UNIVERSITY OF THE PHILIPPINES
CENTER FOR INTEGRATIVE AND DEVELOPMENT STUDIES

DISCUSSION PAPER SERIES 2026-34

AI for whom?

Interdisciplinary Governance Framework for AI in Philippine Healthcare

Iris Thiele Isip Tan MD, MSc, *Lisa Traboco MD*, *Michael Fong MD*

Program on Health Systems Development

ISSN 2619-7448 (PRINT)
ISSN 2619-7456 (ONLINE)

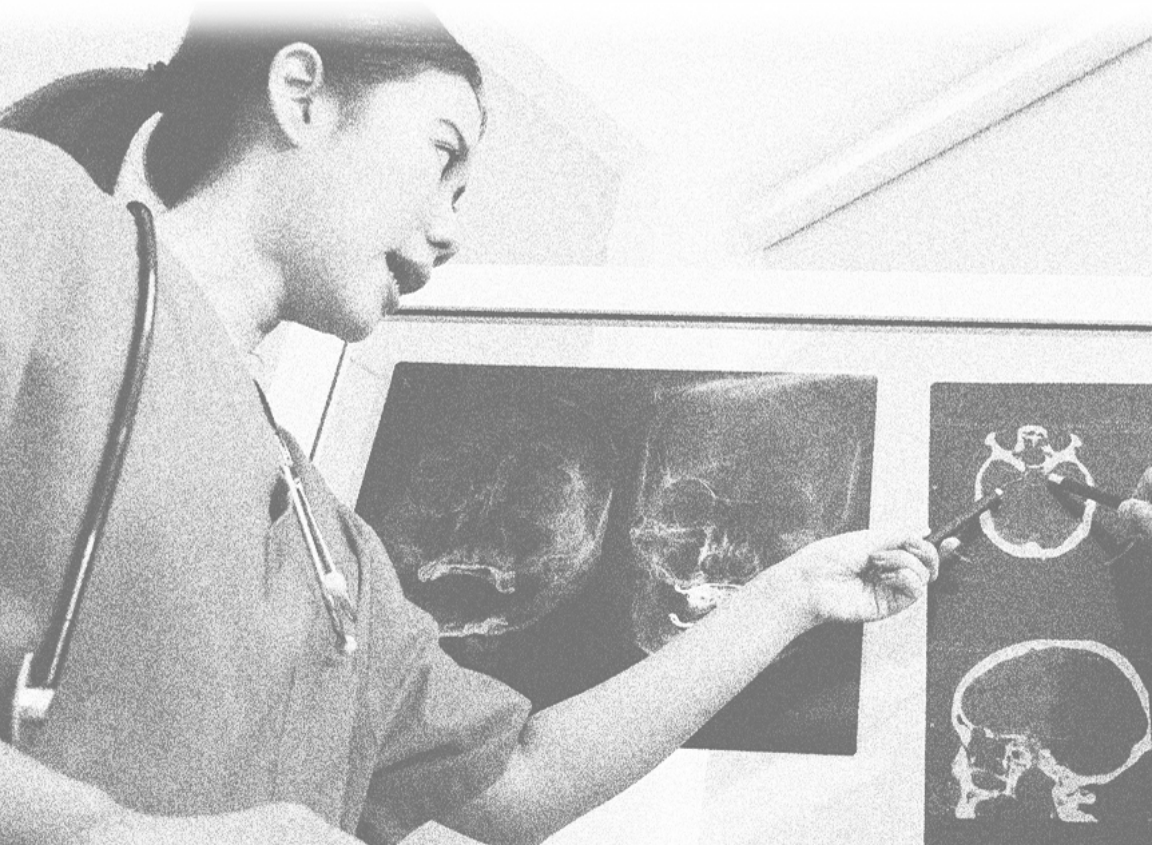
DISCUSSION PAPER SERIES 2026-34

Program on Health Systems Development

AI for whom?

Interdisciplinary Governance Framework for AI in Philippine Healthcare

Iris Thiele Isip Tan MD, MSc, Lisa Traboco MD, Michael Fong MD



UP CIDS DISCUSSION PAPER SERIES
2026-34

UP CIDS Discussion Paper Series is published by the

University of the Philippines
Center for Integrative and Development Studies

Lower Ground Floor, Ang Bahay ng Alumni
Magsaysay Avenue, University of the Philippines
Diliman, Quezon City 1101

Telephone: (02) 8981-8500 loc. 4266 to 4268 / (02) 8426-0955
Email: cidspublications@up.edu.ph
Website: cids.up.edu.ph

cids.up.edu.ph

ISSN 2619-7448 (Print)
ISSN 2619-7456 (Online)

Cover Image Credit

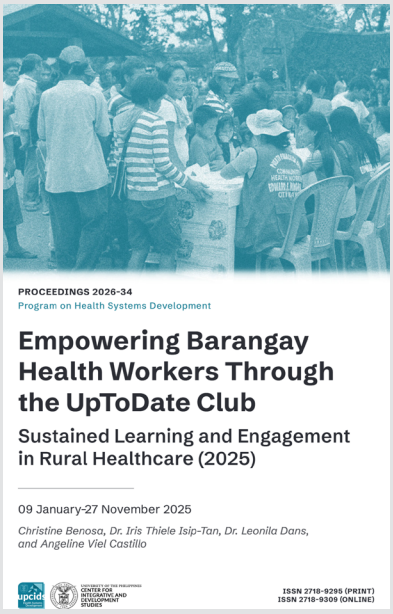
"Doctor and medical assistants discussing about diagnosis result on x-ray film"

https://www.vecteezy.com/photo/11271665-doctor-and-medical-assistants-discussing-about-diagnosis-result-on-x-ray-film

Table of Contents

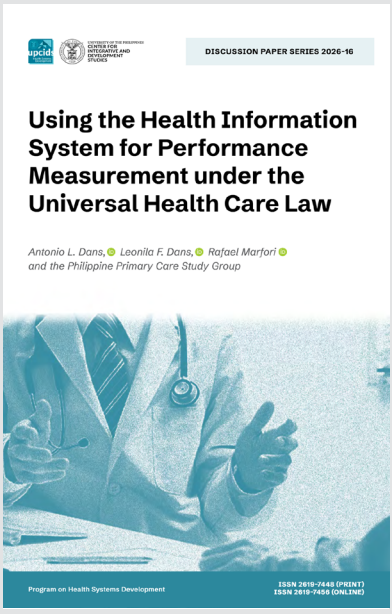
Download related policy papers for FREE

cids.up.edu.ph/collection-database



PROCEEDINGS

Empowering Barangay Health Workers Through the UpToDate Club: Sustained Learning and Engagement in Rural Healthcare



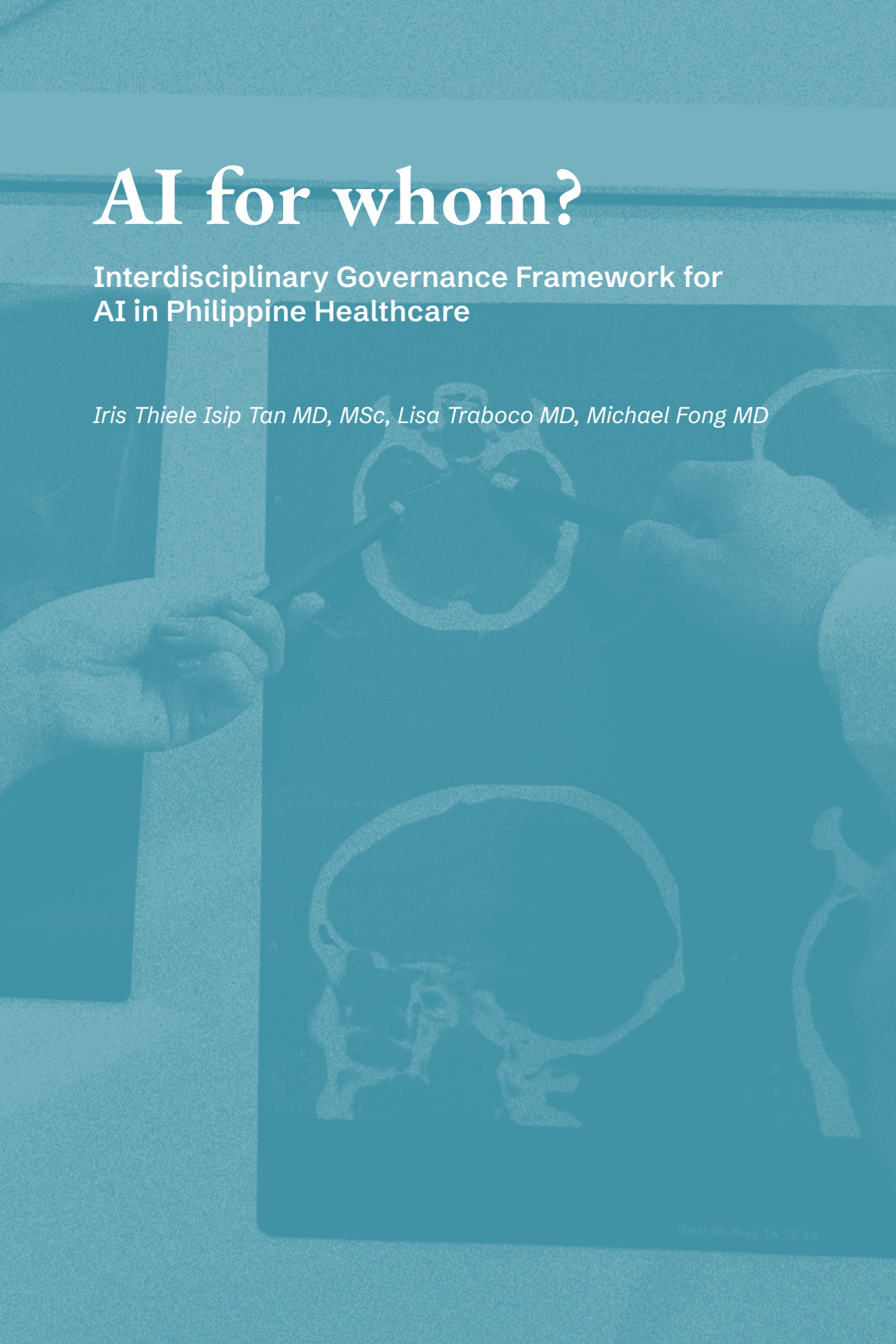
DISCUSSION PAPER

Using the Health Information System for Performance Measurement under the Universal Health Care Law

AI for whom?

Interdisciplinary Governance Framework for AI in Philippine Healthcare

Iris Thiele Isip Tan MD, MSc, Lisa Traboco MD, Michael Fong MD



Key Highlights

Policy Landscape

The Philippines enacted the Universal Healthcare Act and developed the National AI Strategy (NAIS PH) and NICS-AI Literacy Framework, though digital infrastructure remains uneven and most lacking in geographically isolated and disadvantaged areas, where the risk of AI widening existing health inequities is greatest.

Main Argument

Siloed development of AI tools, where developers work in isolation from clinicians, patients, and communities, neglects the social, ethical, and organizational factors essential for successful real-world implementation. Interdisciplinary collaboration across the AI lifecycle, with patients as essential partners, is necessary for effective, safe, and equitable AI in healthcare. While several frameworks advocate for sustained interdisciplinary engagement, most lack concrete mechanisms for maintaining it beyond the design stage. Localization adequacy, relational trust, and accountable decision authority are cross-cutting governance requirements. These need to be embedded in institutional AI governance structures and continuously enacted across the entire AI lifecycle.

Methodology

A search for relevant AI frameworks was conducted using PubMed, Google Scholar, and Research Rabbit. Frameworks were included if they addressed the AI lifecycle, interdisciplinary collaboration, or both. The included frameworks were synthesized across two dimensions: lifecycle coverage and mechanisms for interdisciplinary engagement, including patient inclusion. The AI-CAD (artificial intelligence–computer-aided detection) tuberculosis screening implementation in the Philippines was analyzed as an empirical case study illustrating where cross-cutting governance requirements were absent and what gaps resulted.

Findings

While existing AI frameworks cover the full lifecycle, they are designed for high-income countries and lack concrete mechanisms for sustained interdisciplinary engagement. Analysis of the AI-CAD TB screening case study revealed governance gaps.

Policy Recommendations

We propose a dual-layer interdisciplinary framework localized to the Philippine context. The first layer organizes stakeholder questions across the AI lifecycle. The second layer specifies three cross-cutting governance requirements that must be revisited at every stage. The framework addresses contextual challenges unique to LMICs. It incorporates *kapwa* (shared Filipino identity and relationality) and *pakikiramdam* (empathetic attunement), which respects the communal and relational nature of Filipino healthcare.

Introduction

A. Ensuring the Promise of AI in Healthcare

Artificial intelligence (AI), through rapid processing of medical data, is envisioned to improve diagnostic precision, aid personalized treatment, and drive robotics automation (Olawade et al. 2024). Health professions educators hope that AI can augment human intelligence by harnessing collective knowledge from healthcare systems (Lomis et al. 2021). The premise is that AI can assist health workers who face information overload from electronic health records, imaging, and monitoring devices, potentially improving the quality of care. Analysis of large research datasets using AI tools may accelerate drug development and innovation (Ocana et al. 2025). While AI for clinical care faces the hurdles of acceptance, validation, and regulation (Mennella et al. 2024), there appears to be less resistance in using AI for administrative tasks such as scheduling and documentation (Henry, 2025). The potential of AI, however, is matched by its potential to cause harm. The World Health Organization (WHO 2021) warns that unchecked optimism can lead to inappropriate use of AI, perpetuating bias and inequity, and causing harm.

For the Philippines, the Department of Science and Technology (DOST) spearheads the National AI Strategy (NAIS) PH (Marfal 2025), which intends to build a collaborative AI ecosystem through workforce development, infrastructure, governance, and research and innovation. However, past implementation of digital health tools, such as electronic medical records, for example, has varied widely (Elepaño et al. 2025). Most government-run health facilities lag in digital transformation, especially those in geographically isolated and disadvantaged areas. Partly to mitigate this, and because the Philippines is an archipelagic nation, the DOST will establish a national High-Performance Computing (HPC) center, with regional HPC sites (Marfal 2025). The problem will remain, however, to be "the last mile" as the country's digital infrastructure remains lacking where it is most needed. AI will just be another technology that widens the digital divide. Localization and accounting for rural primary-care physician workflows constitute another challenge when AI is primarily developed in highly urbanized and high-income countries (Ciecierski-Holmes et al. 2022).

The Philippines is also transitioning to universal healthcare. Mandatory health information systems and the National Health Data Repository (NHDR) are specified in the Universal Health Care (UHC) Act (The LAWPHiL Project

2019). However, interoperability has remained elusive despite the specification of standards by the Department of Health. The mandate for comprehensive health data collection through the NHDR could, if successfully implemented, provide datasets for AI model development and validation. However, data quality, completeness, and representativeness remain significant challenges. Health facilities vary widely in their data collection capacity and data quality, with underreporting and incomplete documentation common in resource-constrained settings (Murai et al. 2022).

The UHC Act also provided for the creation of the Health Technology Assessment Council (HTAC), which assesses the value for money of any health technology to be funded by the government (The LAWPHiL Project 2019). This poses a problem for AI in healthcare as the HTAC might require validation of efficacy by the Philippine Food and Drug Administration (FDA) and cost-effectiveness analyses. The regulation of AI falls under FDA guidance on software as a medical device (FDA 2025).

With the transition to UHC, colleges for health professions need to pivot to "UHC-ready" curricula (Pepito et al. 2025). This means including more digital health skills, and possibly, through this, adding AI competencies. After all, the promise of AI in healthcare can only be achieved if there are healthcare workers capable of building, testing, and implementing AI solutions.

B. The Problem: Siloed Development Across the AI Lifecycle

Healthcare systems are complex environments with many moving parts. The design, development, validation, deployment, and monitoring of AI in healthcare typically require different stakeholders and priorities for each of these phases. While AI scientists prioritize algorithms and technical performance, clinicians emphasize patient safety and workflow (Gisselbaek et al. 2025). Implementation teams may prioritize feasibility and resource constraints, patients may emphasize acceptability and equity, while program managers focus on scalability and sustainability (Jacob et al. 2025). These stakeholders are often siloed in their expertise, with limited engagement with each other across the AI lifecycle.

This fragmentation is problematic in resource-constrained settings like the Philippines, where the gap between AI promise and implementation reality is widest. Infrastructure limitations, regulatory uncertainty, variable facility

capacity, and competing priorities mean that AI tools designed in isolation will predictably fail to translate into equitable and sustainable impact.

C. The Gap

Current AI frameworks were designed for high-income countries and do not account for the specific contextual challenges facing low- and middle-income countries. These challenges include archipelagic infrastructure limitations, regulatory uncertainty, variable facility capacity with limited health workforce, and the lack of culturally appropriate solutions. Interdisciplinary governance mechanisms for AI in healthcare are needed to avoid widening health inequities and harming patients.

D. Objectives

This discussion paper reviews existing interdisciplinary frameworks for AI in healthcare and illustrates their limitations in low- and middle-income country contexts using a purposively selected Philippine case study. An interdisciplinary AI framework localized to the Philippines will be proposed.

Methodology

A literature review of interdisciplinary frameworks for AI in healthcare was conducted using PubMed, Google Scholar, and Research Rabbit. Search terms included: artificial intelligence, healthcare tools, medical tools, interdisciplinary framework, and collaboration. Research Rabbit was used to visualize citation networks and identify additional relevant frameworks through reference mapping. Frameworks were included if they addressed the AI lifecycle, interdisciplinary collaboration, or both. Frameworks were synthesized across two dimensions: lifecycle coverage and mechanisms for interdisciplinary engagement, including patient inclusion.

To ground the framework analysis in the Philippine context, a case study was purposively selected from published peer-reviewed literature. The AI-CAD tuberculosis screening implementation was selected as the only documented large-scale AI deployment in Philippine healthcare with sufficient published evidence to analyze across multiple lifecycle stages. Published accounts of the implementation were reviewed to identify where cross-cutting governance requirements were absent and what gaps resulted.

Interdisciplinary Frameworks for AI in Healthcare

The literature search revealed interdisciplinary AI frameworks that ranged in coverage from broad international principles to specific country or institutional guidelines. A synthesis of frameworks is necessary to navigate the multidimensional and interdisciplinary challenges of healthcare AI implementation globally, and particularly in resource-constrained settings like the Philippines. Table 1 maps elements of the framework or guideline according to phases of the AI lifecycle: design, development, validation, deployment, and monitoring. The frameworks were further analyzed to determine if they addressed interdisciplinarity and patient inclusion.

Table 1. AI Frameworks Addressing the AI Lifecycle

Framework / Guideline	AI Life Cycle				
	Design	**Development**	**Validation**	**Deployment**	**Monitoring**
ASEAN[a] Guide on AI Governance (ASEAN 2024)	Emphasis on alignment with ethical policies & to conduct risk assessments; human-centric design; privacy by design	Guidance on data collection, processing, and modelling, including how to train and check for biases Adversarial testing to prepare against unexpected input	Stakeholders' involvement in performance acceptance tests so their outcomes may align with the purpose of the AI (Fit for purpose) Separate validation dataset recommended	Establish training, tests and risk alignment General disclosure provided to users	Conduct risk assessments to flag errors over time, with avenues for users to request human review Metrics defined by the organization
AICC[b] (NAM[c] 2025)	Identify goals, stakeholders and data needs; incorporate society and cultural goals into technical implementation	Acquire large, high-quality datasets with consideration of privacy and representation to avoid bias	Frequent validation using statistical performance; consider clinical/ economic outcomes	Focus on user-centered clinical workflow so that the AI tool improves the well-being of the workforce	Continuous "algorithmovigilance" Post-implementation monitoring is essential to detect problems, including unsafe systems, and plan for decommissioning
FUTURE-AI Framework (Lekadir et al. 2023)	Engagement with stakeholders to focus on a human-centered, risk-aware strategy	Prioritize data representation, standardization (such as DICOMe/FHIR[f]), and mitigation strategies	External validation for generalizability across sites	Focus on regulatory compliance, providing training, and recommending oversight committees "AI information leaflets" for transparency	Establish a logging system for accountability and auditing Check the AI tool for performance degradation
Singapore AIHGle[g] (MOH[h] & HSA[i] 2026)	Include clinical inputs relevant to the use of AI from individuals with relevant expertise	Factor in ethical considerations; use high-quality and representative datasets Developers register with the government	Periodically evaluate and validate AI to ensure it meets clinical practice standards Different training and test data sets	Establish triggers and escalation pathways Train staff with contingency protocols	Periodic evaluation for model drift, develop strict exit management once AI tools will no longer be used.
OECD[j] AI Lifecycle (OECD 2024)	Design AI in a way that respects the rule of law, human rights, democratic values, and diversity, including safeguards	Ensure traceability of datasets and processes during development	An independent government body will ensure the standards of the AI tool	Implement training programs to upskill users	Establish reporting for positive and negative incidents
Stanford FURM[k] Assessment (Callahan et al. 2024)	Identify problems to target, clinical use, ethical considerations, and financial impact APLUS[l] is a Stanford simulation tool program for checking usefulness.	Check across past historical data, training methods, to select appropriate algorithms for specific use-cases	Conduct simulations; verify output before releasing to the public "Silent deployment" before full integration	Organizational integration to identify leaders and make a rollout plan "Loud deployment:" workflow actions are fully executed Randomized departments to check outcomes and personnel reaction.	Check impact, adherence, and outcomes Continuous experiments on impact even after go-live. Prepare plans for retraining and retirement of the AI tool
WHO Ethics & Governance of AI for Health (WHO 2021)	Design for values and incorporate dignity and autonomy as requirements	"Pre-mortem" and "red team" testing to identify vulnerabilities before they happen ISO[m] or IEEE[n] standards for privacy and security	Prospective randomized trials are needed, and not just training data Include diverse demographics	Human supervision on AI output Health care professionals are educated on AI limitations "Social license:" consent & trust of the community involved	Continued risk-based quality monitoring and post-market assessment Assessment by independent third parties and published
AI for IMPACTS[o] Framework (Jacob et al. 2025)	Consider "user-centricity," fairness, and interoperability when designing AI systems	Focus on environmental sustainability when developing AI systems	Measure performance metrics depending on the tool	Consider the environmental impact of deployment, including whether it will require additional technology (hardware or software)	Post-deployment monitoring, including checking for model drift, data security Maintain liability checks

RUAIH[p] (Joint Commission and CHAI[q] 2025)	Focused on vendor transparency and privacy design	Acknowledges the bias that can occur in development, but leaves the process to the vendor only	Validation is mentioned, but focused on the vendor's responsibility to check against local data	Organizations are to provide training sessions to help with integration into the clinical workflow	Continuous monitoring to detect performance drift, reevaluation of when to terminate, and blind safety reporting
AMA[r] STEPS Governance Toolkit (Lozovatsky and Thomas 2025)	Governance group formation to set organizational goals Establish guardrails for ethical standards, bias mitigation, and potential problems	Consider off-the-shelf or internal development, depending on intention and resources	Vendor evaluation using peer-reviewed studies and internal pilot results Recommend diverse demographics and targeted testing	Go-live planning, change management, and a workshop for organizational readiness across departments	Track for hallucinations, data drift (performance degradation), and concept drift (changes in practice) Ensure channels for reporting and turning off AI tools if needed
NICE[s] Evidence Standards Framework (Unsworth et al. 2021)	Clinical & social care professionals involved in design	Pilot strategies can be done safely while full evidence is being gathered	Comparative studies, randomized Clinical Trials, and real-world performances for validation	Staff trained depending on the AI tool's purpose (for systemic or individual use)	Adaptive AI algorithms and post-deployment analysis Economic impact to check sustainability
CHUM[t] Framework (CHUM 2024)	Co-design workshops for AI Experts, healthcare, and education	Field experience is incorporated in development as continuous, not just a one-time consult.	Validation is mentioned in relation to the readiness of human users or organizations to understand AI ethically	Human-centered deployment: focus on the competency of users Multisectoral collaboration	Continuous monitoring that is human-centric (how it affects patients) rather than focused on technical outcomes

a. ASEAN: Association of Southeast Asian Nations

b. AICC: AI Code of Conduct

c. NAM: National Academy of Medicine

d. FUTURE-AI: Fairness, Universality, Traceability, Usability, Robustness, and Explainability - AI

e. DICOM: Digital Imaging and Communications in Medicine

f. FHIR: Fast Healthcare Interoperability Resources

g. AIHGle: Artificial Intelligence in Healthcare Guidelines

h. MOH: Ministry of Health

i. HSA: Health Sciences Authority

j. OECD: Organisation for Economic Co-operation and Development

k. FURM: Fair, Useful, and Reliable AI Model

l. APLUS: Achievable Performance via Utility-Based Simulation

m. ISO: International Organization for Standardization

n. IEEE: Institute of Electrical and Electronics Engineers

o. IMPACTS: (1) I—integration, interoperability, and workflow; (2) M—monitoring, governance, and accountability; (3) P—performance and quality metrics; (4) A—acceptability, trust, and training; (5) C—cost and economic evaluation; (6) T—technological safety and transparency; and (7) S—scalability and impact

p. RUAIH: Responsible Use of AI in Healthcare

q. CHAI: Coalition for Health AI

r. AMA: American Medical Association

s. NICE: National Institute for Health and Care Excellence

CHUM: Centre hospitalier de l'Université de Montréal

The reviewed frameworks reveal several consistent findings. First, all frameworks address AI governance across the full lifecycle rather than as a one-time check. Emphasis is placed on ethical policy alignment, stakeholder needs, and risk assessments at the design stage to anticipate failure before it occurs. In development, many frameworks recommend large, high-quality datasets and bias checks to ensure fairness. Continuous oversight is recommended for monitoring. Second, the interdisciplinary team envisioned by these frameworks goes beyond consultation and beyond clinicians and engineers. Technical performance alone is insufficient, as teams must include competencies in ethics, legal aspects, human resources, etc.

Design

Across almost all frameworks, the design stage emphasizes moving beyond technical requirements to include ethical and social safeguards. Frameworks like AICC (NAM 2025), FUTURE-AI (Lekadir et al. 2023), and CHUM (2024) prioritize involving clinicians, patients, and multi-sectoral experts from the start. WHO and OECD (OECD, 2024) stress that AI must respect human rights, dignity, and autonomy as foundational requirements. The ASEAN (ASEAN, 2024) guide and AMA STEPS (Lozovatsky and Thomas 2025) recommend identifying potential biases and harms before data collection, processing, and modelling.

Development

The reviewed frameworks focus on data integrity and bias mitigation. AIHGle (MOH and HSA 2026) and AICC (NAM 2025) emphasize using representative, large-scale datasets to ensure the AI works for diverse populations. FUTURE-AI (Lekadir et al. 2023) specifically mentions technical standards like DICOM and FHIR to ensure interoperability across different healthcare systems. RUAIH (Joint Commission and CHAI 2025) and OECD (2024) demand traceability, ensuring that the data collection and modeling processes are documented and auditable.

Validation

Rigorous testing beyond accuracy is recommended across the frameworks. Frameworks advocate for diverse testing methods such as adversarial testing (ASEAN 2024), randomized clinical trials (WHO 2021; Unsworth et al. 2021), and simulations (Callahan et al. 2024). Adversarial testing is described

as a "series of tests to expose the system to a broad range of unexpected inputs and mitigate any unintended behaviour" before deployment. FUTURE-AI (Lekadir et al. 2023) and AMA STEPS (Lozovatsky and Thomas 2025) stress "external validation," testing the AI on data from different locations to ensure it wasn't just "overfit" to its original training site.

Deployment

The reviewed frameworks recognize that the transition of AI from a development environment to a live clinical setting requires a socio-technical approach centered on workforce readiness, phased integration, and clear communication. Nearly all frameworks (OECD, 2024; Joint Commission and CHAI 2025; CHUM 2024) highlight that staff must be trained not just on how to use the tool, but on its limitations and contingency protocols. The ASEAN Guide (ASEAN, 2024) and AMA STEPS (Lozovatsky and Thomas, 2025) recommend that organizations treat deployment as an exercise in change management, encouraging leaders to redesign existing jobs so that AI augments rather than replaces human skillsets. FURM (Callahan et al., 2024) suggests a "silent deployment" (running in the background) before a "loud deployment" (full integration). FUTURE-AI (Lekadir et al. 2023) suggests "AI information leaflets" to help users understand the tool's intended use, instructions, and risks, such as potential biases.

Monitoring

The AI lifecycle does not end at "go-live." AMA STEPS (Lozovatsky and Thomas, 2025) and AIHGle (MOH and HSA 2026) focus on "model drift" or "performance degradation," where an AI's accuracy drops as real-world data changes over time. AICC (NAM 2025) and WHO (2021) advocate for continuous, post-market assessment and "blind safety reporting" to catch errors. Critically, AICC (NAM 2025), AIHGle (MOH and HSA 2026), and FURM (Callahan et al. 2024) include plans for decommissioning, knowing exactly when and how to turn an AI tool off if it is no longer retrainable, or has become unsafe or obsolete.

While all frameworks reject siloed AI development, they differ significantly in both the composition and continuity of interdisciplinary engagement across the lifecycle. On composition, there is broad consensus that a core team must bridge the technical and clinical worlds through diverse expertise spanning ethics, law, medicine, IT, finance, and data science. Most frameworks formalize

this through designated structures. Stanford's FURM (Callahan et al. 2024) uses "component leads" such as clinicians, ethicists, and IT professionals. AMA STEPS (Lozovatsky and Thomas 2025) forms interdisciplinary working groups, with representatives from legal, finance, and pharmacy departments. RUAIH (Joint Commission and CHAI 2025) specifies a team including experts in cybersecurity, regulatory compliance, and clinical operations. The CHUM framework (CHUM 2024) extends this further by embedding cross-sectoral perspectives from education and AI research into the team's competency design from the outset.

The frameworks differ significantly in how they include patients. There is a clear spectrum of involvement from patients as core leaders/partners, as consultative stakeholders, or as "end-users" for acceptance. The AICC (NAM 2025) is the most progressive in calling for "patient-led representation" at every single stage of the AI lifecycle. RUAIH (Joint Commission and CHAI 2025) specifically includes patients and caregivers as part of the "AI team" to reflect the needs of impacted populations. WHO (2021) uses the concept of a "human warranty," implying that patients must evaluate the technology before it is considered responsible. FURM (Callahan et al. 2024) relegates patients to a consultative role where ethicists conduct interviews to identify potential "value collisions." AIHGle (MOH and HSA 2026) suggests developers seek patient input to clarify wording and ensure the AI design is holistic, but patients are not listed as "core" developers. FUTURE-AI (Lekadir et al. 2023) recommends "advisory boards" or interviews to define requirements early on. Rather than involving patients in the build, AI for IMPACTS (Jacob et al. 2025) focuses on "acceptance," measuring how well a tool is embraced by patients. AMA STEPS (Lozovatsky and Thomas 2025) encourages patients to join the monitoring team after the tool is live to report problems, rather than designing the tool itself.

The more consequential pattern cuts across both dimensions. Regardless of where frameworks fall on the patient inclusion spectrum, concrete mechanisms for interdisciplinary engagement, whether of patients or the broader team, are most specified at design and progressively underspecified at validation, deployment, and monitoring. While frameworks assign responsibility to teams in these stages, the mechanisms for interdisciplinary collaboration are less granular. For validation, while frameworks specify technical metrics (e.g., AUROC, F1 scores, or 95 percent confidence intervals), they often lack instructions on how interdisciplinary teams can reconcile conflicting technical and clinical goals. For deployment, frameworks mandate "organizational integration" but often fail to define the real-time decision-making hierarchy

during the transition to a live environment. FURM (Callahan et al. 2024) suggests a "loud deployment" where personnel take action on AI outputs, but it does not specify who has the final authority to immediately "shut down" the system if a patient-safety event occurs, only that "relevant personnel" should be able to assume control. The ASEAN Guide (ASEAN 2024) mentions that "available capacity" should be reconciled, but they do not provide a workflow for resolving disputes when an AI tool increases the workload for one department (e.g., more nurse-led screenings) while benefiting another (e.g., fewer specialist referrals). For monitoring, FUTURE-AI (Lekadir et al. 2023) advises to "assign the role of auditor(s)" and update the model, but it does not specify the interdisciplinary workflow for how the team collectively decides on a "performance threshold" or a "decommissioning trigger" as concretely as it specifies the design-phase workshop models.

The Philippine Case Study: AI-CAD Tuberculosis Screening Implementation

Tuberculosis (TB) remains a public health crisis in the Philippines. The Philippines recorded 739,000 new TB cases in 2023 with an incidence rate of 643 per 100,000 people. In response, the Department of Health promoted a strategy to enhance case-finding efforts by utilizing innovative and more sensitive screening tools to detect TB presumptively, leading to the introduction of AI technologies in chest X-ray screening across communities, health facilities, and workplaces (DOH 2025).

The Philippines has approximately 1,500 radiologists serving a population of over 100 million, below the ideal of one per 10,000 (Flores et al. 2024). These are the same radiologists expected to function as "humans-in-the-loop" by checking AI outputs in imaging-based screening. Interoperability across health facilities remains elusive despite legislative mandates, limiting the feedback infrastructure necessary to detect model drift or implementation failure early. The need to address the gaps in interdisciplinarity and patient inclusion revealed in the framework review is even more acute for the Philippines, where resources are limited and community trust in health institutions cannot be taken for granted.

The published AI-CAD TB screening implementation study by Marquez et al. (2025) offers the only documented large-scale AI deployment in Philippine healthcare to examine the AI lifecycle, interdisciplinarity, and patient

inclusion. It is, however, a retrospective cross-sectional clinical study designed to evaluate diagnostic performance. The study was not designed to document interdisciplinary processes, stakeholder deliberation, threshold-setting rationales, community engagement, or patient experience. Gaps identified in the analyses below should therefore be understood as unanswered governance questions, rather than definitive evidence that the corresponding practices were absent. The analyses evaluate governance structures and framework adequacy, not of the implementation team's conduct or intentions.

The Marquez et al study (2025) is a retrospective cross-sectional study of individuals 15 years and older participating in systematic TB case-finding activities in four regions of the Philippines (Regions 3, 4A, 7, and the National Capital Region) between May 2021 and March 2024. It included 5,740 individuals who had complete chest X-rays and molecular World Health Organization-recommended rapid diagnostic (mWRD) test results. The study compared the performance of chest X-ray with computer-aided detection powered by deep learning-based artificial intelligence (AI-CAD) for the identification of presumptive TB cases against mWRD tests (such as GeneXpert) as the gold standard for microbiological confirmation. The tool used was qXR version 3 (developed by Qure.ai), which provided a quantitative abnormality score between 0 and 1. The primary implementation utilized a fixed abnormality threshold of 0.50 to classify individuals as "TB presumptive."

A total of 52,840 individuals were screened, of which 47,817 underwent chest X-ray. The 40,486 individuals who were negative for PTB in both the AI-CAD did not undergo further testing with mWRD. In the study's real-world workflow, only individuals identified as "TB presumptive" (e.g., those who were CAD-positive, had symptoms, or had abnormal radiologist readings) were referred for the "gold standard" confirmatory test (mWRD). The primary outcomes measured were sensitivity, specificity, and positive predictive value.

At the 0.5 threshold, the AI-CAD had a sensitivity of 95.6 percent, specificity of 28.1 percent, and positive predictive value (PPV) of 16.6 percent. The high sensitivity meant that the tool was highly effective in identifying individuals with TB, though with the low specificity, 83 percent of those flagged by AI-CAD did not have TB. A major reason for this is that the AI-CAD often flags scars from previous TB infections as active disease. The low PPV meant that for every one person diagnosed with TB, the program had to perform expensive mWRD tests on five other people who ultimately tested negative, leading to a potential overutilization of diagnostic resources.

The authors used the terms "pseudo-sensitivity" and "pseudo-specificity" to denote that the diagnostic performance metrics were calculated using an incomplete or biased sample resulting from the programmatic screening algorithm. Because those who were CAD-negative were excluded from further testing, the data suffered from partial verification. When a study confirms more CAD-positive than CAD-negative cases, it tends to lead to an overestimation of sensitivity and an underestimation of specificity.

Analysis of AI-CAD by AI Lifecycle Stages

DESIGN. The AI-CAD used was qXR, which was trained on globally sourced data (Qure.AI 2020). The abnormality threshold was set at 0.50 based on a 2020 pilot implementation to "align with the national context." However, frameworks like AICC (NAM 2025) and FUTURE-AI (Lekadir et al. 2023) emphasize that the design stage must involve a deep understanding of clinical workflows and stakeholder needs to define precise problems. The study retrospectively found that the 0.5 threshold resulted in low pseudo-specificity (28.1 percent), leading to a high volume of unnecessary confirmatory tests. A rigorous design stage, as suggested by FURM (Callahan et al. 2024), would have used workflow simulations to generate quantitative estimates of "achievable utility" and resource constraints before deployment.

DEVELOPMENT AND VALIDATION. Marquez et al. (2025) found that pseudo-specificity was not uniform across the population and was significantly affected by demographic and clinical factors. For example, individuals with a history of TB treatment had a significantly lower pseudo-specificity of 9.3 percent, compared to 19.2 percent for those without previous treatment. Pseudo-specificity was much higher among individuals with symptoms (66.2 percent) than those without symptoms (13.3 percent). Pseudo-specificity decreased with age, peaking at 49.4 percent for the youngest group (15–23) and dropping to 25.3 percent for those 50 and older. Females showed a higher pseudo-specificity (34.4 percent) than males (22.4 percent). Pseudo-specificity was highest in community settings (35.2 percent) and lowest in workplaces (14.0 percent). These results demonstrate that during the development stage, there may have been a need to fine-tune the model with local demographic data to account for regional health profiles (like higher rates of healed lung scars in older Filipinos) before widespread use. Marquez et al. (2025) suggested a threshold value of 0.68, resulting in a tradeoff between mWRD tests saved and missed cases. With this higher threshold value, 42 percent mWRD tests would have been saved but missing 74 cases vs 33 missed cases with the 0.5

threshold value. Threshold adjustments affect not just case detection but also financial costs.

DEPLOYMENT. Mobile vans equipped with X-ray machines and AI-CAD, along with trained radiographers and other health workers, were sent to different cities in the country to screen presumptive TB cases and high-risk individuals across cities (Modi and Suresh 2019). GeneXpert cartridges were prioritized for those flagged as "presumptive" by the triage algorithm. The low specificity of the AI-CAD led to the use of GeneXpert cartridges on individuals with an 83 percent false-positive rate. Frameworks like the ASEAN (2024) and AICC (NAM 2025) mandate that "available capacity" be reconciled with anticipated patient volume. In resource-limited settings, the absence of documented capacity reconciliation represents a gap against what these frameworks require, where technical performance (high sensitivity) may overwhelm diagnostic infrastructure without structural safeguards. FURM (Callahan et al. 2024) recommends a "silent deployment" (running in the background without clinical action) to verify if model outputs match expectations.

MONITORING. Although the AI-CAD's performance metrics showed significant annual variation, the abnormality threshold remained static at 0.5 from 2021 to 2024. The study noted a general "downward trend" in detected TB yields over the period; hence, the authors underscore the need for context-specific calibration and "near-continuous" monitoring. Marquez et al. (2025) do not report the use of real-time monitoring mechanisms such as the "AI-QI (Artificial Intelligence Quality Improvement)" units or dashboards recommended by AICC (NAM 2025) and AIGle, respectively, to trigger recalibration when PPV falls below acceptable levels. The recommendation to "retrain" the model was made after the retrospective analysis was complete.

Analysis of AI-CAD by Interdisciplinarity and Patient Inclusion

Frameworks such as the ASEAN Guide (ASEAN 2024), AMA STEPS (Lozovatsky and Thomas 2025), and RUAIH (Joint Commission and CHAI 2025) emphasize that AI should be governed by a "multi-disciplinary, central governing body" or "interdisciplinary AI working group" including representatives from medicine, finance, IT, data science, and legal departments. In Marquez et al.'s (2025) study, the abnormality threshold remained static at 0.50 from 2021 to 2024 based on a 2020 pilot, even as yearly performance

varied significantly and PPV eventually stagnated. Whether interdisciplinary oversight mechanisms were in place to detect and act on these variations is not reported. Against the standard set by frameworks such as FURM (Callahan et al. 2024) and AICC (NAM 2025), which require close, ongoing collaboration between clinicians, administrators, and model developers to ensure context-sensitive calibration, the documented absence of threshold adjustment over this period represents an unmet governance expectation, regardless of its cause.

Applying the lens of patient inclusion to the Marquez et al. study (2025) is difficult, given that it is a diagnostic performance study. As such, it does not report on patient roles in co-creation, threshold deliberation, consent beyond standard TB screening procedures, or patient-reported outcomes.

The Interdisciplinary Governance Framework for AI in Philippine Healthcare

The AI-CAD case study illustrates a pattern that is structurally predictable rather than incidental. Analyzed against the governance standards of the reviewed frameworks, the following were either unmet or undocumented across the lifecycle: multi-stakeholder deliberation over threshold calibration; interrogation of training data for local representativeness; culturally appropriate communication of uncertainty to communities and health workers; and structured community roles in validation and monitoring. We propose the following as cross-cutting governance requirements: localization adequacy, relational trust, and accountable decision authority. These principles must be revisited at every lifecycle stage, not resolved once and carried forward. Together with a structured set of stakeholder questions organized across the AI lifecycle, they form the Interdisciplinary AI Governance Framework for Philippine Healthcare, a dual-layer framework localized to the constraints and cultural context of the Philippine health system.

The proposed framework incorporates concepts of *kapwa* and *pakikiramdam* as fundamental cultural anchors within relational trust. *Kapwa* (shared Filipino identity and relationality) serves as the core value of the framework, where patients and their communities are treated not as units of health data (Gozum 2025) but as fellow (*kapwa*) humans with inherent dignity. As the violation of *kapwa* is a relational transgression for Filipinos (Yacat 2017), it is introduced early in the design stage to maintain trust necessary for AI adoption. We argue

that *kapwa* means that the patient is not separable from his/her community. Thus, as the framework's core value, *kapwa* requires that the patient and community voices be present in the deliberation when resolving conflicts. Complementing this is *pakikiramdam*, or empathetic attunement in clinical care (Gozum 2025). *Pakikiramdam,* like *kapwa*, must be evident in all stages of the AI lifecycle, but critical in the monitoring stage to assess community feedback. This ensures that governance remains sensitive to unspoken concerns that technical metrics alone cannot detect. Together, these values provide a human warranty that honors the communal and relational nature of Filipino healthcare.

To optimize the use of the framework, it should be integrated into the institutional process, specifically through an interdisciplinary working group. At each critical transition in the lifecycle, the group must ask these questions to clarify decision authorities, processes, and responsibilities.

Table 2. Interdisciplinary Governance Framework for AI in Philippine Healthcare

AI Governance Principle	AI Life Cycle	
	Life Cycle Stage	**Governance Questions**
Localization Adequacy: Ensuring that the AI system will be applicable to the specific clinical situation or geographical setting. *Does this AI tool belong here?*	**Design**	What institutional process determines whether this AI tool aligns with organizational policy, and who has final authority to approve or reject adoption? How was the health need this tool addresses defined, and which community or local health system representatives were part of that process?
	Development	Who selects the technical standards, and who ensures data interoperability? What process defines equity and sustainability for the local community?
	Validation	Who certifies that the training data represents the local community, and which local stakeholders review these results? What is the process to check the tool across health literacy levels?
	Deployment	Who has the final approval for deployment? Which institutional leaders will confirm the readiness of the local infrastructure or resources? What are the processes to manage resources?
	Monitoring	Who is responsible for health disparity tracking & who reviews this data? What is the process to re-evaluate the local value of the tool?

Relational Trust: Fostering a community that accepts the tool and trusts it enough to believe the findings. *Does the community accept and trust the AI tool?*	**Design**	Who conducts the ethical risk assessment, and which interdisciplinary committee reviews it? What process ensures that community perspectives and the shared identity of *kapwa* are integrated in the tool?
	Development	Who is responsible for ensuring patient autonomy? What is the process for an "opt-out" mechanism?
	Validation	Who is responsible for training clinical staff? What processes will ensure that clinical staff can explain AI outputs to patients? What is the process for periodic re-evaluation of clinical staff?
	Deployment	Who manages the deployment of the AI tool? What is the process to address feedback from users?
	Monitoring	Which user group provides feedback? How is *pakikiramdam* used to assess feedback? Who has the final decision on modifications?
Accountable Decision Authority Defining who decides what is good enough for an AI tool. *Who decides what good enough looks like for this AI tool?*	**Design**	Who decides on the specific performance metrics of the tool? What is the process to resolve interdepartmental differences in metric priority?
	Development	Who is responsible for simulation parameters, who reviews these, and what is the process for reviewing the results of the simulation?
	Validation	Who leads the oversight committee, and which committee investigates these errors? What is the process for when errors occur?
	Deployment	Who is in charge of escalations, who can override AI decisions, and what process resolves disagreements with AI?
	Monitoring	Who is responsible for auditing and deciding on updates or decommissioning? What processes will trigger a mandatory review?

AI Use Disclosure

AI-assisted tools were used during the preparation of this work. Research Rabbit and NotebookLM (Google) supported the synthesis of existing literature. Socratic dialogue with Claude Sonnet 4.6 (Anthropic) was used to stress-test the paper's argument and structure. The authors declare that all AI-mediated outputs were subjected to strict human oversight and verification. The authors take full responsibility for this paper.

References

Association of Southeast Asian Nations. 2024. *ASEAN Guide to AI Governance and Ethics*. Association of Southeast Asian Nations. https://asean.org/wp-content/uploads/2024/02/ASEAN-Guide-on-AI-Governance-and-Ethics_beautified_201223_v2.pdf.

Callahan, Alison, Duncan McElfresh, Juan M. Banda, et al. 2024. "Standing on FURM Ground: A Framework for Evaluating Fair, Useful, and Reliable AI Models in Health Care Systems. *NEJM Catalyst Innovations in Care Delivery* 5 (10). https://doi.org/10.1056/cat.24.0131.

Centre hospitalier de l'Université de Montréal. 2024. "Competency Framework for AI Integration in Healthcare." School of Artificial Intelligence in Healthcare of the Centre hospitalier de l'Université de Montréal. https://eiaschum.ca/en/competency-framework-eias/.

Ciecierski-Holmes, Tadeusz, Ritvij Singh, Miriam Axt, Stephan Brenner, and Sandra Barteit. 2022. "Artificial Intelligence for Strengthening Healthcare Systems in Low- and Middle-Income Countries: A Systematic Scoping Review." *NPJ Digital Medicine* 5 (162). https://doi.org/10.1038/s41746-022-00700-y.

Department of Health (Philippines). 2025. Guidelines on the Use of Ultraportable Chest X-ray (UP-CXR) in Tuberculosis (TB) Screening under the National Tuberculosis Control Program (NTP). Dept. Cir. No. 2025-0048 (14 January 2025). https://ntp.doh.gov.ph/download/dc2025-0048/.

Elepaño, Anton, Carol Stephanie Tan-Lim, Mark Anthony Javelosa, et al. 2025. "Implementing Electronic Health Records in Philippine Primary Care Settings: Mixed-Methods Pilot Study." *JMIR Medical Informatics* 13 (e63036). https://doi.org/10.2196/63036.

Food and Drugs Administration (Philippines). 2025, July 20). Guidelines on the Regulation of Medical Device Software (MDSW) by the Food and Drug Administration (FDA). Draft for comments, 20 July 2025. https://www.fda.gov.ph/draft-for-comments-guidelines-on-the-regulation-of-medical-device-software-mdsw-by-the-food-and-drug-administration-fda/.

Flores, Gabrielle P., Reiner Lorenzo J. Tamayo, Robert Neil F. Leong, et al. 2024. "Diagnostic Performance of a Computer-aided System for Tuberculosis Screening in Two Philippine Cities." *Acta Medica Philippina* 59 (2): 33–40. https://doi.org/10.47895/amp.vi0.8950.

Gisselbaek, Mia, Joana Berger-Estilita, Arnout Devos, Pierre Luigi Ingrassia, Peter Dieckmann, and Sarah Saxena. 2025. "Bridging the Gap Between Scientists and Clinicians: Addressing Collaboration Challenges in Clinical AI Integration." *BMC Anesthesiology* 25 (269). https://doi.org/10.1186/s12871-025-03130-x.

Gozum, Ivan Efreaim A. 2025. "When AI Meets *Pakikiramdam*: Humanizing Digital Health in Philippine Family Practice." *Journal of General and Family Medicine* 26 (6): 663–64. https://doi.org/10.1002/jgf2.70078.

Henry, Tanya Albert. 2025. "Physicians' Greatest Use for AI? Cutting Administrative Burdens." American Medical Association, 20 March 2025. https://www.ama-assn.org/practice-management/digital-health/physicians-greatest-use-ai-cutting-administrative-burdens.

Jacob, Christine, Noé Brasier, Emanuele Laurenzi, et al. (2025). "AI for IMPACTS Framework for Evaluating the Long-Term Real-World Impacts of AI-Powered Clinician Tools: Systematic Review and Narrative Synthesis." *Journal of Medical Internet Research* 27 (e67485). https://doi.org/10.2196/67485.

Joint Commission and Coalition for Health AI. 2025. *The Responsible Use of AI in Healthcare (RUAIH)*. Joint Commission. https://digitalassets.jointcommission.org/api/public/content/dcfcf4f1a0cc45cdb526b3cb034c68c2.

Lekadir, Karim, Aasa Feragen, Abdul Joseph Fofanah, A. J., et al. (2023, August 11). *FUTURE-AI: International Consensus Guideline for Trustworthy and Deployable Artificial Intelligence in Healthcare*. ArXiv.org, 11 August 2023. https://doi.org/10.48550/arXiv.2309.12325.

Lomis, Kimberly, Pamela Jeffries, Anthony Palatta, et al. 2021. "Artificial Intelligence for Health Professions Educators." *NAM Perspectives*, 8 September 2021. https://doi.org/10.31478/202109a.

Lozovatsky, Margaret, and Michele Thomas. 2025. "Governance for Augmented Intelligence: Establish a Governance Framework to Implement, Manage, and Scale AI Solutions." American Medical Association STEPS Forward, AMA Ed Hub, 29 April 2025. https://edhub.ama-assn.org/steps-forward/module/2833560.

Marfal, Allan Mauro V. 2025. "DOST Builds on 'AI' National Strategy." Department of Science and Technology (Philippines), 2 May 2025. https://www.stii.dost.gov.ph/2025-dost-builds-on-ai-national-strategy.

Marquez Nichel, Erwin John Carpio, Mary Rosary Santiago, et al. 2025. "Performance of Chest X-ray with Computer-Aided Detection Powered by Deep Learning-Based Artificial Intelligence for Tuberculosis Presumptive Identification During Case Finding in the Philippines." *BMC Global and Public Health* 3 (74). https://doi.org/10.1186/s44263-025-00198-y.

Mennella, Ciro, Umberto Maniscalco, Giuseppe De Pietro, and Massimo Esposito. 2024. "Ethical and Regulatory Challenges of AI Technologies in Healthcare: A Narrative Review." *Heliyon* 10 (4): e26297. https://doi.org/10.1016/j.heliyon.2024.e26297.

Ministry of Health (Singapore) and Health Sciences Authority. 2026. *Artificial Intelligence in Healthcare Guidelines*. Singapore Ministry of Health and Health Sciences Authority. https://isomer-user-content.by.gov.sg/3/696493a4-446c-4f91-844c-475b6a570e23/Updated_AIHGle%202.0.pdf.

Modi, Ankit, and Reshma Suresh. 2019. "Scaling Up TB Screening with AI: Deploying Automated X-ray Screening in Remote Regions." Qure.ai, 15 April 2019. https://www.qure.ai/blog/scaling-up-tb-screening-with-ai-deploying-automated-x-ray-screening-in-remote-regions.

Murai, Shinsuke, Ray Justin C. Ventura, and Julita T. Gaite. 2022. "Timeliness of Reporting Process in the National Routine Health Information System: The Case of 19-Year Experience of Field Health Services Information System in Palawan, the Philippines." *PLOS One* 17 (2): e0264681. https://doi.org/10.1371/journal.pone.0264681.

National Academy of Medicine. 2026. "Health Care Artificial Intelligence Code of Conduct." https://nam.edu/our-work/programs/leadership-consortium/health-care-artificial-intelligence-code-of-conduct/.

Ocana, Alberto, Atanasio Pandiella, Cristian Privat, et al. 2025. "Integrating Artificial Intelligence in Drug Discovery and Early Drug Development: A Transformative Approach." *Biomarker Research* 13 (45). https://doi.org/10.1186/s40364-025-00758-2.

Olawade, D. B., David-Olawade, A. C., Wada, O. Z., Asaolu, A. J., Adereni, T., & Ling, J. (2024). Artificial Intelligence in Healthcare Delivery: Prospects and Pitfalls. Journal of Medicine, Surgery, and Public Health, 3(100108), 100108. https://doi.org/10.1016/j.glmedi.2024.100108

Organisation for Economic Co-operation and Development. 2024. "OECD AI Principles." Organisation for Economic Co-operation and Development. https://www.oecd.org/en/topics/ai-principles.html.

Pepito, Veincent Christian F., Arianna Maever Loreche, Ruth Shane Legaspi, et al. 2025. "Health Workforce Issues and Recommended Practices in the Implementation of Universal Health Coverage in the Philippines: A Qualitative Study." *Human Resources for Health* 23 (21). https://doi.org/10.1186/s12960-025-00988-3

Qure.AI. 2020. "Virtual Innovation Spotlight: AI for Screening, Contact Tracing & Triaging for TB & COVID-19, QURE.AI Questions & Answers Summary." Stop TB Partnership, 8 April 2020. https://www.stoptb.org/sites/default/files/imported/document/20200408_VIS_AI_Qure.ai_QASummary_Final.pdf.

Universal Healthcare Act. 2019. RA No. 11223 (20 February 2019). https://lawphil.net/statutes/repacts/ra2019/ra_11223_2019.html.

Unsworth, Harriet, Bernice Dillon, Lucie Collinson, et al. 2021. "The NICE Evidence Standards Framework for Digital Health and Care Technologies—Developing and Maintaining an Innovative Evidence Framework with Global Impact. *Digital Health* 7. https://doi.org/10.1177/20552076211018617.

World Health Organization. 2021. *Ethics and Governance of Artificial Intelligence for Health*. World Health Organization, 28 June 2021. https://www.who.int/publications/i/item/9789240029200.

Yacat, Jay A. 2017. "Walang Pakisama o Walang Kapwa-Tao?: Isang Sikolohikal na Pagsusuri sa Tindi ng Paglabag at Ugnayan sa Relasyong Panlipunan." *DIWA E-Journal* 5: 1–22. https://www.pssp.org.ph/diwa/wp-content/uploads/2017/11/1-Artikulo-Yacat.pdf.

THE UP CIDS DISCUSSION PAPER SERIES

The UP CIDS Discussion Paper Series features preliminary researches that may be subject to further revisions and is circulated to elicit comments and suggestions for enrichment and refinement. They contain findings on issues that are aligned with the core agenda of the research programs under the University of the Philippines Center for Integrative and Development Studies (UP CIDS).

Center for Integrative and Development Studies

Established in 1985 by University of the Philippines (UP) President Edgardo J. Angara, the UP Center for Integrative and Development Studies (UP CIDS) is the policy research unit of the University that connects disciplines and scholars across the several units of the UP System. It is mandated to encourage collaborative and rigorous research addressing issues of national significance by supporting scholars and securing funding, enabling them to produce outputs and recommendations for public policy.

The UP CIDS currently has twelve research programs that are clustered under the areas of education and capacity building, development, and social, political, and cultural studies. It publishes policy briefs, monographs, webinar/conference/forum proceedings, and the Philippine Journal for Public Policy, all of which can be downloaded free from the UP CIDS website.

The Program

The Program on Health Systems Development (PHSD) investigates problems of and innovates solutions for health systems and hopes to bridge some of the persistent divides that have hampered health systems in the Philippines. These include, among others, barriers between public health and clinical-surgical medicine; among levels and types of health professionals; among the differing perspectives and sometimes opposing interests of patients, health professionals, health policymakers, and other stakeholders; and between the public and private sector.

Get your policy papers published. Download open-access articles.

The Philippine Journal of Public Policy: Interdisciplinary Development Perspectives (PJPP), the annual peer-reviewed journal of the UP Center for Integrative and Development Studies (UP CIDS), welcomes submissions in the form of full-length policy-oriented manuscripts, book reviews, essays, and commentaries. The PJPP provides a multidisciplinary forum for examining contemporary social, cultural, economic, and political issues in the Philippines and elsewhere. Submissions are welcome year-round.

For more information, visit cids.up.edu.ph.
All issues/articles of the PJPP can be downloaded for free.

Get news and the latest publications.

Join our mailing list to get our publications delivered straight to your inbox! Also, you'll receive news of upcoming webinars and other updates.

bit.ly/signup_cids

We need your feedback.

Have our publications been useful? Tell us what you think.

bit.ly/dearcids

UNIVERSITY OF THE PHILIPPINES
CENTER FOR INTEGRATIVE AND DEVELOPMENT STUDIES

Lower Ground Floor, Ang Bahay ng Alumni, Magsaysay Avenue
University of the Philippines Diliman, Quezon City 1101

Telephone (02) 8981-8500 loc. 4266 to 4268
(02) 8426-0955

Email cids@up.edu.ph
cidspublications@up.edu.ph

Website cids.up.edu.ph