World Health Organization
South-East Asia Region

intelehealth

Telemedicine In Action: Transforming healthcare in LMICs

Webinar Topics and Dates

Sno	Date	Topic
1	06 March 2025	What is Telemedicine and How Are Health Systems Using It Globally? A Primer for Health System Leaders
2	10 April, 2025	Brick-and-mortar to Brick-and-click - Designing & Implementing Quality, Effective, and Impactful Telemedicine Programs
3	08 May, 2025	Evaluating telemedicine interventions: Evidence so far, and Methodologies
4	5 June, 2025	Creating a Telemedicine-Ready Healthcare Workforce: Training for Healthcare Providers
5	10 July, 2025	Telemedicine Policy: How Telemedicine is Regulated in Asia
6	7 August, 2025	Choosing a Telemedicine Software: The case for standards-compliant, interoperable & open-source Digital Public Goods (DPGs)
7	11 September, 2025	Ensuring Quality of Care & Patient safety in Telemedicine
8	9 October, 2025	Telemedicine Adoption by Communities – How Might We Drive Uptake of Telemedicine (TM) by Citizens?
9	6 November, 2025	Artificial Intelligence and Machine Learning in Telemedicine
10	11 December, 2025	Financing Telemedicine and ROI – The Business Case for Telemedicine
11	8 January, 2026	Telemedicine use cases to advance the SDGs - Part 1 Applications for Non-Communicable Diseases (Diabetes, Hypertension, Cardiovascular disease, Cancer and Mental Health)
12	5 February, 2026	Telemedicine uses to advance the SDGs - Part 2 Applications for Communicable Diseases (Tuberculosis, HIV)
13	12 March, 2026	Telemedicine use cases to advance the SDGs - Part 3 Applications for Primary Healthcare

Artificial Intelligence and Machine Learning in Telemedicine

Objectives and Outcomes

Objectives:

- Introduce participants to the use of AI/ML in telemedicine.
- Highlight current real-world use cases and innovations.
- Explore ethical, regulatory, and technical considerations for AI integration.

Expected Outcomes: By the end of the webinar, participants will:

- Gain a foundational understanding of how AI and ML are being applied in telemedicine.
- Why do we need AI and ML in telemedicine in order to address key challenges in the sustainability and scalability of telemedicine programs.
- Learn from case studies of AI-driven telemedicine solutions in LMICs.
- Understand the risks, limitations, and ethical challenges of using AI in healthcare.

Dr. Saif Khairat

Dr. Saif Khairat is the Beerstecher-Blackwell Distinguished Professor at UNC-Chapel Hill, and the principal investigator of the NIH-funded Center for Virtual Care Value and Equity (ViVE). Dr. Khairat is an expert in digital health who has authored over 100 publications, secured $7.5M in funding, and is a digital health advisor to the WHO. He holds a PhD in Health Informatics and master's degrees in Computer Science and Public Health.

Ms. Divya Kamerkar

Divya Balaji Kamerkar is a visionary digital health entrepreneur committed to transforming women's sexual and reproductive healthcare through technology. As the Co-Founder and CEO of Pinky Promise, India's largest women's health platform that uses AI and great doctors to provide instant healthcare on a woman's fingertips. She blends her background as a biologist from Yale and an MBA from Wharton to create innovative, scalable, and accessible models of care. Her work harnesses AI and digital platforms to close healthcare gaps and empower women across India.

Dr. Neha Verma

Dr. Neha Verma is the Co-founder and CEO of Intelehealth, a telemedicine technology non-profit that delivers health services where there is no doctor. She is an entrepreneur and medical information engineer. She earned an MS in Applied Health Sciences and a PhD in Health Informatics from the Johns Hopkins University School of Medicine. Neha is also an active contributor for Women@Forbes, writing about women in tech, product development, organizational strategy, social impact and nonprofits.

Speaker I

Integration of ML & AI in Telehealth

Saif Khairat, PhD, MPH
Professor, University of North Carolina at Chapel Hill
Director, Center for Virtual Care Value and Excellence (ViVE)

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

VIVE
CENTER FOR VIRTUAL CARE VALUE AND EXCELLENCE

The Strategic Imperative : Bridging the Access Gap

Rural Access Gap
Roughly 50% live in rural areas, far from specialty care

Workforce Gap
Physician density ~7 per 10,000 vs ~16–17 globally

Disease Burden
Non-communicable diseases are a majority of mortality alongside infectious threats

Efficiency
Target AI where it extends reach and productivity toward UHC with safety and equity

Example (tele-triage hub): Village clinics send cases to a district telehealth hub; an AI triage bot prioritizes same-day video consults, reducing avoidable travel.

Sources: Rural access (regional context): World Bank – Rural population (% of total), South Asia viewPhysician density (global comparator; regional pages link to country detail): WHO Data – Density of physicians (per 10,000)NCD burden (SEARO): WHO South-East Asia — Noncommunicable diseases (regional facts)Telehealth at scale example for framing (national program): eSanjeevani (India) overview in BMJ GH commentary

AI in Telehealth

- **NCD prevention in LMICs**
- **Telemedicine solutions** for remote diagnosis and care.
- Chatbot-based tools to deliver personalized health information at scale.

How it works?

1. Prioritise condition/risk
2. Localise messages
3. Integrate with national systems
4. Monitor real-time KPIs

Source: https://www.itu.int/en/ITU-D/ICT-Applications/Pages/mhealth-for-ncd-behealthy-bemobile.aspx
https://www.who.int/initiatives/behealthy

ITU World Health Organization

Symptom checking and triage

ML and NLP Powered Chatbots

- Uses NLP to interpret patient-reported symptoms
- Uses ML to Triage
- Integrated into teleconsultation services in many countries (Germany, UK, USA, Singapore)

Impact: Improved Access, Improved Efficiency Reduces Physician Workload

- High Accuracy and safety[1, 2]
- Potential to relieve pressure on Emergency Departments [3]
- Potential to benefit the diagnostic efficacy of clinicians and improve quality of care.[4]

ada

Four System Challenges

1

Data

Generalizability, sovereignty, privacy, and interoperability issues

2

Bias

Under-representation leading to errors for specific groups

3

Trust

Clinical validity and transparency concerns

4

Skills

Clinical, technical, and regulatory capacity gaps

These determine where AI succeeds or fails in telehealth.

Example (national tele-clinic): Outcomes vary by language and handset quality, exposing intertwined data, bias, trust, and skills gaps to fix before scale.

Sources: Barriers to digital uptake (device, affordability, literacy, content): GSMA SOMIC hub https://www.gsma.com/somic

Challenge 1: Data

Generalizability risk

Models trained on non-local data underperform in new populations

Sovereignty, privacy, interoperability

cure, standards-based exchange

Fragmentation

Silos limit representativeness; invest in IDs, registries, consent, and FHIR profiles

Example (tele-ICU): A risk model trained abroad underperforms on remote feeds; adding local telemonitoring data and retraining restores accuracy.

Sources: Generalizability & bias from training data: Obermeyer et al., Science (2019)Data governance/sovereignty anchor: WHO Global Strategy on Digital Health 2020–2025Overview: https://www.who.int/publications/i/item/9789240020924

Challenge 2: Bias

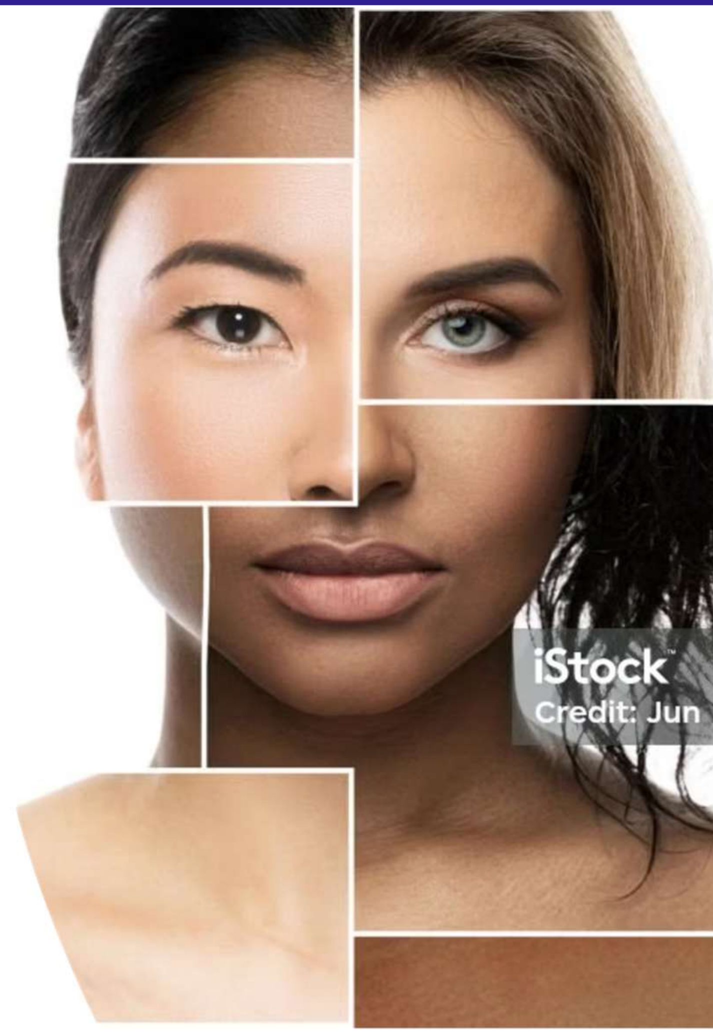
Under-representation
Skin tone, language, demographics leads to errors for specific groups

Digital Divide

Without equity checks, AI widens the digital quality divide

Equity Checks
Require subgroup analyses pre-deployment and monitor after launch

Example (tele-dermatology): A photo-based service misses lesions in darker-skin patients; local images plus subgroup testing improve sensitivity.



Source: Dermatology disparities example (skin tones): Adamson & Smith, JAMA Dermatology (2018)

Challenge 3: Trust

Clinical validity needs independent, clinically relevant evaluation

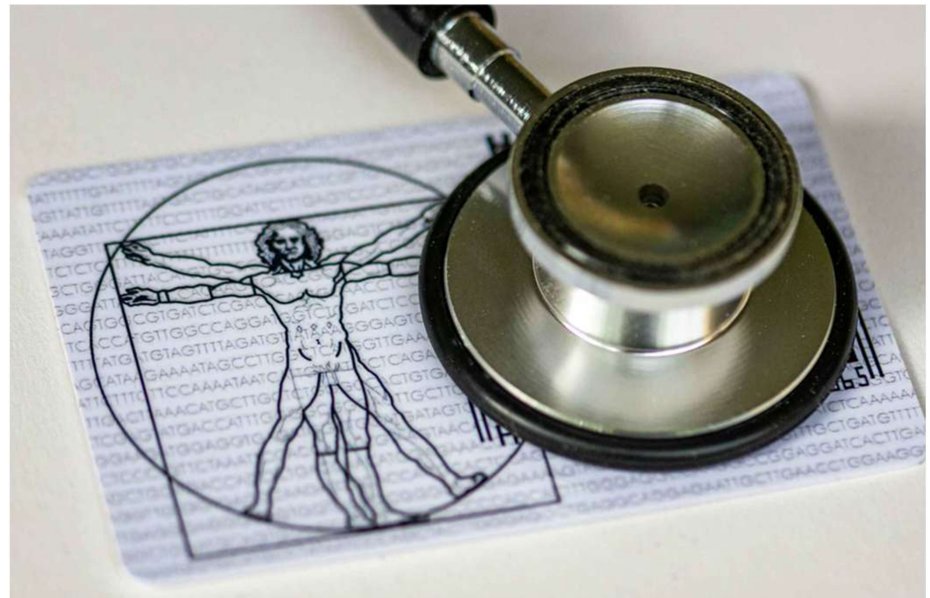
Not vendor claims

Black-box concerns reduce adoption

Ensure fit-for-purpose transparency

Establish incident reporting

Escalation and rollback paths



Example (tele-triage line): Add "why this advice" summaries, a clinical safety owner, and incident reporting; adherence and acceptance rise.

Source: Explainability & multidisciplinary view: Amann et al., BMC Med Inform Decis Mak (2020)

Challenge 4: Skills

Clinical

Train providers to interpret outputs and override when needed

Technical

Grow local data/AI engineering for tuning, validation, monitoring

Regulatory

Build authorities for review and post-market surveillance

Example (tele-ophthalmology network): Nurses capture fundus photos; a "telehealth AI lead" and data team review flags weekly and retrain quarterly.

Sources:

Regulatory capacity & SaMD oversight: WHO (2023) Regulatory considerations on AI for health

ITU Focus Group background (precursor to GI-AI4H)

Turning Challenges into Opportunities : Four Opportunity

Scale

Extend reach to underserved populations

Speed

Accelerate diagnosis and triage using AI-based tools

Accuracy

Enhance clinical decision-making

Equity

Ensure inclusive access and outcomes

Map each to telehealth workflows.

Example (integrated telefront door): AI triage (Scale), auto-intake notes (Speed), risk scores in e-visits (Accuracy), IVR in local languages (Equity).

Sources:

- Symptom checker accuracy vs GPs (vignettes): Gilbert et al., BMJ Open (2020)
- WHO digital interventions incl. client-to-provider & provider-to-provider telemedicine, health worker decision support: WHO Guideline (2019)

Opportunity: Scale

AI-based Triage and symptom advice

Manage primary-care flow and escalate urgent cases

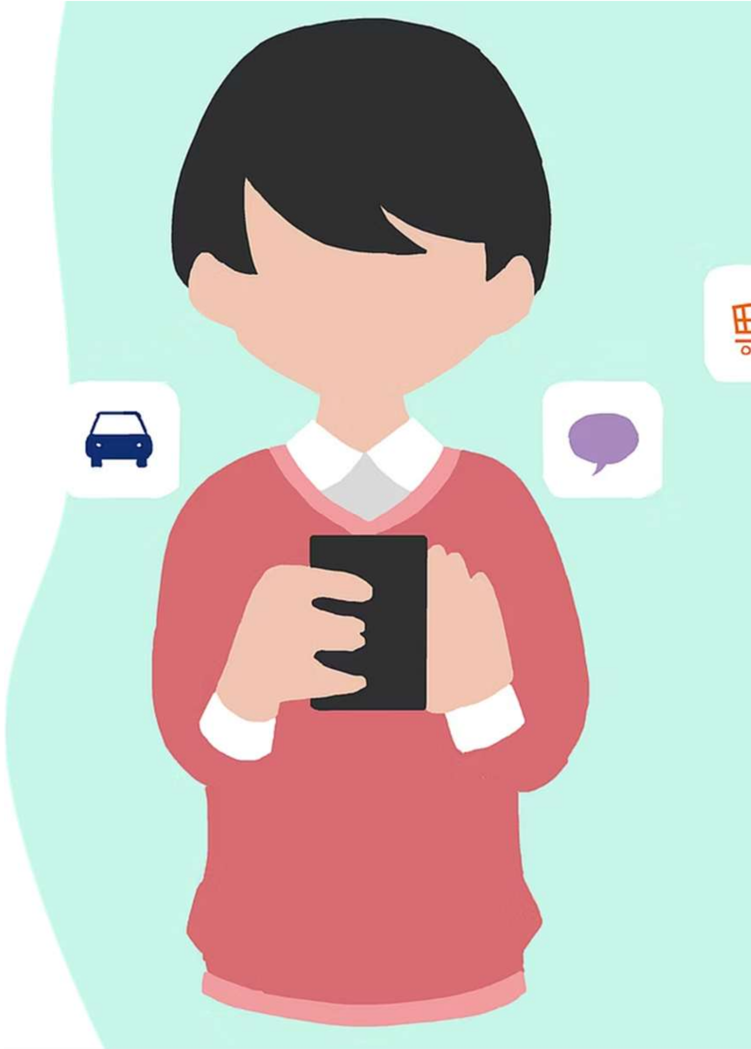
Chronic NCD support

Always-on chatbot-based coaching and monitoring (with oversight)

Frontline worker aids

Standardize assessment and referral across villages

Example (community tele-visits): CHWs use a tablet triage app; red-flag cases get same-day tele-consults; stable NCD patients receive tele-coaching and refills.

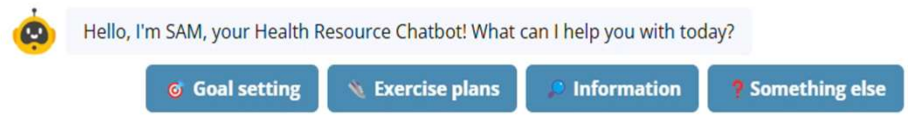


Sources:

- WHO recommendation to use CAD with chest radiography for TB screening (since 2021; updated 2025)
- Deep learning for diabetic retinopathy screening (multi-ethnic datasets): Ting et al., JAMA (2017)

AI in Telehealth: SAM
Opportunity : Speed

Locally hosted LLM based Chatbot for cancer coaching



Goal: Think, Reason, and Respond to support patients

- SAM's dynamically tailored responses are based on the user's persona and purpose
  - Doctor
  - Nurse
  - Pharmacist
- Trained on medical and scientific literature

VIVE
CENTER FOR VIRTUAL CARE VALUE AND EXCELLENCE

Opportunity : Accuracy

Augment, not replace

AI-based Decision support reduces cognitive load and surfaces relevant data

Measure patient-relevant outcomes

Not just model metrics

Symptom checkers assist triage

Use as adjuncts with oversight

Example (tele-cardiology): During e-visits, an AI panel surfaces vitals, meds, and labs; fewer missed drug-disease interactions.

Sources: Digital inclusion barriers (device cost, literacy, relevant content);WHO Digital Health Guideline (equity considerations embedded in recommendations & implementation notes):

Opportunity : Equity

Bring specialist-level support to remote centers

Via telehealth plus AI, with subgroup monitoring

Multilingual and voice interfaces

For low-literacy settings, tested with intended users

Equity by design

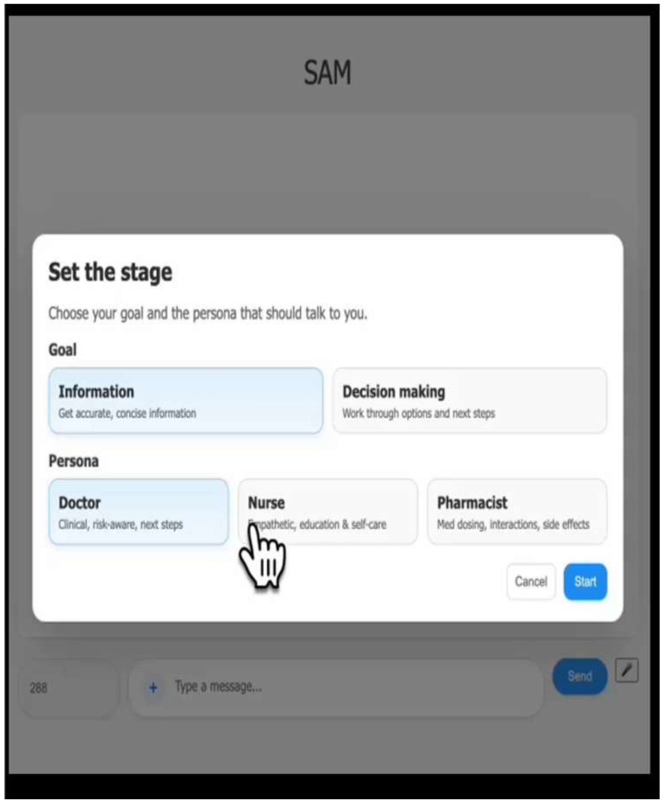
Parity goals defined before scale-up

Example (maternal tele-line): Voice bots in local languages with low-literacy menus; AI flags danger signs and schedules nurse callbacks.

Sources: GSMA mobile internet adoption barriers (language/handset), equity guardrails in WHO guidance. gsmaintelligence.com +1

SAM: Health Assistant

- Handles everything from general health questions and consultations to diagnosis support and wellness guidance
- Personalized Personas
- Conversation history saved in the database
- Prepare patients pre-telehealth encounter



CENTER FOR VIRTUAL CARE VALUE AND EXCELLENCE

UNC RESEARCH

From Principles to Practice

Two Guides

WHO ethics and governance

Compass for autonomy, safety, equity, accountability, transparency, sustainability

UK buyer's/playbook approach

Checklists for problem definition, data readiness, evaluation, procurement, safety

Example (tele-triage pilot): Map the pilot to ethics/governance checkpoints; adopt a buyer's checklist before expansion.

Sources: WHO (2021) Ethics and governance of AI for health
NHS AI Lab — A buyer's guide to AI in health and care"How to get it right"

WHO Guidance → Actionable Checkpoints

Human oversight

Clinician reviews outputs before action

Equity by design

Predefined subgroup thresholds and rollback criteria

Transparency

Clear capabilities, limits, and rationale in plain language

Accountability

Defined liability and incident reporting before deployment

Example (tele-derm safety case): Require human-in-loop review, patient notices, parity thresholds, and rollback triggers before scale.

Sources: HealthBench (LLM benchmark for health)MedPerf (federated benchmarking)Equity evaluation framework (HEAL): Google Research explainer

UK Playbook → AI in Telehealth Template

1

Problem first

Clear clinical need and success metrics

2

Data readiness

Provenance, quality, privacy, representativeness

3

Efficacy and generalizability

External validation on local populations

4

Safety and regulation

Treat as SaMD where applicable; lifecycle monitoring

Action: create your buyer's guide with shared checklists and templates

Example (tele-procurement): Define the use case, verify data representativeness, require local external validation, specify post-market monitoring.

Sources: UK "Buyer's guide to AI in health and care." GOV.UK

Benchmarking Efforts for AI in Health

AI for Health
ITU-WHO Focus Group
World Health Organization

AI for Health
An ITU Focus Group
In collaboration with WHO

WIPO

Enable
Standards, governance, policies, and guidance on evidence-based AI4H

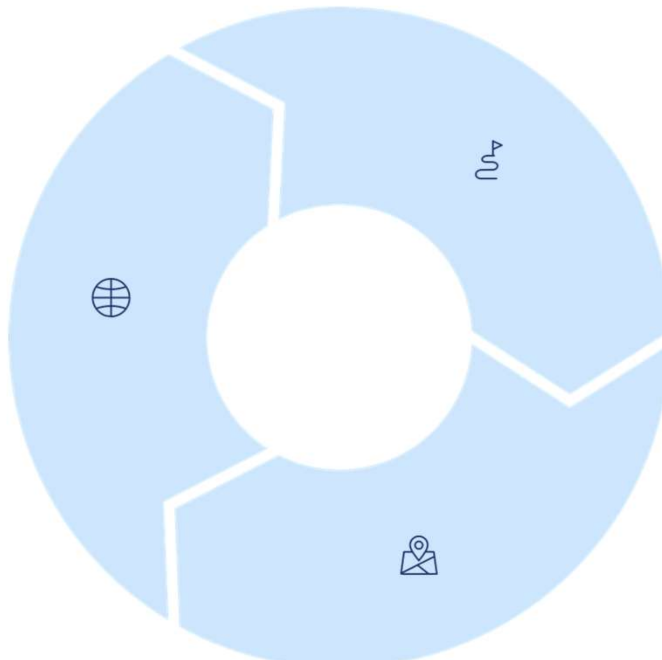
Facilitate
Pooled Investments & a global community of experts

Implement
Sustainable models of AI programs implementation at the country level

Global Initiative on AI for Health (GI-AI4H).

- Launched in 2023
- long-term institutional structure,
- Mission: to enable, facilitate, and implement AI in healthcare.

Coordination Engine



Local Context

Adaptive AI tools; Workflow Integration; Usability; User Experience; Clinical Outcomes

Global initiative for AI in health

Align standards, benchmarks, and multi-country collaboration

National

Pathway from principle to regulation to implementation

Example (regional tele-DR protocol): Countries align a shared tele-retinopathy validation protocol, pool metrics, and publish a joint implementation note.

Source: WHO's Global Initiative on AI for Health (GI-AI4H). World Health Organization

References

1. Gilbert S, Mehl A, Baluch A, Cawley C, Challiner J, Fraser H, Millen E, Montazeri M, Multmeier J, Pick F, Richter C, Türk E, Upadhyay S, Virani V, Vona N, Wicks P, Novorol C. How accurate are digital symptom assessment apps for suggesting conditions and urgency advice? A clinical vignettes comparison to GPs. BMJ Open. 2020 Dec 16;10(12):e040269. doi: 10.1136/bmjopen-2020-040269. PMID: 33328258; PMCID: PMC7745523.

1. Gräf M, Knitza J, Leipe J, Krusche M, Welcker M, Kuhn S, Mucke J, Hueber AJ, Hornig J, Klemm P, Kleinert S, Aries P, Vuillerme N, Simon D, Kleyer A, Schett G, Callhoff J. Comparison of physician and artificial intelligence-based symptom checker diagnostic accuracy. Rheumatol Int. 2022 Dec;42(12):2167-2176. doi: 10.1007/s00296-022-05202-4. Epub 2022 Sep 10. PMID: 36087130; PMCID: PMC9548469.

1. Cotte F, Mueller T, Gilbert S, Blümke B, Multmeier J, Hirsch M, Wicks P, Wolanski J, Tutschkow D, Schade Brittinger C, Timmermann L, Jerrentrup A, Safety of Triage Self-assessment Using a Symptom Assessment App for Walk-in Patients in the Emergency Care Setting: Observational Prospective Cross-sectional Study: JMIR Mhealth Uhealth 2022;10(3):e32340 URL: https://mhealth.jmir.org/2022/3/e32340 DOI: 10.2196/32340

1. Faqar-Uz-Zaman, Sara F. MD*; Anantharajah, Luxia MD*; Baumartz, Philipp MD*; Sobotta, Paula MD*; Filmann, Natalie Msc†; Zmuc, Dora MD‡; von Wagner, Michael MD§,∥; Detemble, Charlotte MD*; Sliwinski, Svenja MD*; Marschall, Ursula MD¶; Bechstein, Wolf O. MD*; Schnitzbauer, Andreas A. MD*. The Diagnostic Efficacy of an App-based Diagnostic Health Care Application in the Emergency Room: eRadaR-Trial. A prospective, Double-blinded, Observational Study. Annals of Surgery 276(5):p 935-942, November 2022. | DOI: 10.1097/SLA.0000000000005614

Thank You!

Prof Dr Saif Khairat
saif@unc.edu

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

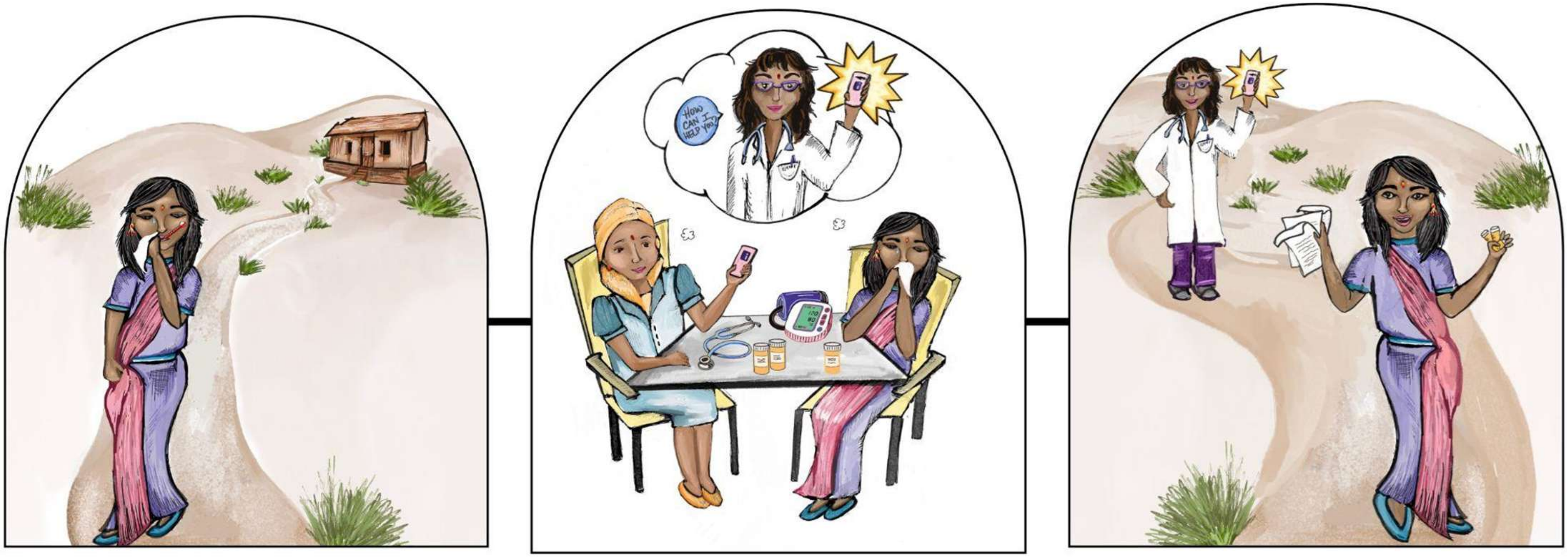
VIVE
CENTER FOR VIRTUAL CARE VALUE AND EXCELLENCE

Speaker II

https://www.beautiful.ai/player/-OdMxyIwY2sgb6GTecOb

Speaker III

GenAI-enabled Digital Assistant: Improving the Accuracy and Efficiency of Clinical Diagnosis and Treatment Planning in rural Telemedicine


HOW CAN I HELP YOU?

Ayu 2.0 – a GenAI enabled digital assistant

Differential Diagnosis (DDx)

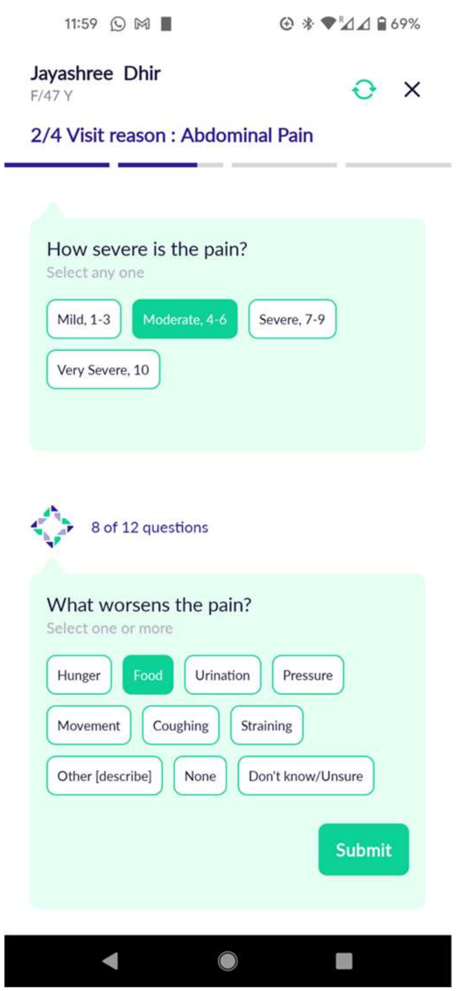
Suggests top 5 possible diagnoses based on symptoms, vitals, and history.

Treatment Plan (TTx)

Recommends medications, tests, referrals, follow-up, and lifestyle tips, aligned with medical protocols.

Patient Story

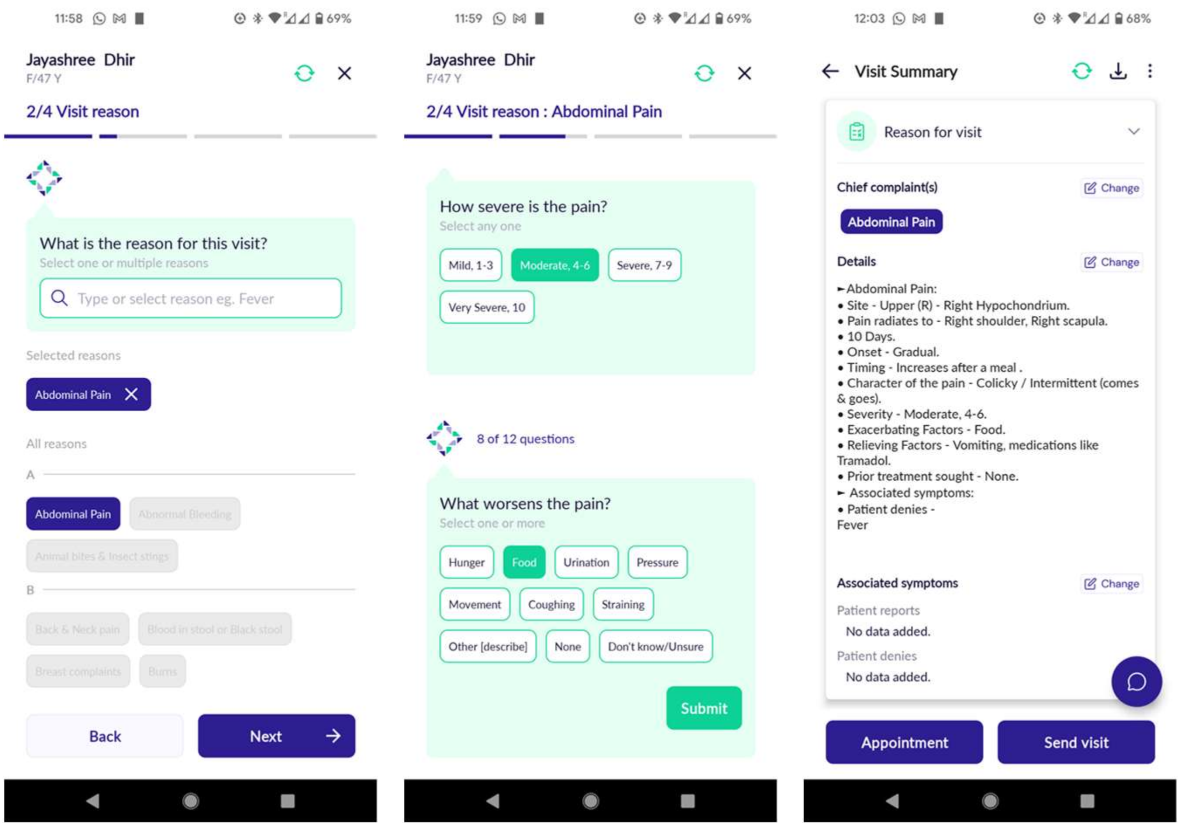
Generates a quick, clear summary of the case for doctors to fast-track decision-making.



Development & Testing Approach

01	Benchmarking Datasets	• Curated test and training datasets
02	Evaluation Methodology	• Top-5 DDx accuracy (doctor-reviewed and LLM-validated) • QUEST scoring
03	Model Testing	• Benchmarked 12 different AI models
04	Prompt Optimization	• DSPy • In-context learning • Few shot learning
05	Alternative Techniques No improvement in accuracy	• Medically fine-tuned LLMs • Reasoning-based models • Reinforcement learning

Ayu Platform – Rules based history for patients & frontline health workers

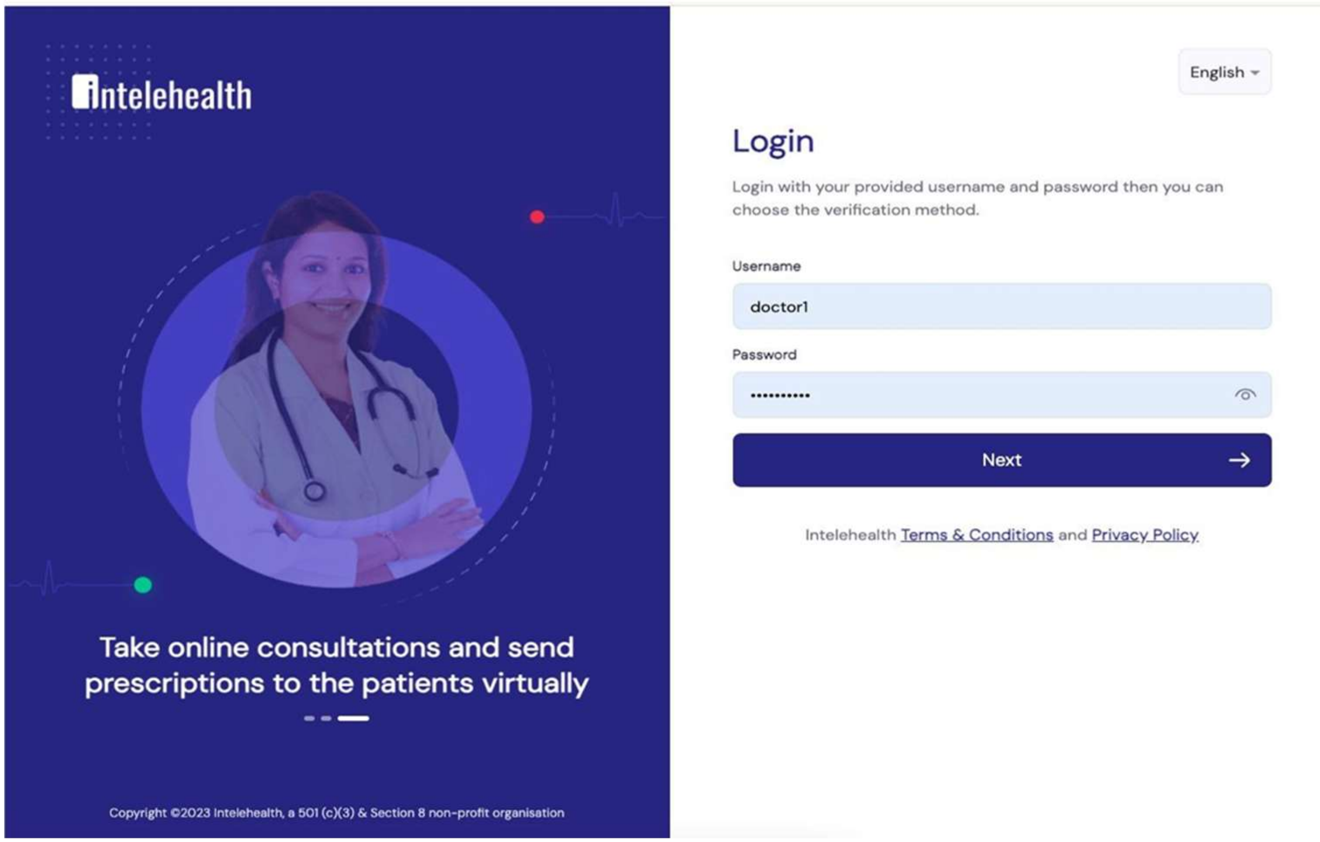


250+ clinical protocols

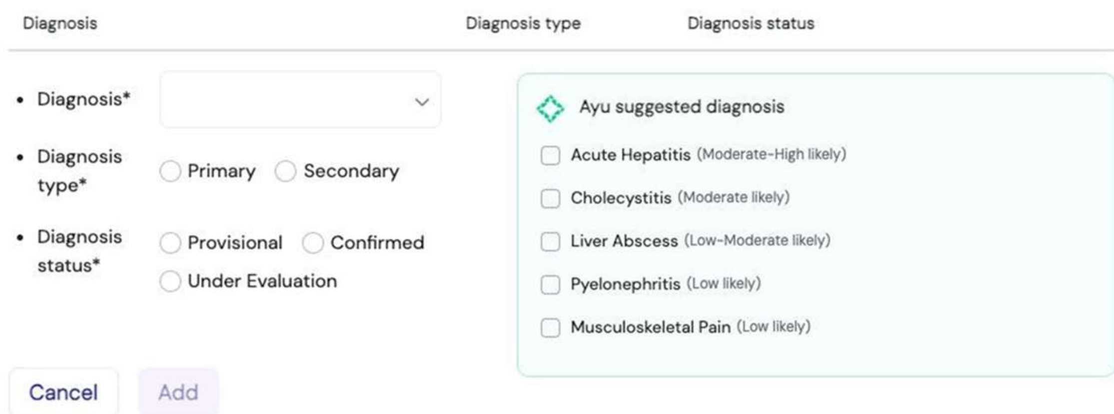
15 languages

for history taking, physical examination and clinical decision support for common conditions.

Ayu Platform – Web Interface for Doctors



How do the Doctors make use of AI generated Differential Diagnosis



Model	Top 1 DDx Accuracy	Top 5 DDx Accuracy
Gemini 2.5 Flash	67%	99%
Open AI GPT 4.1	54%	94%
LLama 4 Maverick 17B 128E	75%	98%

✅ The service is pluggable to these LLMs

1.Gemini 2.5 Flash

2.Open AI GPT 4.1

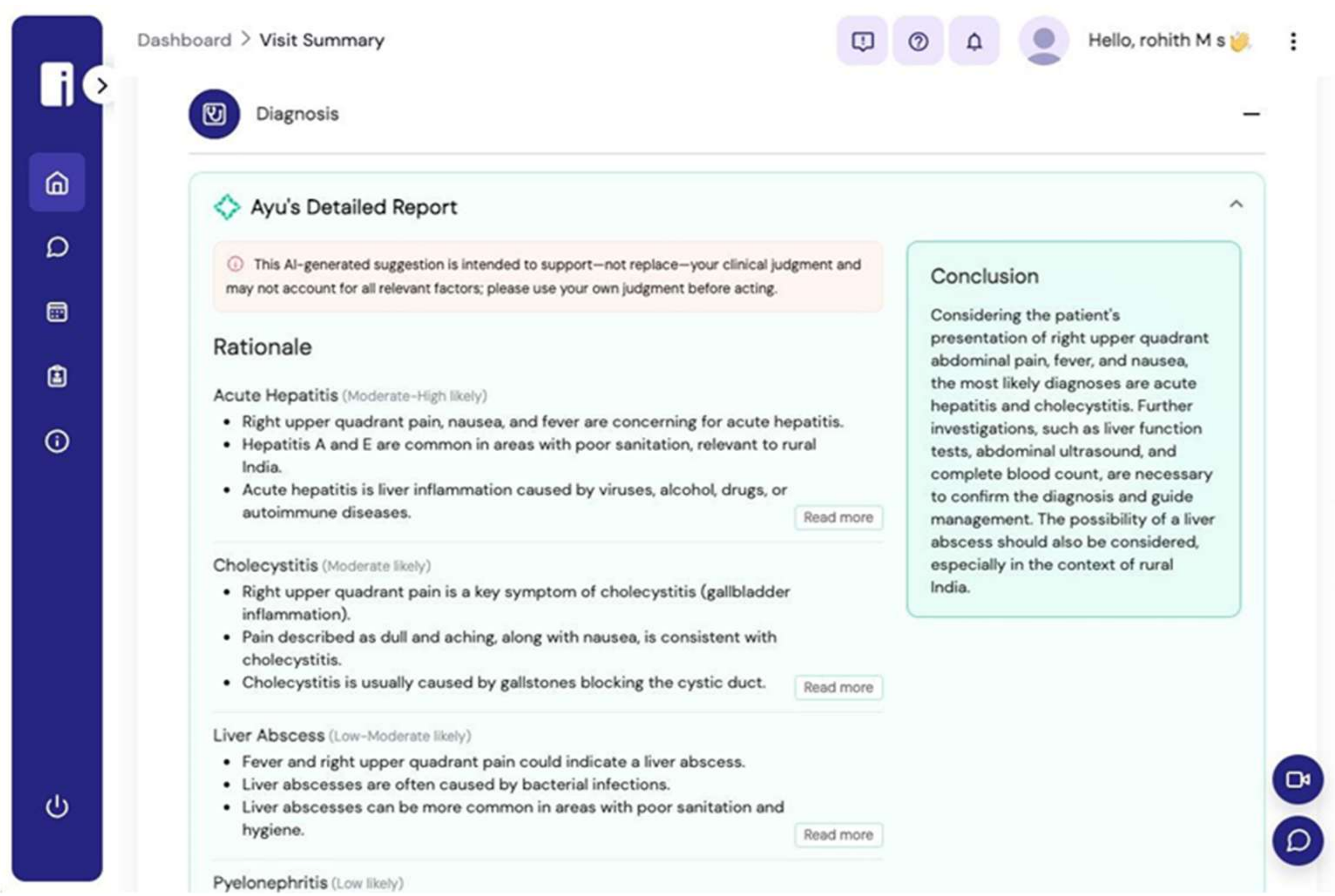
3.Llama 4 Maverick MOE 128 Experts

✅ The models have a contextualization layer on top of the base model to tune their outputs and safety-gate outputs

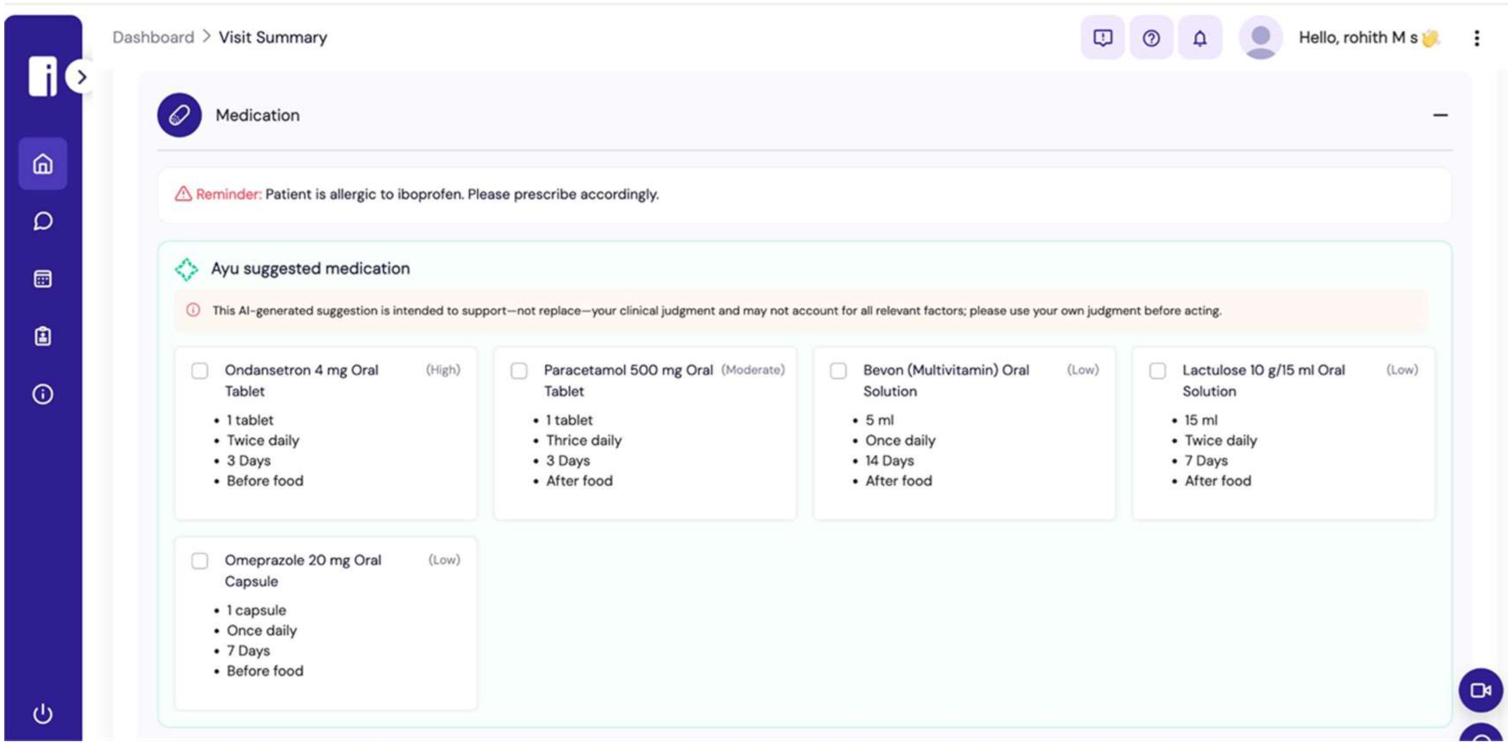
✅ The contextualization layer is trained on real-world telemedicine use data.

✅ **The system - base model + contextualization layer** is evaluated on carefully curated datasets from Intelehealth's telemedicine programs

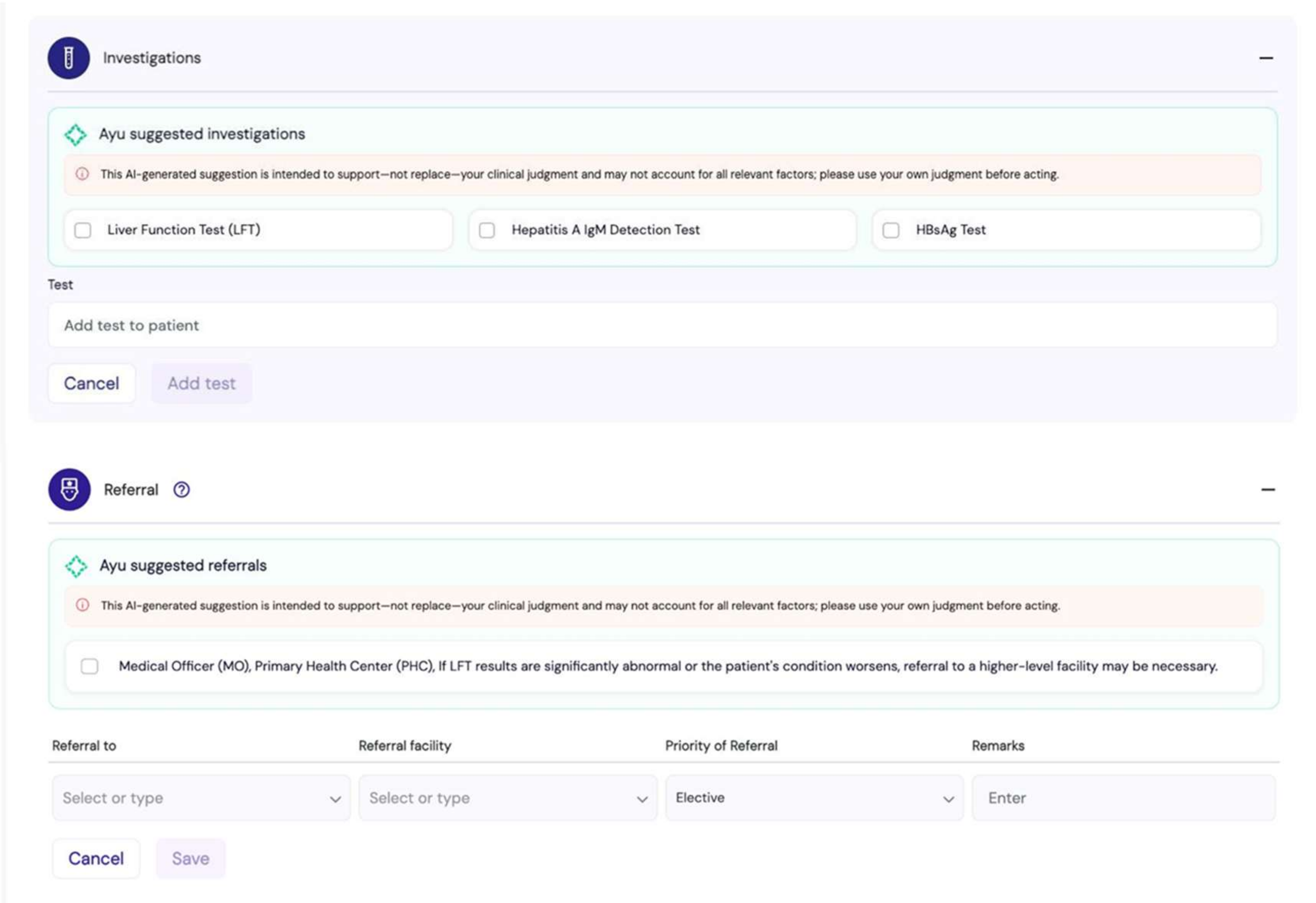
DDx Interface In Ayu Platform



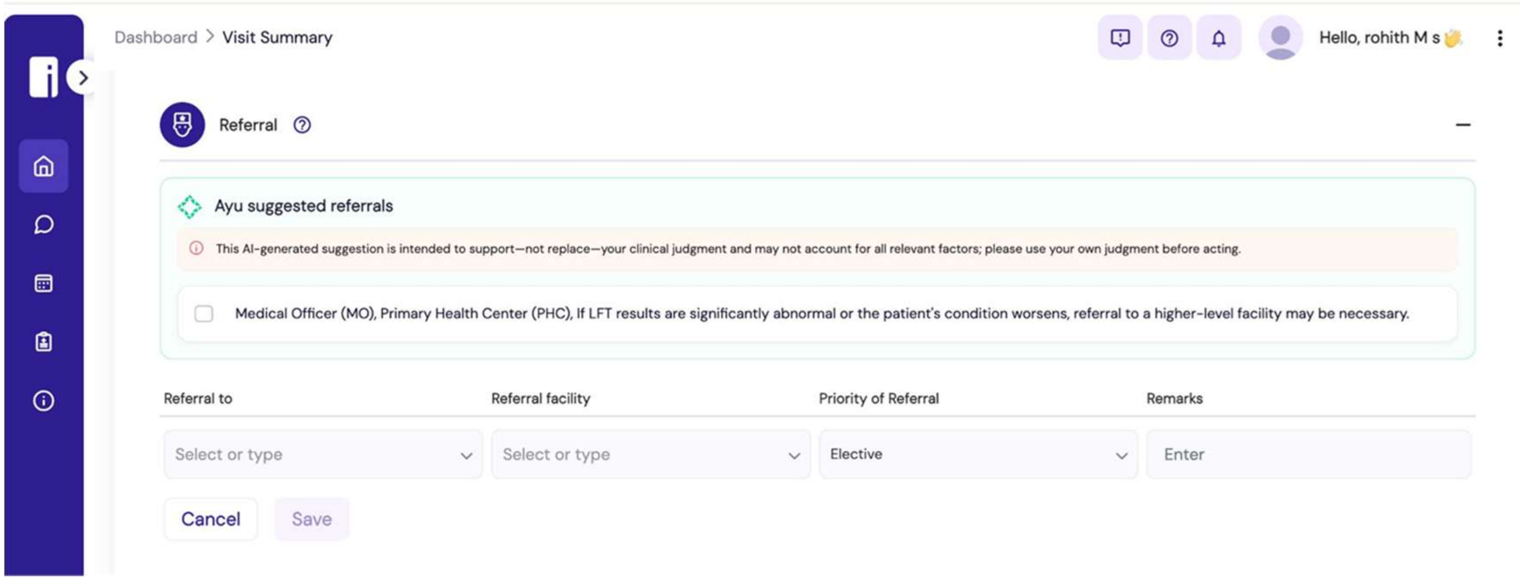
Medication prediction in Treatment Plan (TTx)



Investigations & Referrals Prediction in Treatment Plan (TTx)



Referral Prediction in Treatment Plan (TTx)



Background and Rationale

Access to healthcare in India's rural and tribal regions remains limited due to geography, infrastructure gaps, and shortages of doctors.

Arogya Sampada, implemented by Intelehealth, uses telemedicine to extend medical consultations to underserved communities.

Clinicians delivering remote care often face challenges such as high patient loads, incomplete case histories, and limited diagnostic inputs, leading to variability in care quality.

Artificial intelligence (AI) offers an opportunity to strengthen these systems. Such tools could enhance diagnostic accuracy, consistency, and efficiency in telemedicine. However, evidence from real-world, low-resource settings remains scarce

Thus, this evaluation assessed the impact of integrating an AI-powered diagnostic assistant within *Arogya Sampada* on physicians' diagnostic accuracy, consultation time, and treatment quality.

Study Objectives

Objective

To evaluate the effect of large language model (LLM) assistance on diagnostic accuracy using a **randomized evaluation with a matched case-control design**, where each clinical case is reviewed once by an AI-assisted doctor and once by a non-assisted doctor.

Research Questions

- Does AI assistance optimised for enhancing diagnostic accuracy and treatment planning compared to clinicians using traditional and/or non-LLM-assisted resources?
- Does the AI assistance impact the time taken to diagnose?
- Does the AI assistance influence the clinical appropriateness and quality of prescriptions and treatment plans provided by clinicians?

Study Design

Hypothesis

LLM assistance improves diagnostic accuracy from 80% to 95%.

Statistical Parameters

- Outcome: Binary (correct vs incorrect diagnosis)
- Significance level (α): 0.05 (two-sided)
- Power (1 – β): 80%

Sample Size Calculation

Using the formula for matched-pair binary outcome studies:

$Z_{1-\alpha/2}$ is 1.96 for α = 0.05
$Z_{1-\beta}$ is 0.84 for power = 80%
Odds Ratio (ψ) is 0.21
Required matched pairs according to McNemar's sample size calculation : 689

Final Case Requirements

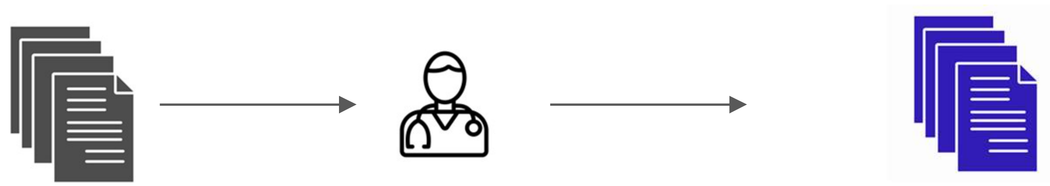
- Each of the **10 pairs** will evaluate **100 unique clinical cases**
- Each case is seen by **2 doctors only** (one LLM-assisted, one without LLM assistance)
- Each doctor evaluates **100 cases**

Total cases=10 pairs× 100 =1000

$$N_{pairs} = \frac{\left(Z_{1-\alpha/2}(\psi+1)+Z_{1-\beta}\sqrt{\left\{(\psi+1)^2-(\psi-1)^2\pi_{discordant}\right\}}\right)^2}{(\psi-1)^2\cdot\pi_{discordant}}$$

Study Design

Step 1: Creating a benchmarking dataset

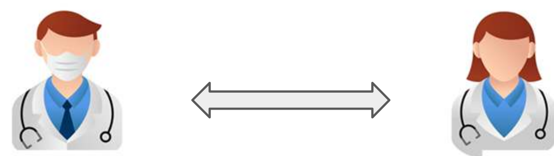


EHR data of 50000 patients

Clinicians' Review

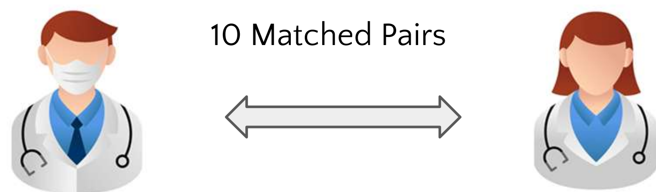
1000 cases selected as a benchmarking dataset with ground truth

Step 2: Recruiting study participants



10 Matched Pairs

Step 3: Randomizing cases and AI assistance



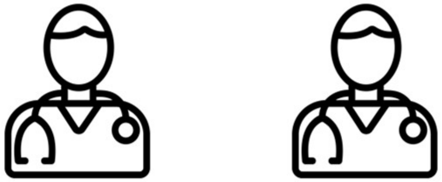
Case 1	Assisted	Not assisted
Case 2	Not assisted	Assisted
Case 3	Assisted	Not assisted
.	.	.
.	.	.
.	.	.
Case 100	Not assisted	Assisted

Study Design

Step 4: Completing consultations

Step 5: Evaluating case results

Independent clinicians' evaluation of case results



1. Consultation details captured
2. DDx Captured
3. Diagnosis and treatment plan captured

Overview of Evaluation Parameters and Scoring Methods

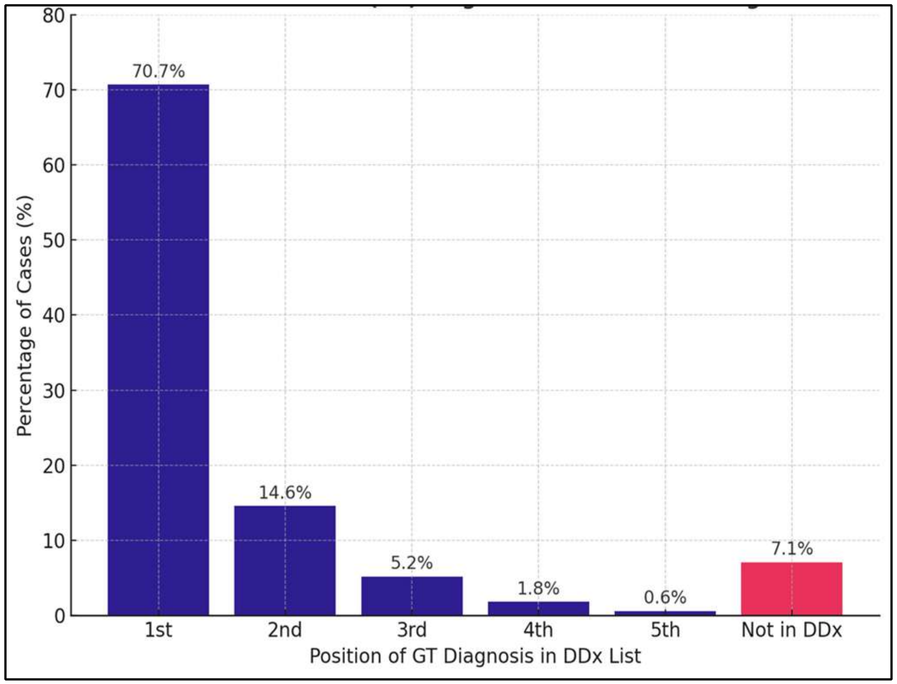
Parameter	Description
Diagnostic Accuracy	Binary: Correct if the clinician's diagnosis matched the ground truth; Incorrect otherwise.
Turn around time	Duration from case initiation to final submission, measured in minutes and seconds, available from the telemedicine platform.
Quality of Differential Diagnoses	Evaluated only in the LLM-assisted group;Top-5 Accuracy(whether Ground truth diagnosis present in top 5 + rank), Appropriateness*, Comprehensiveness*
Quality of Treatment Plan	Degree of alignment between clinician's plan and treatment plan based on ground truth diagnosis. Completeness*and Medical Appropriateness Index (MAI) as given by Hanlon et. al. (Hanlon, 1992)
Quality of Medical Tests	Relevance and necessity of investigations.*
Quality of Medical Advice	Clinical relevance, clarity, and safety of advice given.*
Quality of Referral Advice	Necessity and appropriateness of referral decisions or omissions.*

*Using a 5-point scoring scale (Bond, 2012; McDuff, 2025)

Key Findings

Quality of LLM generated DDx

Figure 3: Rank of The Ground Truth diagnosis in the LLM-generated Differential Diagnosis list

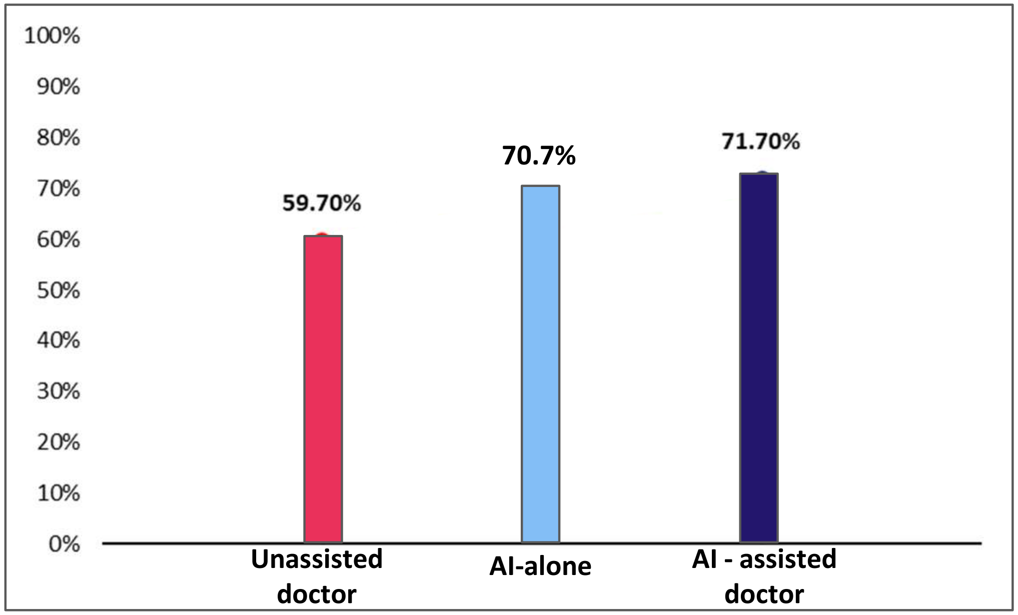


- **The ground truth diagnosis appeared within the top three positions in 90.5% of cases**, most commonly ranked first (70.7%), while only in 7.1% of cases it was not listed in the DDx.
- **90.8% of AI-generated DDx were rated as very appropriate**, 4.8% as appropriate, and 4.4% as neutral, indicating a high overall standard of diagnostic relevance.

Key Findings

Diagnostic Accuracy

Figure 2: Comparison of Diagnostic accuracy Between Unassisted and AI-assisted groups.



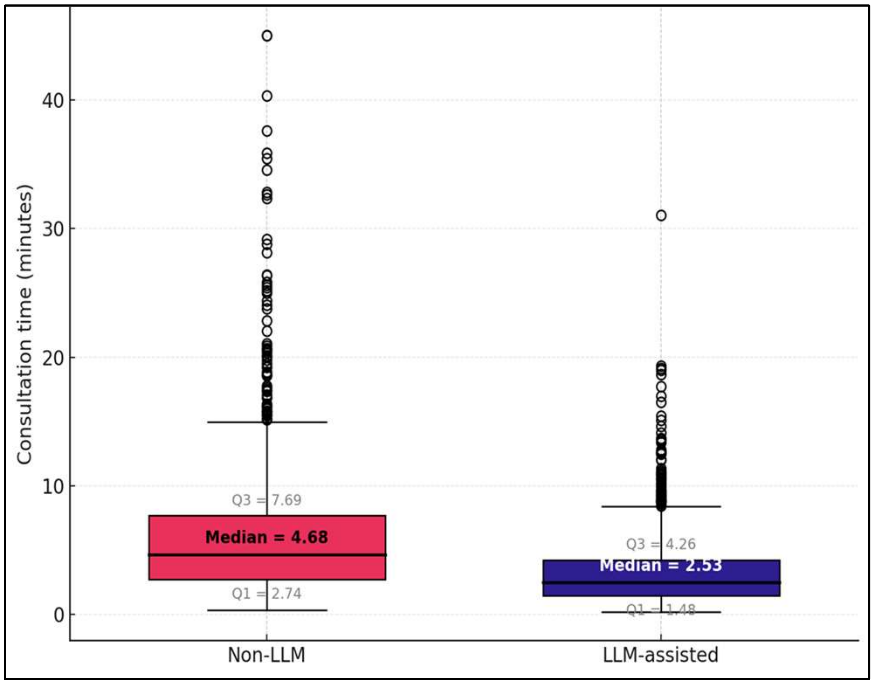
Diagnostic accuracy was significantly higher with AI assistance (71.7%) compared to the unassisted group (59.7%), **showing a 12% absolute improvement** (N=1000, χ^2=39.34, p<0.001).

Also noteworthy was that AI alone had the right diagnosis in the number 1 position **70.7%** of the time.

Key Findings

Time to diagnose

Figure 2: Comparison of Time taken Between Unassisted and AI-assisted arms.



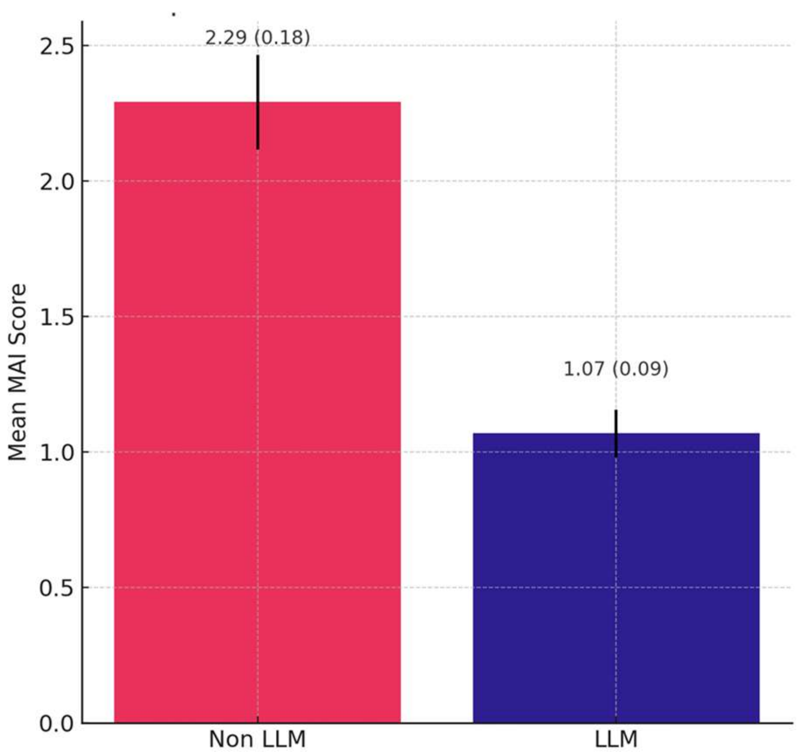
The mean time taken was notably shorter with LLM support (2 minutes 25 seconds vs. 4 minutes 35 seconds), reflecting a 47% reduction.

Mean difference = **2.16 minutes i.e 2 minutes 10 seconds**, t(925)=12.15, p<0.001.

Key Findings

Treatment Plan

Figure 4: Comparison of Medication Appropriateness Index (MAI) Between Unassisted and AI-assisted arms.



- Treatment plans developed with AI assistance demonstrated greater **completeness, with a median score of 4 (mean = 3.94 ± 0.01) compared to 3 (mean = 3.26 ± 0.01)** in the unassited arm.
- The Medication Appropriateness Index (MAI), used to assess prescribing quality, was significantly lower (indicating higher appropriateness) in the **AI-assisted arm (mean = 1.07 ± 0.09) than in the unassisted arm (mean = 2.29 ± 0.18)**.

Key Findings

Medical Test, Medical Advice and Referral Advice

Among 1000 cases, the appropriateness of medical recommendations was consistently higher in the unassisted arm across all domains. Notably, a higher proportion of unassisted responses received the top rating of 5 compared to the assisted group.

Table 1: Comparison of Appropriateness Scores for Medical Test, Medical Advice, and Referral Advice Between Unassisted and Assisted arms.

Appropriateness of (N=1,000):	Unassisted Score Mean (SE)	Assisted Score Mean (SE)	Unassisted: % Rated 5	Assisted: % Rated 5	p-value
Medical Test	4.47 (0.03)	3.86 (0.03)	78.30%	42.30%	< 0.001
Medical Advice	4.88 (0.02)	4.17 (0.05)	95.10%	70.50%	< 0.001
Referral Advice	4.80 (0.01)	4.53 (0.04)	90.50%	81.60%	< 0.001

Key Findings

Qualitative Insights: Clinicians' perceptions

The AI as a "Cognitive Co-Pilot"

"...it will remind us of the diseases we skip or we forget, so it will not let us make any human error."

"But whatever I experienced, whatever we thought in our mind during making a diagnosis, it was already there in that your section, in provisional diagnosis and medication."

"I usually go with my thought process... and then I cross-check if it's there or not in the diagnosis list."

Key Findings

Qualitative Insights: Clinicians' perceptions

The "Reality Gap" of Context

"It will be better if you design it for a more complicated case."

"Some of the medications are not listed... so I have to give my own input."

Barriers to Widespread Adoption

"If there is training it is possible to use it regularly"

Key Learnings

Successes

- **Enhanced Performance:** AI significantly enhanced **diagnostic accuracy** and **efficiency** (reduced consultation time by nearly half), strengthening structured assessments.
- **Cognitive Co-Pilot:** Clinicians viewed the AI as a "cognitive co-pilot" that **validated reasoning** and **expanded diagnostic thinking**, not a replacement for judgment.
- **Proven Safety:** No unsafe or clinically inappropriate outputs were reported in the 1000 interactions, demonstrating the AI assistant is a **safe and valuable support tool**.

Areas for Refinement

- **Contextual Adaptation:** Need to further **train the model on locally relevant data**, building awareness of referral networks, infrastructure constraints, and context-specific clinical pathways.
- The primary design and deployment challenge is ensuring the model's outputs are aligned with **feasible options** and **local health realities**, not just ideal clinical standards.
- **Integration and Training:** Essential next steps include improving **integration within existing clinical workflows** and ensuring adequate **user training** before broader field deployment.

Limitations

- **Controlled & Retrospective Design:** The absolute diagnostic accuracy values should be interpreted with caution. Clinicians were diagnosing based only on written case vignettes and could not interact with patients or gather additional clinical context. These results therefore reflect performance under controlled, retrospective conditions rather than real-world teleconsultations.

- **Limited Contextual Familiarity:** Participants were **not directly embedded** in the communities represented, resulting in limited familiarity with local disease patterns, patient context, and health-system norms. This unfamiliarity may have affected the clinicians' ability to provide fully contextualized diagnoses and appropriate **referral decisions**.

- **AI Model's Contextual Gap**: Non-contextual AI suggestions may have **biased clinician decisions** toward ideal but locally irrelevant choices.

Way Forward

- **Deepen Analytical Understanding:** Conduct further analysis to assess how the quality and clinical depth of case information affects accuracy for clinicians in both the LLM and non-LLM arms.
- **Validate Performance In-Field:** Execute a prospective field study to better understand performance, integration, and safety in real-life settings, moving beyond controlled, retrospective vignettes.
- **Prioritize Contextual Model Training:** Ensure LLM suggestions align with feasible options and local realities, not just ideal clinical standards, to improve test and referral appropriateness.
- **Develop Integration & Training Strategy:** Focus on seamlessly integrating the AI into existing clinical workflows for maximum efficiency and adoption.

Thank you

Conclusion: Telemedicine Can Be A Crucial Role In LMIC's Through Transforming Health Equity

intelehealth

www.intelehealth.org

WHO SEARO + Intelehealth webinar series

www.intelehealth.org/webinars

Objectives:
Learn how telemedicine can address challenges and enhance health systems

Expected Outcomes:
By the end of the session, participants will:

- Gain a foundational understanding of telemedicine and its key components.
- Learn from successful case studies of national and sub-national public sector telemedicine implementations.
- Understand key policy and regulatory considerations for integrating telemedicine into national health systems.
- Be equipped with practical insights to explore and implement telemedicine solutions in your contexts.

Webinar Topics and Dates

Sno	Date	Topic
1	06 March 2025	What is Telemedicine and How Are Health Systems Using It Globally? A Primer for Health System Leaders
2	10 April, 2025	Brick-and-mortar to Brick-and-click - Designing & Implementing Quality, Effective, and Impactful Telemedicine Programs
3	08 May, 2025	Evaluating telemedicine interventions: Evidence so far, and Methodologies
4	5 June, 2025	Creating a Telemedicine-Ready Healthcare Workforce: Training for Healthcare Providers
5	10 July, 2025	Telemedicine Policy: How Telemedicine is Regulated in Asia
6	7 August, 2025	Choosing a Telemedicine Software: The case for standards-compliant, interoperable & open-source Digital Public Goods (DPGs)
7	11 September, 2025	Ensuring Quality of Care & Patient safety in Telemedicine
8	9 October, 2025	Telemedicine Adoption by Communities – How Might We Drive Uptake of Telemedicine (TM) by Citizens?
9	6 November, 2025	Artificial Intelligence and Machine Learning in Telemedicine
10	11 December, 2025	Financing Telemedicine and ROI – The Business Case for Telemedicine
11	8 January, 2026	Telemedicine use cases to advance the SDGs - Part 1 Applications for Non-Communicable Diseases (Diabetes, Hypertension, Cardiovascular disease, Cancer and Mental Health)
12	5 February, 2026	Telemedicine uses to advance the SDGs - Part 2 Applications for Communicable Diseases (Tuberculosis, HIV)
13	12 March, 2026	Telemedicine use cases to advance the SDGs - Part 3 Applications for Primary Healthcare

Webinar Evaluation and Feedback

Thank You for Attending!

Access the recording and slides at: https://intelehealth.org/webinars/

Please take a few minutes to fill out our feedback form – your input is invaluable!

https://forms.gle/9bn92whEPArfRgHG6

Q&A Session

Thank You For Joining Us!

www.intelehealth.org | sindhura@intelehealth.org , rishi@intelehealth.org, neha@intelehealth.org