

Masci Scholar Personal Testimony



Ma. Sergia Fatima Sucaldito, MD

Clinical Faculty, Division of Adult Medicine, Department of Medicine, UP / PGH

Placement Site: Stanford University

Placement Date: Virtual AI in Healthcare Specialization (4 courses) June - December 2025

With the growing prominence of artificial intelligence (AI) applications in health and the consequent transformation of healthcare systems, it is imperative for providers working within them to have a firm grasp of the fundamental principles underpinning AI. This is a precursor to knowing the nuances of building, testing, and using AI for health purposes and, more broadly, knowing how to maximize the benefit/risk profile of AI for patients, providers, and communities.

If these fundamentals are not established, health providers, researchers, and decision-makers run the risk of embedding themselves into AI use paradigms that were built to serve the ruling tech oligarchy. Unfortunately, the goals and values pursued by these few prominent AI companies may deviate from the goals of patients and health systems, necessitating strong regulatory frameworks.

Though the four-course specialization I undertook through Stanford University's online programming was US-centric, there was a good progression of topics from a macro-level overview of the actors and interactions within the U.S. healthcare system, funneling into more micro-level knowledge and skills on clinical data and machine learning (ML).

The course wrapped up with a practical discussion of evaluating applications and how this evaluation was standardized by the U.S. FDA through regulatory pathways. Some courses were contextualized for the U.S. setting, but these remained interesting to me as I was able to compare and contrast the Philippine and U.S. healthcare systems. This comparison helped to inform unique use cases for AI and to deduce country-specific issues that could impact any future AI development and deployment.

Notably, in the U.S., most FDA-approved AI applications are intended for specialist use cases, such as detecting retinopathy or classifying lung lesions. I believe the Philippine healthcare system could stand to benefit more from AI if the use cases were at the primary care and public health levels.



In September 2025, Dr. Sucaldito (2nd from left) delivered a short lecture on the fundamentals of AI for medical education at the University of Northern Philippines.

The majority of AI's promise is predicated on the availability and accessibility of substantial, high-quality, and representative clinical data. The development of AI without this prerequisite sets the stage for expensive and time-consuming projects that will likely build an application that performs poorly at the intended purpose.

Understanding the constraints of our healthcare system, where a majority of clinical data is

locked away in paper records and digital records exist in vertical siloes, there is a discouraging dearth of high-quality clinical data.

As large tertiary academic center, the UP / PGH is in a unique position to help address the clinical data problem – a position that it is already starting to leverage through inter-institutional linkages.

Finding, collecting, curating, transforming, and cleaning data are resource-intensive tasks that require inputs from an interdisciplinary team with members from both tech and health sectors. Beyond making specialized data sets (i.e., data sets of brain tumors, optic fundi), I hope to contribute to building an ethically sourced and high-quality clinical data repository from routinely collected clinical data.

Currently, Filipino researchers interested in doing ML go through a grueling process to create their data sets from scratch. While this may be necessary for specialized or rare diseases, the potential for developing models from data sets of routinely collected data, such as demographics, vital signs, and lab values, remains untapped. If such a repository becomes available, this would remove many barriers for Filipino researchers interested in

research questions that could be answered by ML and help develop and test applications intended for the Filipino people.

For example, a standardized data set like the MIMIC-IV – a data set of ICU patients from the Beth Israel Deaconess Medical Center – is parent to hundreds of studies exploring the interactions among different patient factors. Though many of these studies did not translate to solid changes in clinical practice, the existence of the data set created fertile soil for studies that might do so in the future.

Though I mentioned the potential of building a clinical repository, the course also made me cognizant of the care we must take when deciding to use ML/AI in any endeavor. ML algorithms and other AI models are tools to complete specific tasks – whether classifying into groups, calculating probabilities, or generating human-like speech. Unlike typical statistical tests however, building these models is more time- and resource-consuming, requires greater computational reserve, and therefore generates substantially more carbon emissions.



In February 2026, a class of Year III medical students from UP Manila worked through individual case simulations using an LLM.

Because of this, for problems that can be answered using more frugal tools (i.e., good old logistic regression), the benefit/risk of ML becomes highly questionable. If the purpose is fulfilled just as well with simple tools, ML becomes superfluous and wasteful. I am therefore opposed to using AI/ML routinely for any use case or as a first choice to answer a research question, without prior investigation of other tools and methods. As faculty, I advocate for more conscientious guidance for our trainees and colleagues when AI projects are proposed.

As to how we can provide this guidance, we need to build skills in terms of evaluating AI project ideas and studies. Next to fundamentals, knowing how to evaluate the performance of these applications and deciphering the findings of AI studies is probably of similar importance. While not all providers will develop AI applications in their career, it is likely every provider will be faced with the decision to use or not use these applications in a digitally transformed healthcare system.

These AI applications are intended to affect clinical judgment and, at times, even guide management, making these tools just as impactful to patient care as diagnostic exams or medications. As such, we should be systematic and rigorous in the evaluation of these applications the same way we would critically gauge if we should give a drug – or not – to a particular patient.

Unlike the typical trials we interpret for drug efficacy, the “efficacy” of AI applications – that is, their performance – is expressed in diverse ways that may be unfamiliar to the provider, such as FI score, AUROC, calibration plots, and decision-curve analyses, among others. Furthermore, these applications are not immune to issues on ethics and privacy, transparency and interpretability, social responsiveness and inclusivity, fairness, and equity - necessitating the inclusion of these aspects in evaluation. The evaluation approach to this relatively novel type of study must be incorporated in existing evidence-based medicine modules.



Dr. Sucaldito helps facilitate hands-on activities and discussions on the use of AI tools in the research workflow among obstetrician-gynecologists in the Philippines on March 6, 2026.

As a result of my participation in this specialization, I have worked to share and disseminate what I have learned in the Philippines. These activities have to date included:

- **AI workshops with the Medical Informatics Unit (MIU):** I joined the UP MIU as facilitator for a series of workshops intended to build capacity for appropriately and responsibly using AI tools among two groups – medical educators and researchers. These workshops had participants from a wide range of ages, stages of career, and different institutions across the Philippines. The medical education workshop focused on allowing participants to use AI tools in curriculum development and teaching and learning activities in a controlled environment, as well as to think up their own AI use policy for courses. Similarly, the research workshop introduced and allowed experimentation with various MIU-vetted AI tools at different stages of the research process.
- **AI workshop for research with internal medicine residents:** Our residents have a research requirement for promotion to the next year level, so we incorporated a workshop on AI tools for research into our existing research workshop. This workshop introduced the basic principles of AI and how it can be applied in the research cycle as well as the risks of using AI applications in research. We incorporated periods of supervised experimentation to allow residents to apply some previously vetted AI applications to their own research questions.

- **AI-augmented learning activities for medical students:** As one of the coordinators for the internal medicine module for medical students, I have incorporated AI-augmented learning activities for diagnostic stewardship and telemedicine into the course to provide individualized case simulations followed by debriefing and synthesis.
- **Lectures at national and local conventions on basic principles and responsible use of AI:** I have assisted capacity-building efforts across different academic institutions and professional societies by delivering lectures at local and national conventions. These lectures focus on simplifying the basic principles of how AI works and describing real-world evidence and implications of applying AI at different points in the healthcare system. I aimed to build understanding as well as a healthy distrust of this disruptive technology.

Finally, I am also currently beginning work on my thesis for my master's degree in health informatics at UP College of Medicine and my initial proposal will be on equity evaluation of large language models (LLMs).

LLMs are widely available and are being used by providers, patients, and decision-makers alike. "Doctor Google" has evolved into "Doctor GPT". While prior benchmarking has focused mostly on LLM performance in terms of accuracy and efficiency, I believe that equity considerations should be as important, especially during the provision of healthcare-related information. If approved by my thesis panel, the study will aim to evaluate whether LLM-generated clinical information and reasoning introduce equity-related risks when applied in the Philippine health system context.