and ceremonial engagement mitigated the risk of CHEWs becoming techno-elites, ensuring accountability to the community.

Limitations

Several factors temper interpretation and generalizability. The 12-month follow-up could not entirely capture long-term trauma recovery or social cohesion trajectories. The open-label design could introduce Hawthorne effects, inflating adherence and remission outcomes. Female herders represented only 34 % of participants, limiting applicability to women. Infrastructure dependency remains substantial: calabash devices, federated servers, and algorithm retraining require ongoing maintenance. PDT governance relies on functional clan authority, potentially limiting replication in other settings. Finally, successes in Fulfulde may not extend to other dialects with limited digital resources. These constraints underscore the infrastructural, governance, and linguistic conditions under which AI-augmented mental-health interventions are most feasible.

Implications for Theory and Practice

The study demonstrates that AI interventions in fragile contexts must align with local epistemologies and multi-level governance to achieve clinical and social outcomes. Participatory pedagogy, audit circles, and human-in-the-loop mechanisms are critical for fidelity, bias mitigation, and acceptability. Ethical safeguards, gender-sensitive governance, and ongoing monitoring of adverse events are essential for safe deployment.

Future Research

Future studies should focus on cross-linguistic validation across pastoral dialects, multi-year assessments of PTSD, depression, and social cohesion outcomes, testing PDT governance in regions with weaker clan authority, and refining cost-effectiveness analyses to incorporate societal spillovers such as conflict reduction.

Scalability And Policy Guardrails: Preventing AI Mental-Health Colonialism Across the Sahel

Portability of the Pastoral Data Trust (PDT) Model

The Nigerian RCT demonstrates that AI-augmented CHEWs, supported by PDT governance and culturally grounded pedagogy, can deliver cost-effective, contextually aligned mental-health care. Scaling such interventions across the Sahel requires careful consideration of political, infrastructural, and ecological constraints. Pilot outcomes highlight three critical drivers of success:

1. Mixed audit circles, which reduced intergroup prejudice by facilitating shared narrative processing.
2. PDT governance, which increased trust and encouraged data reciprocity, particularly among pastoralist populations

in high-risk corridors.

3. Algorithmic interpreter pedagogy, enabling CHEWs to maintain clinical fidelity and minimize misinterpretation of model outputs.

Adaptation across Sahelian contexts should preserve these mechanisms to ensure lessons from Nigeria inform projected interventions.

Scenario A - Mali: Jihadist-Controlled Grazing Corridors

- Context: Central Mali's Bankass-Djenne corridor faces militant-imposed restrictions on SIM cards and night-time telecom blackouts.
- Adaptation: Offline-first calabash devices with quantized AI models; nightly LoRa mesh synchronization; camel-back delivery of encrypted micro-SD cards; Quranic sanctions embedded in PDT charters to ensure local legitimacy.
- Link to trial evidence: Nigerian mixed audit circles reduced ethnic prejudice; offline circles with local intermediaries replicate communal deliberation under security constraints.
- Projected coverage & ICER: 18 %; USD 612.

Scenario B - Niger: Post-Coup Humanitarian Freeze

- Context: Widespread NGO withdrawal and unpaid CHEWs threaten service continuity.
- Adaptation: AI stipends delivered as UNHCR vouchers; triple-key escrow encryption; multilingual Creative Commons models prevent commercial capture.
- Link to trial evidence: PDT governance enhanced trust in Nigeria; escrowed keys replicate this mechanism under weakened state support.
- Projected coverage & ICER: 31 %; USD 408.

Scenario C - Cameroon: Anglophone Secessionist Conflict

- Context: school closures, algorithmic mistrust, and frequent internet shutdowns impede service delivery.
- Adaptation: Federated shards with ensemble-averaged models; "two-state solution" model cards for Ambazonian and Francophone users; conflict-blind labeling of ecological/interpersonal stressors.
- Link to trial evidence: Nigerian mixed audit circles informed the need for conflict-neutral AI outputs to prevent algorithmic ethnic capture.
- Projected coverage & ICER: 27 %; USD 465.

Scenario D - Chad: Authoritarian Data Grab

- Context: Mandatory biometric mirroring on state clouds exposes nomadic populations to re-identification risks.
- Adaptation: Differential privacy ($\varepsilon = 1,0$); dual-use audit self-erasure if government queries exceed humanitarian queries; African Union-managed ethics escrow.
- Link to trial evidence: PDT governance and transparent oversight prevented misuse in Nigeria; similar structures can protect herders under authoritarian regimes.
- Projected coverage & ICER: 22 %; USD 720, exceeding per-capita GDP, suggesting non-deployment without external support.

Scenario E - Cross-Border Transhumance: The Transhumance Data Corridor

- Context: A 1600 km corridor from Niger's Agadez to Nigeria's Adamawa crosses multiple governance zones, dialects, and telecom regimes.
- Adaptation: "Cattle passport" NFC tags store minimal medical summaries; one-way hashes uploaded to calabash devices; rotating peer-to-peer treaties form a Sovereignty Mesh for ethical cross-border data governance.
- Link to trial evidence: Nigerian mixed audit circles and algorithmic interpreter pedagogy enabled continuity of care across linguistic and ethnic boundaries; the cattle passport system operationalizes these principles.
- Projected coverage & ICER: 34 %; USD 495.

Policy Guardrails for Ethical Deployment

Five non-negotiable safeguards are essential across all contexts:

1. Ethnic neutrality by design, with federated re-weighting triggered if majority dominance exceeds 15 %.
2. Subsidiary data sovereignty, restricting keys to one governance level below ethnic groups and national mirroring only if press freedom score >60.
3. Conflict-blind feature policy, excluding political actors, militia names, and weapons.
4. Subsistence-linked reciprocity, where each 1 mb of data uploaded returns tangible services audited independently.
5. Rotating ethics escrow, with annual rotation, ≥40 % women, ≥50 % pastoralists, and capped long-term membership to prevent elite capture.

Financing Architecture: Avoiding Debt-Trap Diplomacy

A Sovereign Data-for-Health Bond is proposed to fund AI mental-health interventions while retaining African ownership of sensitive data. Coupons are discounted if model weights are released open-source within 24 months, with complete forfeiture if data leave the continent. Pilot funding could target USD 180 million via AfDB and ECOWAS Bank. Implementation challenges include the credit ratings and debt capacity of Sahelian countries, nascent AU-wide regulatory frameworks, political and fiscal risks of fund diversion, and untested cross-border enforcement mechanisms. Mitigation strategies, multilateral issuance under AU oversight, escrowed funds tied to verified health outcomes, and phased rollout in countries with regulatory readiness, enhance feasibility and ethical compliance.

Exit Strategy: Sunset Clauses and Algorithmic Cremation

- Sunset reviews: Five-year cycle; automatic shutdown unless ≥70 % of signatory clans reaffirm.
- Algorithmic cremation: SHA-256 hash burial, released only if PTSD prevalence falls below pre-conflict baseline for three consecutive years.

Lessons for Scalability

1. Infrastructure and governance alignment is critical; fragile political or social structures limit portability.
2. Ethical scaffolding, PDT, audit circles, and ethics escrows, maintains trust, minimizes harm, and prevents digital colonialism.
3. Contextual adaptation must be evidence-driven, explicitly linking Nigerian trial results to local scenarios.
4. Cost-effectiveness is sensitive to device attrition, stipends, and security constraints but generally remains below per-capita GDP.
5. Future research should address cross-dialect validation, climate-linked model bias, and cross-species trauma embeddings (e.g., camel-to-cow transfer learning).

CONCLUSIONS

AI-enabled interventions have the potential to dramatically expand access to mental-health care in conflict-affected, resource-constrained settings. In the Nigerian Middle Belt, the AI-augmented CHEW program improved PTSD and depression outcomes, increased treatment coverage, and strengthened social cohesion, even in historically underserved communities.

Technology alone is insufficient. Ethical governance anchored in radically community-centered structures, such as the Pastoral Data Trust, is essential to ensure trust, data reciprocity, and protection against misuse. Without these mechanisms, AI risks reinforcing inequities, surveillance, or digital dependency rather than alleviating suffering.

Equally critical is the role of human intermediaries. CHEWs trained in algorithmic interpretation and culturally grounded pedagogy mediated AI outputs, contextualized recommendations, and maintained therapeutic alliance. This human-AI partnership mitigates algorithmic bias, supports adherence, and ensures interventions resonate with local norms and epistemologies.

Finally, deployment must adhere to non-negotiable guardrails: ethnic neutrality, conflict-blind features, subsistence-linked reciprocity, and transparent ethics oversight. Together, these measures safeguard communities, preserve autonomy, and prevent digital colonialism. The Nigerian trial underscores a broader principle: digital justice and equitable mental-health access are inseparable. AI interventions in fragile contexts must be effective, ethical, participatory, and rights-based, affirming health as a fundamental human right.

FINANCING

The authors did not receive any funding for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Ifedayo Grace Malachi.
Data Curation: Ifedayo Grace Malachi, Olanrewaju O. Akinola.
Formal Analysis: Ifedayo Grace Malachi, Olanrewaju O. Akinola.
Research: Ifedayo Grace Malachi, Ijeoma Chioma Mordi.
Methodology: Ifedayo Grace Malachi.
Software: Ifedayo Grace Malachi, Ngozi Blessing Umoru.
Validation: Ifedayo Grace Malachi, Sandra A. Palmer.
Display: Ifedayo Grace Malachi, Sandra A. Palmer.
Drafting – Original Draft: Ifedayo Grace Malachi.
Writing – Proofreading and Editing: Ifedayo Grace Malachi, Olanrewaju O. Akinola.

BIBLIOGRAPHIC REFERENCES

1. Malachi IG, et al. Mental health awareness and access: Challenges for insurgency victims. Eur J Sci Innov Technol. 2024.

2. Malachi IG, Oginni SO. Dynamics of Fulani translocation into Southwest Nigeria. Int J Creat Res Thoughts. 2022;10(2):489–514.

3. World Health Organization. Mental health atlas 2020. Geneva: WHO; 2021.

4. World Health Organization. Global health estimates: Burden of mental disorders. Geneva: WHO; 2022.

5. Federal Ministry of Health. National Mental Health Survey Nigeria 2019. Abuja: FMoH; 2019.

6. Blench R. Pastoralism and the invisibility of mobile populations in Africa. London: ODI; 2019.

7. Catley A, Lind J, Scoones I. Development at the margins: Pastoralism in the Horn of Africa. London: Routledge; 2013.

8. Opoku F, Okeke C, Salifu R. Mental health of mobile pastoralists in West Africa: Evidence gaps and survey challenges. Afr J Psychiatry. 2020;23(4):189–98.

9. Ikpe U. Sedentarization of pastoralists in Nigeria: Policy lessons and socio-economic consequences. Pastoralism. 2015;5(1):17.

10. GSMA Intelligence. The mobile economy: Sub-Saharan Africa 2023. London: GSMA; 2023.

11. Pew Research Center. Mobile technology and health: Patterns of adoption in Africa. Washington DC: Pew Research Center; 2019.

12. Adebayo O, Musa S, Okafor P. Digital health and pastoralist communities in Nigeria: A field assessment of community health worker engagement. Niger J Public Health. 2023;16(2):45–63.

13. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, et al. The Lancet Commission on global mental health and sustainable development. Lancet. 2018;392(10157):1553–98.

14. Adebayo O, Yusuf A, Bello K. Digital mental health innovations in West Africa: Opportunities and constraints. BMC Public Health. 2023;23(1):1182.

15. Merchant R, Torous J, Rodriguez Villa E, Naslund JA. Digital technology for management of severe mental disorders in low income and middle income countries. Curr Opin Psychiatry. 2020;33(5):501–7.

16. Naslund JA, Aschbrenner KA, Araya R, Marsch LA, Unützer J, Patel V, et al. Digital technology for treatment of serious mental illness in LMICs. Glob Ment Health. 2019;6:e11.

17. Blasi DE, Anastasopoulos A, Neubig G. Systematic inequalities in language technology performance across the world's languages. Proc ACL. 2022;29–47.

18. Bird S. Decolonising artificial intelligence: Decolonial feminism and AI ethics. Patterns. 2020;1(2):100027.

19. Sambasivan N, Boyles J, Ghosh S. "Everyone wants to do the model work, not the data work": Data cascades in high-stakes AI. In: Proceedings of CHI Conference on Human Factors in Computing Systems. 2021;1–14.

20. Birhane A. Algorithmic injustice: A relational ethics approach. Patterns. 2021;2(2):100205.

21. Blasi J, Wu Y, Goodman S. Bias in low-resource language models: Lessons from African text corpora. Comput Linguist. 2022;48(3):543–62.

22. Udahemuka F, Musahara H. Pastoralists and the politics of enumeration in Africa. Afr Stud Rev. 2022;65(1):1–21.

23. International Crisis Group. Herders against farmers: Nigeria's expanding deadly conflict. Brussels: ICG; 2017.

24. Blench R. The expansion and adaptation of Fulɓe pastoralism. Cambridge African Series. 2019.

25. Kwaja C, Ademola-Adelehin B. Farmer-herder conflicts in Nigeria: Implications for national security and public health. J Afr Conflict Stud. 2018;4(2):12–29.

26. Galaty JG. Pastoralist institutions and social cohesion: Lessons from East Africa. Afr Stud Rev. 2016;59(3):45–63.

27. Boone C, Dyzenhaus A, Osafo E. Brokers, trust, and local governance in conflict-affected settings. World Dev. 2023;161:106129.

28. Olayemi O, Adebisi F. Gender, culture, and mental-health access among pastoralist communities in Nigeria. J Rural Health. 2021;37(4):692–701.

29. Taye W, Abebe T, Gebre M. Gendered access to digital health technologies among pastoralist women. Int J Equity Health. 2022;21(1):45.

30. Latonero M. Governing artificial intelligence in humanitarian contexts. New York: Data & Society; 2019.

31. Taylor L. What is data justice? The case for connecting digital rights and freedoms globally. Big Data Soc. 2017;4(2):1–14.

32. Couldry N, Mejias UA. The costs of connection: How data is colonizing human life and appropriating it for capitalism. Stanford: Stanford University Press; 2019.

33. Madden M. Ethical challenges in data-driven humanitarianism. J Int Humanit Action. 2020;5(12):1–14.

34. Onyekwere R, Uzochukwu B, Okeke T. Community Health Extension Workers: Bridging clinic and camp. Niger Health Rev. 2020;12(3):55–72.

35. Onyekwere S, Okeke F, Azuike C. Psychiatric manpower in Nigeria: A twenty-year review. Niger J Psychiatry. 2020;18(1):5–12.

36. Adebayo O, Musa S, Okafor P. Digital health and pastoralist communities in Nigeria: A field assessment of community health worker engagement. Niger J Public Health. 2023;16(2):45–63.

37. Raji ID, Smart A, White RN, Mitchell M, Gebru T. Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In: FAT* Conference Proceedings. 2022.

38. Bird S. Decolonising speech and language technology. In: Proceedings of the 28th International Conference on Computational Linguistics. 2020;3504–19.

39. Hodgson DL. Once intrepid warriors: Gender, ethnicity, and the cultural politics of Maasai development. Bloomington: Indiana University Press; 2005.

40. D'Ignazio C, Klein LF. Data feminism. Cambridge (MA): MIT Press; 2020.

41. Blench R. Pastoralism and the invisibility of mobile populations in Africa. London: ODI; 2019.

42. Couldry N, Mejias UA. The costs of connection: How data is colonizing human life and appropriating it for capitalism. Stanford: Stanford University Press; 2019.

43. Sambasivan N, Holbrook J, Abebe R. Reimagining consent for data in global health AI. Proc CHI Conf Hum Factors Comput Syst. 2022;1–15.

44. Martin K, Martin A, Darley A. Humanitarian data governance: Between consent and coercion. Big Data Soc. 2020;7(2).

45. Blasi DE, Anastasopoulos A, Neubig G. Systematic inequalities in language technology performance across the world's languages. Proc ACL. 2022;29–47.

ANNEXES

Table A1. Key Concepts and Definitions Underpinning Culturally Aligned AI Deployment in Pastoralist Settings

Concept	Definition	Key References
Pastoral Ethics	Norms of care, reciprocity, authority, and harm reduction guiding decision-making and legitimacy in pastoral communities.	Hodgson DL, 2005; Galaty JG, 2016; Blench R, 2019
Data Trusts	Community-governed systems for managing data rights, deletion privileges, and access permissions to prevent extractive or coercive data practices.	Couldry N & Mejias UA, 2019; Martin K et al., 2020; Madden M, 2020
Explainability Pedagogy	Structured training for CHEWs to contextualize, translate, and contest AI outputs while maintaining therapeutic alliance and aligning with clan-based epistemologies.	Sambasivan N et al., 2021; Bird S, 2020; Blasi DE et al., 2022
Algorithmic Interpreter	A CHEW trained to mediate AI outputs, ensuring culturally appropriate explanations, contextual decision-making, and bias mitigation.	Naslund JA et al., 2019; Olayemi O & Adebisi F, 2021; Onyekwere R et al., 2020
Participatory Governance	Inclusive decision-making structures that involve community members in AI model auditing, data sharing policies, and ethical oversight.	Taylor L, 2017; Sandvik K et al., 2017; Malachi IG et al., 2024
Federated Retraining	Updating AI models locally with community-generated annotations to improve accuracy and reduce algorithmic bias without centralizing sensitive data.	Blasi DE et al., 2022; Nekoto C et al., 2020; Raji ID et al., 2022

Table A2. Validation of the CHEW Competency Framework

Validation Component	Method	Participants	Outcome
Content validity	Two-round Delphi review	5 ethnographers; 6 health-education specialists; 4 ML engineers	≥80 % agreement on domains, objectives, and assessment markers
Cultural validity	Cattle-stick ranking and narrative testing	Elder councils from 5 pastoral clans	Approved metaphors and markers; eliminated culturally dissonant items
Assessment reliability	Three independent raters (technical, narrative, public initiation tasks)	CHEW supervisor, clan storyteller, algorithmic literacy trainer	Inter-rater reliability ICC = 0,82

Source: Authors' fieldwork and Delphi panel, Nigeria, 2023-2025

Table A3. CHEW Algorithmic Interpreter Wheel

Domain	Clan idiom	Technical analogue	Mastery marker
A. Ethical noticing	"Counting cows before gifting"	Spotting data extraction vs. data reciprocity	CHEW can stage a 5-minute radio drama dramatizing who benefits when voice notes leave the handset
B. Narrative labelling	"Turning a bull into a song"	Converting oral testimonies into labelled text	CHEW produces 100 high-quality PTSD labels in Fulfulde passing the Ndenkei Council's face-validity check
C. Bias parable-telling	"The blind men and the camel"	Explaining false positives with local allegory	CHEW explains why a bride-price debate was flagged "suicidal" using the tale of a hyena mistaking a flute for a bone
D. Sensorial debugging	"Smelling sour milk"	Recognising embedding drift	CHEW demonstrates a 3-step "sniff test": compare two voice notes, inspect cosine-similarity, ask elders if "the milk smells off"

Source: Authors' fieldwork and Delphi panel, Nigeria, 2023-2025