npj | digital medicine
Published in partnership with Seoul National University Bundang Hospital

Brief communication

https://doi.org/10.1038/s41746-025-01800-1

How AI is used in FDA-authorized medical devices: a taxonomy across 1,016 authorizations

Check for updates

Rohan Singh[1], Monika Bapna[2], Abdul Rahman Diab[1], Emily S. Ruiz[2] & William Lotter[1,2,3]

We reviewed 1016 FDA authorizations of AI/ML-enabled medical devices to develop a taxonomy capturing key variations in clinical and AI-related features. Quantitative image analysis remains themost common application, but its relative proportion has declined recently. Over 100 devices leverage AI for data generation, though none yet involve LLMs. Our taxonomy clarifies current AI usage in medical devices and provides a foundation for tracking developments as applications evolve.

Over one thousand artificial intelligence (AI)/machine learning (ML)enabled medical devices have been authorized by the US Food and Drug Administration (FDA). Given the vast range of potential applications of AI in clinical care[1–5] and the increasing focus on translating AI into routine practice, it is critical to understand what types of AI devices are currently authorized for clinical use, and how these use cases are evolving over time. The FDA's current classification systems for medical devices, including product codes and device classes, provide a broad characterization that do not fully capture key dimensions of AI use. Recent work has provided insights into subsets of devices[6–11], but a comprehensive characterization of FDA-authorized AI-enabled devices is currently lacking.

We reviewed all 1016 authorizations of AI/ML-enabled medical devices listed by the FDA as of Decembar 20, 2024, and created a taxanomy to describe and quantify major axes of variatoin in current devices (Fig. 1). The taxonomy has three core factors. The first factor is the core clinical function of the device, describing the general medical role or purpose it serves in patient care. The second factor is the AI function, capturing how the device uses AI to assist with the clinical function. The third factor is the data type used as input to perform the AI function. By reviewing public authorization summaries for current devices, we categorized devices into different classesfor eachofthese factors. Asa given device canhave multiple authorizations over time for updated versions, we additionally grouped the 1016 authorizations into a list of unique devices through a semi-automated matching process (see Methods). We primarily report results at the device levelrather thanthe authorizationlevel toenable amore representativeview of the current device distribution.

Across 1016 FDA authorizations between 1995 and 2024, 836 unique devices were identified (Supplementary Table 1). For data type, 621 (84.4%) devices use images as the core input to the AI algorithm, 107 (14.5%) use signals, 5 (0.7%) use 'omics data, and 3 (0.4%) use tabular electronic health records(EHR).Fordevicesbasedonimaging,Radiologywastheleadreview panel for the majority (88.2%), followed by Neurology (2.9%) and Hematology (1.9%). Signals include time series such as ECG and EEG, with cardiovascular (64.5%), neurology (16.8%), and anesthesiology (12.1%) representingthemostcommonreviewpanels.Thefive'omics-baseddevices use data pertaining to RNA expression, DNA variants, and/or antibody assays as input to the AI/ML model. The three current EHR-based devices use tabular data such as treatment information and vital measurements as input.

In terms of clinical function, current devices fall into two broad categories based on whether they assist in patient assessment, such as diagnosis or monitoring, or intervention, such as surgery or treatment guidance. The majority of current devices fall under Assessment (84.1%). Of the 117 (15.9%) Intervention devices, 112 (95.7%) use images as the AI data type, followed by signals (3; 2.6%) and EHR (2; 1.7%). Current intervention applications for signals and EHR include assisting with insulin dosing based on either continuous glucose monitoring (signals) or tabular data (EHR), whereas common use cases for images include assisting with surgical or radiotherapy planning.

For AI function, we first classified devices based on whether the AI assists with the data generation and/or the data analysis process, as these categories often entail broadly different considerations for AI development and clinical implementation. We based this classification on the AI component of the device; for instance, an ultrasound machine that acquires images without using AI but has a built-in AI component to analyze the resulting images would be classified as "Analysis". With this definition, we classified 630 (85.6%) devices as Analysis, 83 (11.3%) as Generation, and 23 (3.1%) as both Analysis and Generation. For the devices that use AI for data generation, 99% had a data type of Images.

Beyond Analysis vs. Generation, we subclassified the AI function of each device into more fine-grained categories based on the device outputs and intended use (Fig. 2). For Analysis, we identified six subclasses across current devices: triage, quantification/feature localization, detection, diagnosis, detection/diagnosis, and predictive. Triage devices output a binary prediction for a given exam or time series that serves as a notification for prioritized review by a clinician. Devices classified as quantification/feature localization calculate a quantitative metric from the input data and/or provide positional information of structural features. This was considered a joint category because these

[1] Dana-Farber Cancer Institute, Boston, MA, USA. [2] Brigham & Women's Hospital, Boston, MA, USA. [3] Harvard Medical School, Boston, MA, USA. e-mail: lotterb@ds.dfci.harvard.edu

functions are often performed in tandem, such as segmenting an anatomical structure on a medical image and then quantifying its volume. Detection devices differ from quantification/feature localization in that they are specifically designed to assist clinicians in finding disease-suspicious regions in the input data, such as highlighting lung nodules on CT images or identifying the time interval of a potentially abnormal heart rhythm in an ECG trace. Devices indicated for diagnosis do not explicitly identify suspicious regions but instead output a score or category across the input data that indicates the likelihood of a specific diagnosis. Some devices assist with both detection and diagnosis, and we retain this as a distinct subclass as it is often considered as such in FDA regulation. Finally, predictive devices were defined as those that generate a score or category to indicate the future risk ofan event or disease rather than the current risk.

Across the 653 Analysis devices, 427 (65%; 58% of all devices) were subclassified as performing quantification/feature localization. Triage was the next most common Analysis subclass with84 devices (12.9% of Analysis devices;11.4%ofalldevices).Theremainingsubclasscountswereasfollow

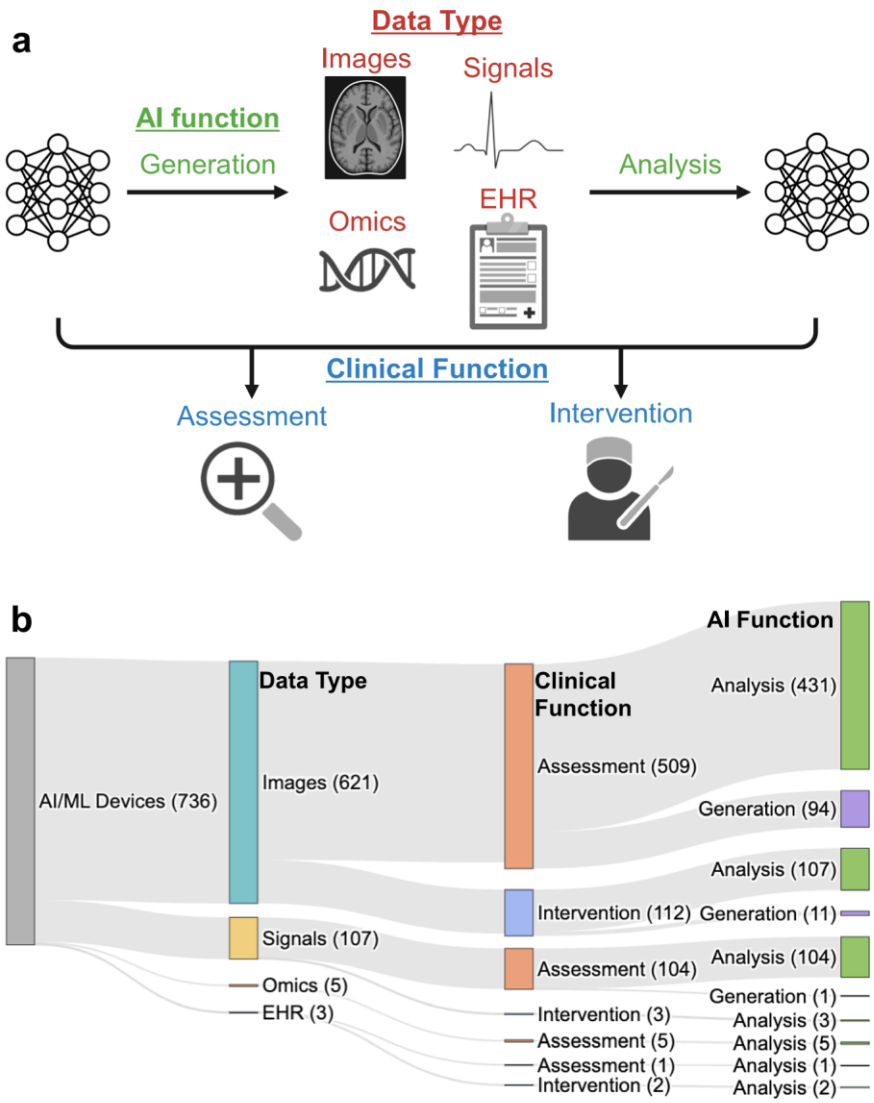

Fig. 1 | Taxonomy of AI/ML-enabled medical devices. a Illustration of the three core factors: data type, AI function, and clinical function. **b** Distribution of factor counts across FDA-authorized devices.

signal-baseddevices have a relatively higher percentage of diagnosis, detection/diagnosis, and predictive.

For generation devices, we identified three current subclasses: image enhancement, acquisition guidance, and synthetic data generation. Image enhancement includes tasks such as image denoising and AI-based reconstruction. Devicesthat use AI for acquisition guidance assist the clinical user in acquiring the appropriate data, such as indicating whether a body part or acquisition device is properly positioned. These devices do not explicitly generate the primary data used for downstream analysis but instead assist in properly capturing the data. Finally, synthetic data generation uses AI to generate a new data element altogether, such as synthesizing a CT medical image from an MRI or synthesizing a particular ECG bandwidth lead from existing leads.

Ofthe106Generationdevices,84(79.2%;11.4%ofalldevices)perfor m image enhancement. Acquisition guidance and synthetic data generation included 17 (16.0/2.3%) and eight (7.5/1.1%) devices, respectively. Images were the data type for 105 ofthe 106 Generation devices, withthe remaining device performing synthetic data generation

s: 47 (7.2/6.4%) diagnosis, 45 (6.9/6.1%) detection, 40 (6.1/5.4%) detection/ diagnosis, and 11 (1.7/1.5%) predictive. The proportion of devices in each Analysis subclass varied by data type and clinical function (Fig. 2). Imagebased devices have a relatively higher percentage of quantification/feature localization and triage, whereas using signals.

In addition to quantifying the current distribution of AI-enabled devices, we explored trends over time. peaked in 2021 at 94% and stands at 81% in 2024 so far (Fig. 3). Similarly, quantification/feature

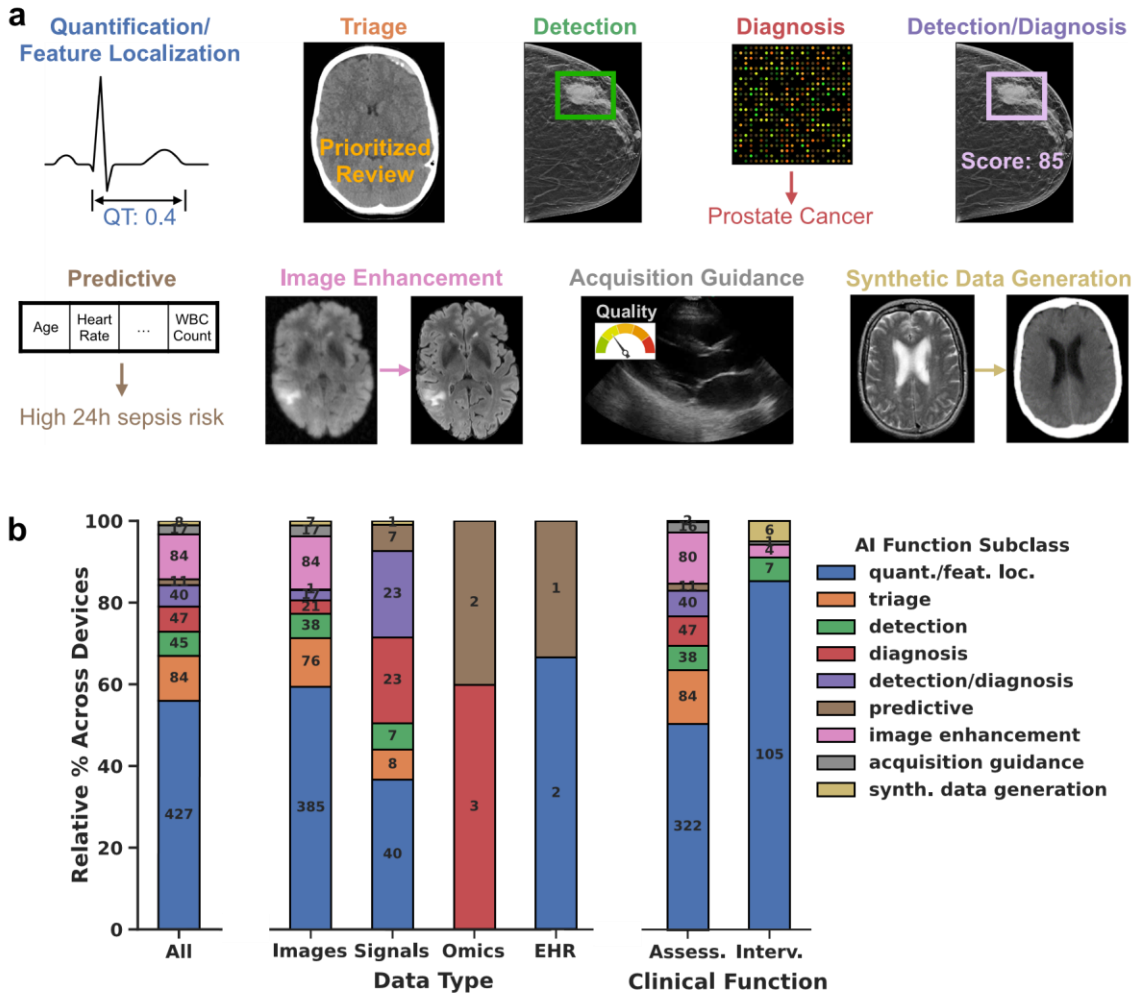

Fig. 2 | Subclasses of AI function in FDA-authorized medical devices. a Illustration of the nine subclasses. **b** Distribution of AI function subclasses across FDA-authorized devices. Counts indicate the number of devices with the AI function subclass.

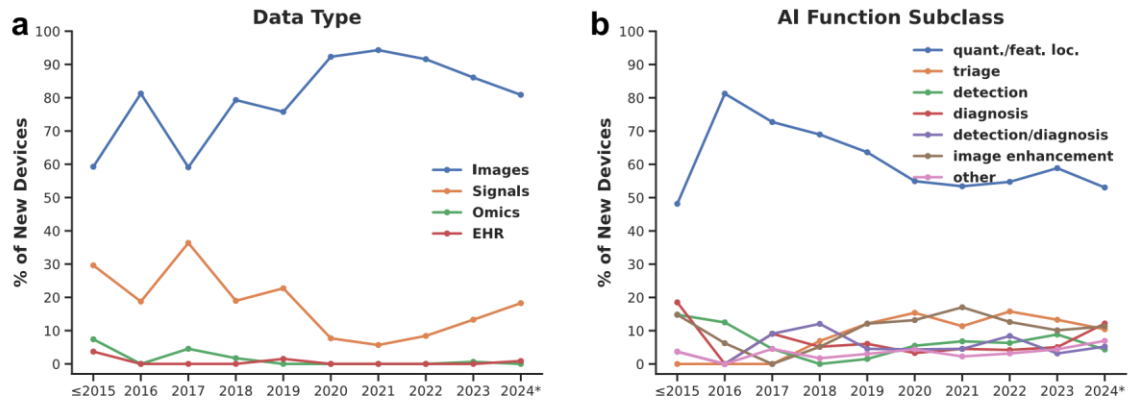

Fig. 3 | Trends in data type and AI function subclass over time. The percentage of new devices per year (*up to Sept. 27 for 2024) is displayed by a data type and b AI function subclass. The AI function subclasses of predictive, synthetic data generation, and acquisition guidance are grouped as "other" given low counts.

While Images remains the most common data type class, its relative proportion among new devices localization is the most common AI function subclass, but its prevalence peaked at 81% of

devices in 2016 and has declined to 51% in 2024. The proportion of triage and image enhancement especially increased between 2017 and 2021, with 2024 now containing a relatively mixed proportion of devices across the subclasses. For clinical function, the ratio of Assessment vs. Intervention has remained relatively consistent over time.

. The proportion of triage and image enhancement especially increased between 2017 and 2021, with 2024 now containing a relatively mixed proportion of devices across the subclasses. For clinical function, the ratio of Assessment vs. Intervention has remained relatively consistent over time. These trends also hold at the individual authorization levelratherthan the device level(Supplementary Figs. 2, 3). The percentage ofauthorizations that represent updated versions of existing devices has averaged 34% over 2022–2024, compared to 14% between 2017–2019 (Supplementary Fig. 1).

Our review of the 1016 authorizations revealed several other notable insights. First, out of the 69 product codes with more than one device, 19 (27.5%) contain non-uniform taxonomy values, where different devices under the same product code have different taxonomy classifications. These 19 codes account for 57% of all devices, driven in part by the popular radiology product codes QIH and LLZ, which together comprise 31% of devices. These findings highlight the importance of more fine-grained analysisbeyondFDAproductcodeswhenstudyingAI/ML-enableddevices. Second, while the specific AI/ML methods used by a device were often unclear, it is evident that the FDA employs a broad definition of AI/ML in creating its current device list. For instance, some devices explicitly refer to the use of traditional (non-deep learning) ML methods in their authorizationsummaries.Infact,theNVIproductcode,whichcurrentlyincludesone 'omics-based device, includes the K-nearest neighbors algorithm in its name: "Diagnostic software, K-nearest neighbor algorithm, autoimmune disease".Whendeeplearningwasused,certaindevicesindicatedthespecific model and/or architecture family, such as convolutional or recurrent neural networks. However, we did not find evidence of large language models (LLMs) in the studied device list. As LLM-based devices begin to receive authorization, they may or may not fit under existing taxonomy classes depending on their intended use. For instance, if an LLM-based device is specifically indicated to estimate sepsis risk from clinical notes, the device would fall under EHR-Assessment-Analysis-Predictive like existing devices performing the same function (Fig. 2). In contrast, novel use cases, such as radiology report generation, may entail the creation of new classes for certain taxonomy factors.

Along with a common lack of clarity regarding the form of AI/ML, it was unclear where AI/ML was used altogether in some authorization summaries. The public-facing authorization summary contains only a snapshotofthedevicecomparedtothefullsubmissionavailabletotheFDA. As such, we used the FDA's published list of authorized AI/ML devices rather than creating a separate list curated from authorization summaries, which may omit devices that describe AI/ML in the full submission but not the summary. This ambiguity regarding AI/ML use was particularly common for devices that perform many functions, some of which may or may not involve AI/ML. In such cases, we estimated the taxonomy classes based on similar devices, domain expertise, and/or available information from manufacturer websites. As a result, our findings should be interpreted in terms of overall trends rather than precise class counts. The varying level of detail present in the authorization summaries also made it infeasible to classify the form ofAI/MLusedfor eachdevice,soour taxonomyfocuses on the end use of the AI/ML algorithm regardless of its form. Our taxonomy also does not capture the full scope of the level of AI involvement for the intended clinical task, such as whether a quantification function is performed fully automatically or with potential clinician input/adjustment, as these nuances are also challenging to reliably determine from the summaries. Another limitation of our study is its reliance on manual curation (see Methods). While we explored various LLMs and prompting strategies for automated classification, these methods proved insufficient given the nuances and variability of the data. Future work could explore whether our curated database can now facilitate fine-tuning and/or in-context learning with LLMs to enable automated classification moving forward. Such approaches may also help overcome limitations in relying on the FDA's published list of AI/ML-enabled devices, which is updated irregularly (in April 2025, the latest authorization listed remains September 27, 2024). Specifically, by comparing authorization summaries of all FDA-authorized devices against those listed as AI/ML-enabled, LLM-assisted strategies combined with rule-based logic could help infer AI/ML usage even when it is not explicitly mentioned. Given that the FDA updates its general databases of authorization summaries weekly, such approaches could facilitate more continuous identification of AI/ML-enabled products and, in turn, updates to our taxonomy database.

Our analysis provides complementary insights to other recent reviews of FDA-authorized devices that have examined aspects such as predicate networks[12], marketing[13], explainability[7], regulatory pathways[14], and clinical validation[9,14–18].Eachoftheseaspectsareimportantforevaluatingindividual devices and the FDA regulatory process as a whole. Our taxonomy helps address the fundamental question—what are these devicesdesigned to do— but a subsequent critical question is—how well do they perform for these applications.Tothis end,recent analyseshavecalledforimprovedreporting of the clinical validation performed for FDA authorization[9,14,16,18] and a greateruse ofprospectivestudies overall[15,17].Ouranalysishelpsinformthese discussions by linking device function to the question of what constitutes appropriate validation for a given use case.

Ratherthan treating AI as a homogeneous technology in medicine, it is essentialtounderstandcurrentuse casesandevolvingtrends.Wecreatedan organized taxonomy of AI/ML-enabled medical devices by reviewing over one thousand FDA authorizations. This taxonomy and accompanying database are intended to provide utility for both clinical and research purposes, such as helping clinicians identify use cases relevant to their practice and aiding researchers in formulating novel AI applications. Our analysis of the database confirms some expected patterns, such as a high proportion of quantitative imaging devices, but also highlights notable recent shifts, such as increases in higher-risk use cases and signal-based devices. While LLMbased generative models are not yet represented, more than one hundred devices leverage AI for medical data generation, including image denoising ("super-resolution"[19]) and synthetic data creation. To facilitate further exploration, we have developed an interactive website (https://fda-aitaxonomy.vercel.app/) to host the curated database. Our taxonomy is designed not only to capture the current landscape but also to be extendable as new devices receive authorization. Continued monitoring of AI-enabled medical devices will be crucial for healthcare professionals, policymakers, and AI developers to fully realize AI's translational potential in medicine.

Methods

Database curation process

To create the database, the FDA's published list of AI/ML-enabled devices (https://www.fda.gov/medical-devices/software-medical-device-samd/ artificial-intelligence-and-machine-learning-aiml-enabled-medicaldevices) was analyzed. This list is maintained by the FDA and includes devices that have been authorized for marketing, which does not capture those that may be used through alternative regulatory pathways such as the investigational device exemption or humanitarian device exemption. The version of the list accessed corresponded to an update on December 20, 2024, which included devices authorized up to September 27, 2024. From this list, we curated the taxonomy over two phases. First, we reviewed all of the unique product codes across the devices, including reviewing the description and example authorizations for each code. The descriptions are available via the FDA's product code classification database, and authorization summaries were accessed using the FDA's 510(k), De-novo, and PMA databases. For each product code, we determined whether the taxonomy values could potentially vary across devices within the product code, or whether all devices would clearly have the same values. For example, the QFMproductcodehasadescriptionof"Radiologicalcomputer-aidedtriage andnotificationsoftware",whichwouldcorrespondtoadatatypeofImages and an AI function subclass of triage for all devices with this code. Every product code was reviewed by at least two members of the study team. Any product code for which any of the taxonomy levels could potentially vary by device were then considered for the second phase of review, which involved reviewingtheindividualauthorizationsummariesforeachdevicewithinthe product code. This review was performed by the senior author.

Taxonomy assignments

Each device was classified according to four factors: data type, clinical function,AIfunction,andAIfunctionsubclass.Thepossibleclassesforeach factor were defined through the review process to encompass major axes of variation acrossthedevices,andwere alsoinformedbyexistingcategoriesof medical devices[7,8], such as computer-aided detection ("CADe") and computer-aided diagnosis ("CADx"). Data type was assigned based on the core input to the AI algorithm in a mutually exclusive manner and could differ from the underlying data modality in some circumstances as our goal was to characterize how AI is actually being used. For instance, if the device involved a chemical assay but only involved AI to quantify the results based on a picture of the assay acquired using a smartphone camera, the data type would be classified as "Images". Clinical function was assigned based on the primary use case for the device according to the indications for use in the authorization summary. For imaging systems such as MRI machines that could theoretically be used for both diagnostic and interventional tasks, we considered the clinical use tobe"Assessment" unless the AI component was specifically designed to assist in interventions. Analogously, we assigned the AI function based on the output of the AI model(s) rather than the overall output(s) of the device. If a device had several AI-enabled components, we assigned the AI function for the device as a union across the individual components (i.e., a device could perform both data generation and analysis using AI).

Determining unique devices

Along with performing the taxonomy assignment for each individual authorization, we grouped authorizations into a list of unique devices, since a given device can have multiple submissions over time (i.e., for updated versions). Authorizations were grouped as the same device if they had matching product codes and similar company names, device names, and indicationsforuse.Thesegroupingswerefirstproposedautomaticallybased on case-insensitive matching of product codes, company names, and device names. The authorization list and proposed matches were then manually reviewed to make any corrections, including comparing the authorization summaries across potential matches to ensure similar indications for use whenconfirmingmatches,andseparatingpreliminarymatches that had the same device name but different indications for use. The latter scenario can occur when the device name on the submission represents a platform product, but the individual submissions pertain to unique AI algorithms (e.g., for detecting different diseases), in which case, we considered each submission unique. The manuscript includes results presented at both the authorization and device levels. When presenting results by device, we use the most recent authorization unless otherwise noted.

Trend analysis

When performing the trend analysis at the device level, the earliest authorization for each device was considered. Taxonomy counts per year were analyzed, where devices/authorizations before 2016 were grouped together, given low counts per year (33 total authorizations between 1995 and 2015, with a maximum of 6 per year over this timeframe).

Data availability

The curated taxonomy database is available as Supplementary Data 1.

Code availability

Code for the analysis in the manuscript is available at https://github.com/ lotterlab/fda-ai-taxonomy.

Received: 26 February 2025; Accepted: 14 June 2025;
Published online: 01 July 2025

References

1. Alowais, S. A. et al. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. BMC Med. Educ. 23, 689 (2023).
2. Rajpurkar, P., Chen, E., Banerjee, O. & Topol, E. J. AI in health and medicine. Nat. Med. 28, 31–38 (2022).
3. Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S. & Acharya, U. R. Deep learning for healthcare applications based on physiological signals: a review. Comput. Methods Prog. Biomed. 161, 1–13 (2018).
4. Lotter, W. et al. Artificial intelligence in oncology: current landscape, challenges, and future directions. Cancer Discov. 14, 711–726 (2024).
5. Sahni, N. R. & Carrus, B. Artificial Intelligence in U.S. Health Care Delivery. N. Engl. J. Med. 389, 348–358 (2023).
6. Milam, M. E. & Koo, C. W. The current status and future of FDA-approved artificial intelligence tools in chest radiology in the United States. Clin. Radiol. 78, 115–122 (2023).
7. McNamara, S. L., Yi, P. H. & Lotter, W. The clinician-AI interface: intended use and explainability in FDA-cleared AI devices for medical image interpretation. NPJ Digit. Med. 7, 80 (2024).
8. Zhang, K., Khosravi, B., Vahdati, S. & Erickson, B. J. FDA review of radiologic AI algorithms: process and challenges. Radiology 310, e230242 (2024).
9. Ebrahimian, S. et al. FDA-regulated AI algorithms: trends, strengths, and gaps of validation studies. Acad. Radiol. 29, 559–566 (2022).
10. Morey, J. et al. FDA reviewed artificial intelligence-enabled products applicable to emergency medicine. Am. J. Emerg. Med. 89, 241–246 (2024).

11. Brewster, R. C. L., Nagy, M., Wunnava, S. & Bourgeois, F. T. US FDA approval of pediatric artificial intelligence and machine learningenabled medical devices. JAMA Pediatr. 179, 212–214 (2025).
12. Muehlematter, U. J., Bluethgen, C. & Vokinger, K. N. FDA-cleared artificial intelligence and machine learning-based medical devices andtheir 510(k)predicate networks.LancetDigit. Health5, e618–e626 (2023).
13. Clark, P., Kim, J. & Aphinyanaphongs, Y. Marketing and US Food and Drug Administration clearance of artificial intelligence and machine learning enabled software in and as medical devices: a systematic review. JAMA Netw. Open 6, e2321792 (2023).
14. Joshi, G. et al. FDA-approved artificial intelligence and machine learning (AI/ML)-enabled medical devices: An updated landscape. Electronics 13, 498 (2024).
15. Chouffani El Fassi, S. et al. Not all AI health tools with regulatory authorization are clinically validated. Nat. Med. 30, 2718–2720 (2024).
16. Muralidharan, V. et al. A scoping review of reporting gaps in FDAapproved AI medical devices. NPJ Digit. Med. 7, 273 (2024).
17. Wu, E. et al. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat. Med. 27, 582–584 (2021).
18. Khunte, M. et al. Trends in clinical validation and usage of US Food and Drug Administration-cleared artificial intelligence algorithms for medical imaging. Clin. Radiol. 78, 123–129 (2023).
19. Umirzakova, S., Ahmad, S., Khan, L. U. & Whangbo, T. Medical image super-resolution for smart healthcare applications: a comprehensive survey. Inf. Fusion 103, 102075 (2024).

Acknowledgements

W.L. acknowledges support from the Ellison Foundation, the Wong Family Award, and the National Institute of Biomedical Imaging and Bioengineering award R21EB035247.

Author contributions

Conceptualization: R.S. and W.L.; Analysis: R.S., M.B., A.R.D. and W.L.; Website creation: R.S.; Drafting of manuscript: M.B. and W.L.; Review and approval of manuscript: all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-025-01800-1.

Correspondence and requests for materials should be addressed to William Lotter.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material isnot includedin thearticle's CreativeCommons licenceandyourintended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/bync-nd/4.0/.

© The Author(s) 2025