

Almost half of FDA-approved AI medical devices are not trained on real patient data, research reveals

August 26 2024, by Kendall Daniels



Credit: Pixabay/CC0 Public Domain

Artificial intelligence (AI) has practically limitless applications in health care, ranging from auto-drafting patient messages in MyChart to

optimizing organ transplantation and [improving tumor removal accuracy](#). Despite their potential benefit to doctors and patients alike, these tools have been met with skepticism because of patient privacy concerns, the possibility of bias, and device accuracy.

In response to the rapidly evolving use and approval of AI medical devices in [health care](#), a multi-institutional team of researchers at the UNC School of Medicine, Duke University, Ally Bank, Oxford University, Columbia University, and University of Miami have been on a mission to build public trust and evaluate how exactly AI and algorithmic technologies are being approved for use in patient care.

Together, Sammy Chouffani El Fassi, a MD candidate at the UNC School of Medicine and research scholar at Duke Heart Center, and Gail E. Henderson, Ph.D., professor at the UNC Department of Social Medicine, led a thorough analysis of clinical validation data for 500+ medical AI devices, revealing that approximately half of the tools authorized by the U.S. Food and Drug Administration (FDA) lacked reported clinical validation data.

Their findings were [published](#) in *Nature Medicine*.

"Although AI device manufacturers boast of the credibility of their technology with FDA authorization, clearance does not mean that the devices have been properly evaluated for clinical effectiveness using real patient data," said Chouffani El Fassi, who was first author on the paper.

"With these findings, we hope to encourage the FDA and industry to boost the credibility of device authorization by conducting clinical validation studies on these technologies and making the results of such studies publicly available."

Since 2016, the average number of medical AI device authorizations by

the FDA per year has increased from two to 69, indicating tremendous growth in commercialization of AI medical technologies. The majority of approved AI medical technologies are being used to assist physicians with diagnosing abnormalities in radiological imaging, pathologic slide analysis, dosing medicine, and predicting disease progression.

Artificial intelligence is able to learn and perform such human-like functions by using combinations of algorithms. The technology is then given a plethora of data and sets of rules to follow, so that it can "learn" how to detect patterns and relationships with ease.

From there, the device manufacturers need to ensure that the technology does not simply memorize the data previously used to train the AI, and that it can accurately produce results using never-before-seen solutions.

Regulation during a rapid proliferation of AI medical devices

Following the rapid proliferation of these devices and applications to the FDA, Chouffani El Fassi and Henderson et al. were curious about how clinically effective and safe the authorized devices are. Their team analyzed all submissions available on the FDA's official database, titled "Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices."

"A lot of the devices that came out after 2016 were created new, or maybe they were similar to a product that already was on the market," said Henderson. "Using these hundreds of devices in this database, we wanted to determine what it really means for an AI medical device to be FDA-authorized."

Of the 521 device authorizations, 144 were labeled as "retrospectively

validated," 148 were "prospectively validated," and 22 were validated using randomized controlled trials. Most notably, 226 of 521 FDA-approved [medical devices](#), or approximately 43%, lacked published clinical validation data.

A few of the devices used "phantom images" or computer-generated images that were not from a real patient, which did not technically meet the requirements for clinical validation.

Furthermore, the researchers found that the latest draft guidance, published by the FDA in September 2023, does not clearly distinguish between different types of clinical validation studies in its recommendations to manufacturers.

Types of clinical validation and a new standard

In the realm of clinical validation, there are three different methods by which researchers and device manufacturers validate the accuracy of their technologies: retrospective validation, prospective validation, and a subset of prospective validation called randomized controlled trials.

Retrospective validation involves feeding the AI model image data from the past, such as patient chest X-rays prior to the COVID-19 pandemic.

Prospective validation, however, typically produces stronger scientific evidence because the AI device is being validated based on real-time data from patients. This is more realistic, according to the researchers, because it allows the AI to account for data variables that were not in existence when it was being trained, such as patient chest X-rays that were impacted by viruses during the COVID pandemic.

Randomized controlled trials are considered the gold standard for clinical validation. This type of prospective study utilizes random

assignment controls for confounding variables that would differentiate the experimental and control groups, thus isolating the therapeutic effect of the device.

For example, researchers could evaluate device performance by randomly assigning patients to have their CT scans read by a radiologist (control group) versus AI (experimental group).

Because retrospective studies, prospective studies, and randomized controlled trials produce various levels of scientific evidence, the researchers involved in the study recommend that the FDA and device manufactures should clearly distinguish between different types of clinical validation studies in its recommendations to manufacturers.

In their *Nature Medicine* publication, Chouffani El Fassi, Henderson and others lay out definitions for the clinical validation methods which can be used as a standard in the field of medical AI.

"We shared our findings with directors at the FDA who oversee medical device regulation, and we expect our work will inform their regulatory decision making," said Chouffani El Fassi.

"We also hope that our publication will inspire researchers and universities globally to conduct clinical validation studies on medical AI to improve the safety and effectiveness of these technologies. We're looking forward to the positive impact this project will have on patient care at a large scale."

Algorithms can save lives

Chouffani El Fassi is currently working with UNC cardiothoracic surgeons Aurelie Merlo and Benjamin Haithcock as well as the executive leadership team at UNC Health to implement an algorithm in their

electronic health record system that automates the organ donor evaluation and referral process.

In contrast to the field's rapid production of AI devices, medicine is lacking basic algorithms, such as computer software that diagnoses patients using simple lab values in electronic health records. Chouffani El Fassi says this is because implementation is often expensive and requires interdisciplinary teams that have expertise in both medicine and computer science.

Despite the challenge, UNC Health is on a mission to improve the organ transplant space.

"Finding a potential organ donor, evaluating their organs, and then having the organ procurement organization come in and coordinate an organ transplant is a lengthy and complicated process," said Chouffani El Fassi.

"If this very basic computer algorithm works, we could optimize the organ donation process. A single additional donor means several lives saved. With such a low threshold for success, we look forward to giving more people a second chance at life."

More information: Not all AI health tools with regulatory authorization are clinically validated, *Nature Medicine* (2024). [DOI: 10.1038/s41591-024-03203-3](https://doi.org/10.1038/s41591-024-03203-3)

Provided by University of North Carolina Health Care

Citation: Almost half of FDA-approved AI medical devices are not trained on real patient data, research reveals (2024, August 26) retrieved 26 August 2024 from

<https://medicalxpress.com/news/2024-08-fda-ai-medical-devices-real.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.