

Metrics and Quality Attributes of AI-Based Software as a Medical Device

Shouki A. Ebad

Center for Scientific Research and Entrepreneurship, Northern Border University, Arar, Saudi Arabia
shouki.abbad@nbu.edu.sa (corresponding author)

Radhia Zaghoud

Department of Computer Science, Northern Border University, Arar, Saudi Arabia
radhia.zaghoud@nbu.edu.sa

Achraf Ben Miled

Department of Computer Science, Northern Border University, Arar, Saudi Arabia
ashraf.benmilad@nbu.edu.sa

Received: 12 April 2025 | Revised: 16 May 2025 | Accepted: 24 May 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.11442>

ABSTRACT

Software as a Medical Device (SaMD) refers to software designed for medical purposes that operates independently of any physical hardware. Artificial Intelligence (AI)-based SaMD refers to standalone software utilizing Machine Learning (ML) and other AI techniques to perform medical functions without being part of a physical device. With the growth in research and practical applications of AI-SaMD, a variety of quality attributes have been proposed. However, the existing body of work lacks a comprehensive and cohesive review from a quality assurance perspective. This research paper aims to fill this gap by conducting a systematic analysis and synthesis of literature published between 2015 and 2024 to identify commonly addressed quality attributes, the metrics used to evaluate them, and emerging directions for future research. The identified key quality attributes included accuracy, AI model transparency, safety, performance, and Specificity and Sensitivity (S&S) as the most frequently used criteria for assessing the quality of AI-SaMD. The findings highlight significant gaps in the current research landscape, particularly the absence of a universal metric to comprehensively evaluate the quality of AI-SaMD. Furthermore, there is a lack of alignment between identified quality attributes, their corresponding metrics, and clinical validation standards, such as Food and Drug Administration (FDA) guidelines. The study concludes by offering valuable insights for stakeholders, including clinicians, technologists, and policymakers, and by outlining promising avenues for future research.

Keywords-Artificial Intelligence (AI); AI-Software as a Medical Device (SaMD); quality; measurement; attributes

I. INTRODUCTION

SaMD refers to software designed for medical purposes that functions independently of physical devices [1]. It operates on general-purpose platforms, such as smartphones or computers, without relying on dedicated hardware. AI is a simulation of human intelligence in machines to perform human tasks, including reasoning, problem-solving, and learning. In contrast, Machine Learning (ML) is a subset of AI that involves training algorithms to recognize patterns and make decisions based on data without being explicitly programmed [2]. Generally, applying AI techniques to software-intensive systems has produced promising results [3]. AI-SaMD refers to standalone software utilizing ML and other AI techniques to perform medical functions without being part of a physical device. Examples include AI tools for image analysis in radiology,

predictive analytics for chronic disease management, and decision-support systems for clinicians.

The increasing adoption of SaMD highlights its significance and transformative role in healthcare. Figure 1 illustrates the sharp increase in the number of approved AI-SaMDs over the past decade, with more than half approved in just the last three years [4]. These technologies offer great potential for improved patient outcomes, enhanced diagnostic accuracy, and personalized treatment strategies. However, their deployment in real-world healthcare settings necessitates stringent quality assurance to ensure that they perform reliably and ethically. From a software engineering perspective, quality refers to the overall attributes of a software product that determine how well it meets the requirements defined by its users [5].

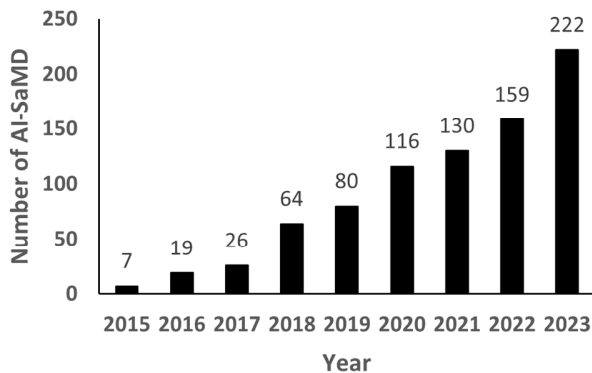


Fig. 1. Growth of AI-SaMDs from 2015 to 2023.

This paper focuses on critical quality attributes essential to AI-SaMD, namely: accuracy, model transparency, safety, performance, and S&S. A deeper understanding of these attributes, along with the challenges in their measurement and assurance, is substantial to developing effective and trustworthy AI-SaMD systems.

II. RESEARCH APPROACH

This study is an extension of our previous work [6], which employed a Systematic Literature Review (SLR) methodology. An SLR is a structured and rigorous approach to identifying, evaluating, synthesizing, and interpreting high-quality research relevant to a specific research question or topic [6]. This methodology has gained widespread adoption across various fields, particularly in interdisciplinary studies that combine Computing and Medicine, such as in [7, 8]. The outline of the five-stage process for conducting SLR is illustrated in Figure 2.

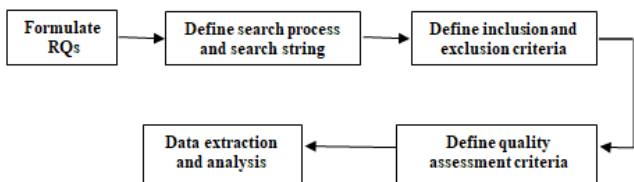


Fig. 2. Stages of an SLR.

A. Question Formulation

This study addresses three Research Questions (RQs), displayed in Table I, along with their motivations.

TABLE I. OVERVIEW OF RQS AND THEIR MOTIVATIONS

ID	Description	Motivation
RQ1	What are the key quality attributes of AI-SaMD?	Understanding the desired attributes is crucial for designing, developing, and evaluating AI-SaMD.
RQ2	What metrics are used to quantify these attributes?	This question addresses the need for quantifiable measures to assess AI-SaMD quality.
RQ3	What are the current research directions and gaps in the area of AI-SaMD quality?	This question encourages exploration of new approaches and challenges in the field.

B. Contribution

The contribution of this study stems from answering RQs in Table I, which form a solid foundation for research and discussion on the quality of AI-SaMD. They cover essential aspects of the development and evaluation process, ranging from identifying desired attributes (RQ1) and quantifying them (RQ2) to exploring future research directions (RQ3). By addressing these RQs, researchers and developers can contribute to the development of high-quality AI-SaMD that can ultimately improve healthcare outcomes.

C. Search Method

The keywords used in this research were: "AI," "artificial intelligence," "SaMD," and "software as a medical device." Using logical operators "AND" and "OR", the search command was constructed as follows: (AI OR artificial intelligence) AND (SaMD OR software as a medical device).

D. Source Selection

The search command was applied to databases featuring high-quality publications on technology and health. Exploratory searches identified PubMed, Semantic Scholar, SpringerLink, Web of Science (WoS), and IEEE Explore as the optimal repositories.

E. Study Selection

Studies were selected iteratively and incrementally, focusing on works written in English and published between 2015 and October 2024 to reflect significant AI advancements. Exclusion criteria ruled out unrelated papers, theses, posters, technical reports, commentaries, duplicates, and non-peer-reviewed studies, such as preprints. In cases of ambiguity, full-text reviews were performed to determine relevance. Out of an initial pool of 128 studies, 62 met all inclusion criteria and were classified as primary studies for this review.

F. Data Extraction

Data extraction, shown in Table II, was standardized, capturing essential details from each study, including title, journal, publication date, authors, methodology, results, objectives, challenges, clinical applications, data analysis, future directions, and recommendations. This structured approach provided comprehensive insights into AI-SaMD research. Details of each iteration are presented in Table III. The references cited within primary studies were not included in this SLR.

III. RESULTS AND DISCUSSION

A. Study Distribution on Quality Attributes

Existing research exhibits a diverse range of focus, with studies investigating various quality attributes of AI-SaMD rather than concentrating on a single one. Figure 3 provides a visual representation of these quality attributes and the proportion of studies dedicated to each. Notably, accuracy emerged as the most emphasized attribute, with 15 studies (24.2%) focusing on it. Model transparency and safety were addressed by 13 studies each (21%), reflecting significant interest in these areas. Performance and S&S were both considered by 11 studies (17.7%). General quality, on the other

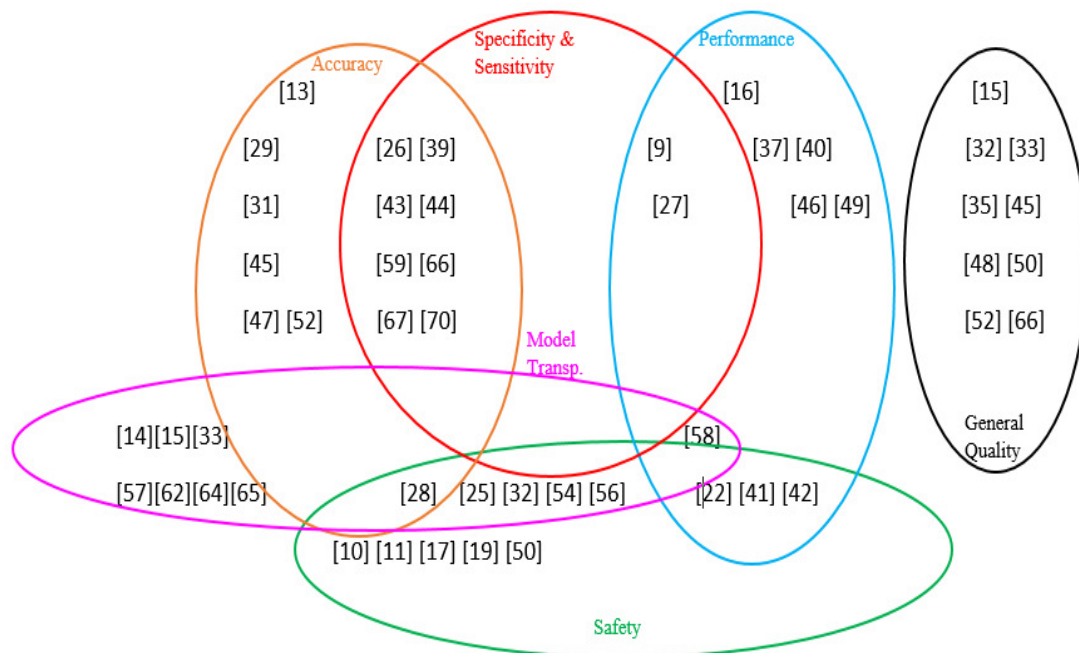


Fig. 4. Allocation of primary studies based on attributes; each enclosed in a circle and highlighted in a distinct color.

C. Measuring Attributes (RQ2)

Quantification of parameters that influence the quality is crucial for advancing research in the field of software-intensive systems [3]. As identified in the primary studies reviewed in this work, several efforts have been made to propose metrics for quantifying the key quality attributes outlined in the response to RQ1. Table V summarizes them in detail.

D. Research Directions (RQ3)

There are three directions for further research regarding AI-SaMD quality.

1) Missing Attributes

According to Figure 4, several potentially important quality attributes have received limited attention in the existing literature, such as regulatory compliance [14, 16], risk management [10, 17], and user experience [12, 58, 61]. Such attributes were excluded from the current analysis due to their infrequent appearance in primary studies.

2) Alignment with Validation Standards

Because the AI-SaMD stakeholders include different people from different fields, such as healthcare professionals, technology developers, and regulatory authorities, the above attributes and their metrics must align with clinical validation standards. Examples of such standards include the FDA guidelines, ISO 13485, and other national and international regulatory frameworks. These bodies play a vital role in ensuring that commercially deployed SaMD systems do not pose risks to public safety, the environment, or the economy. Without such alignment, the societal and market impact of AI-SaMD may remain limited, regardless of their technical innovation.

3) Proposing a Universal Metric

While the metrics outlined in Table V offer valuable insights into individual AI/ML-SaMD quality attributes, a unified evaluation framework is lacking. The complexity of AI-SaMD—arising from variations in use case, healthcare context, and end-user audience—suggests the need for a composite or universal metric that provides a holistic quality assessment. In the present work, it is proposed that such a metric be structured as a weighted composite function, accounting for both the technical (IT part) and clinical (Health part) dimensions of AI-SaMD, as shown in:

$$\begin{aligned} \text{Quality}(\text{AI/ML-SaMD}) = & \\ w_1 \cdot \text{Quality}(\text{IT part of AI/ML-SaMD}) + & \\ w_2 \cdot \text{Quality}(\text{Health part of AI/ML-SaMD}) & \quad (1) \end{aligned}$$

where w_1 and w_2 are weights reflecting the relative importance of each component. The IT part includes technical aspects, such as hardware (e.g., power and storage), AI model properties (e.g., transparency), software concerns (e.g., software business logic [74]), and connectivity (e.g., bandwidth). The Health part encompasses non-technical aspects with direct clinical implications, including regulatory compliance, clinician familiarity, and user experience.

It is worth noting that this metric would need to undergo both theoretical and empirical validation within the AI/ML-SaMD domain. Without such validation, the reliability and practical utility of any universal quality measure would remain questionable.

IV. THREATS TO VALIDITY

As with any SLR, this study is subject to several potential threats to validity, mainly selection bias and data extraction inaccuracies. The process of identifying relevant literature

inherently involves subjective judgment, which may influence the inclusion or exclusion of studies. While the current work made a deliberate effort to minimize this bias by following established SLR guidelines and conducting thorough searches, the accuracy of data extraction and interpretation remains a challenge. A further limitation lies in the dependence on a limited number of search engines and digital repositories. Although five prominent databases known for high-quality publications were selected, PubMed, Semantic Scholar, SpringerLink, Web of Science (WoS), and IEEE Xplore, the exclusion of other potentially relevant databases may have affected the completeness of the study corpus. Additionally, since this study primarily relied on keyword-based queries, the quality and indexing coverage of the selected search engines

could have impacted the comprehensiveness of the obtained results. There is also a possibility that recent publications relevant to this topic were missed due to the rapidly evolving nature of AI-SaMD research and the cut-off point for data collection in October 2024. Although the present study’s SLR includes the most relevant and influential works published during the specified timeframe, it should be acknowledged that this research should be regarded as a living document—one that may need periodic revision and expansion as new studies emerge, and standards evolve. Finally, although all materials, templates, and procedures used in this review can be shared to facilitate transparency and reproducibility, changes in digital library platforms over time may affect the repeatability of the search results.

TABLE IV. DETAILS OF THE MOST IMPORTANT QUALITY PROPERTIES OF AI-SAMD

Attribute	Definition	Affecting factors	Why important
Accuracy	The degree to which AI-SaMD predictions or outputs match the true clinical outcomes or ground truth.	Inaccurate data or inaccurate calculations. According to [71], the probability of a non-accurate output, containing at least one mistake, should not exceed 1%.	High accuracy ensures reliable medical decisions, reducing diagnostic errors and improving patient care. An inaccurate AI-SaMD risks misdiagnosis, inappropriate treatment, or delayed intervention.
Model transparency	The degree to how well the internal operations of the AI-SaMD model can be understood and explained, and why it makes specific decisions.	AI algorithm reliability, bias mitigation [33]	Transparent models make it easier to identify errors, biases, or risks. This is crucial for trust, accountability, and compliance with AI systems.
Safety	Ensuring that AI-SaMD does not cause harm to patients, clinicians, or systems during its operation.	Any hazardous operating situation with AI-SaMD. Unsafe conditions alarm reliability.	Without safety guarantees, these AI-SaMDs may lead to adverse events, jeopardizing patient trust and causing legal or financial repercussions.
Performance	Evaluating the AI-SaMD’s ability to deliver consistent and reliable outputs under various operating conditions.	Unlike performance from a computing perspective that focuses on resource consumption, such as CPU, workload, memory, and throughput [72], performance for AI-SaMD, instead, focuses primarily on functional effectiveness and clinical outcomes.	Consistent performance ensures the device’s usability in diverse clinical environments, safeguarding against unpredictable errors that could undermine its utility.
S&S	The ability of the SaMD to correctly identify True Negative (TN) cases and True Positive (TP) cases, respectively (i.e., correctly diagnosing non-lesions and lesions, respectively).	Quality and quantity of training data, such as imbalanced datasets, data diversity, and label accuracy.	High specificity prevents unnecessary interventions, reducing patient stress, resource wastage, and financial burdens on healthcare systems. In contrast, high sensitivity ensures that critical conditions are not overlooked, enabling timely interventions and better patient outcomes.

TABLE V. METRICS TO ESTIMATE THE QUALITY OF AI-SAMD

Attribute	Metric type	How to quantify
Accuracy	Quantitative	$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$
Model transparency	Qualitative	Conduct user studies where humans rate the transparency of the model explanations or predictions. The results are often scored and averaged. Frameworks, like SHapley Additive exPlanations (SHAP) or Local Interpretable Model-agnostic Explanations (LIME) are commonly used to assess and improve transparency.
Safety	Qualitative	The ordinal scale was selected as an appropriate scale for measuring both attributes, safety and model transparency. The answer to each of the Yes/No questions and the answers to all questions will be scored using the well-known rating scale, 5-point Likert. The minimum set of questions and scores encompasses four categories: (a) requirement specification, (b) functional analysis, (c) synthesis process, and (d) system analysis (as outlined in [71]). The Likert scale ranges from 0 to 5, where 0 represents 'no safety evidence is present' and 5 represents 'extensive evidence is present.'
S&S	Quantitative	Mathematically, specificity is defined as the True Negative Rate (TNR), as expressed in: $\text{Specificity} = \text{TNR} = \frac{TN}{TN+FP} \quad (3)$ $\text{Sensitivity} = \text{TPR} = \frac{TP}{TP+FN} \quad (4)$
Performance	Quantitative and Qualitative	As mentioned in Table I, the performance here does not focus on the resource consumption but on the clinical outcomes and functional effectiveness. Accordingly, the proposed metric to evaluate this performance can be performed through two means: (a) qualitative- a comparison with human experts and non-experts, although the number of clinicians involved was not specified, and (b) quantitative – aggregation metrics: $\text{Performance} = \alpha_1 \cdot \text{accuracy} + \alpha_2 \cdot \text{sensitivity} + \alpha_3 \cdot \text{specificity} \quad (5)$ where α_1 , α_2 , and α_3 are weights reflecting the importance of each metric for the application.

TP: True Positives, TN: True Negatives, FP: False Positives, FN: False Negatives, TNR: True Negative Rate, TPR: True Positive Rate

V. CONCLUSION

The integration of Artificial Intelligence (AI)-based Software as a Medical Device (AI-SaMD) in healthcare presents transformative opportunities for improving patient care. However, ensuring and maintaining high-quality standards is essential to realizing its full potential. This study presented a Systematic Literature Review (SLR) of 62 peer-reviewed publications from 2015 to 2024, aimed at synthesizing the current state of research on the quantification of AI-SaMD quality attributes.

The performed analysis yielded the following key findings: i) The five most frequently discussed quality attributes in the primary studies, in descending order of emphasis, were: accuracy (the most widely addressed by researchers), model transparency/explainability, safety, performance, and Specificity and Sensitivity (S&S). General quality was the least frequently examined attribute. ii) Each attribute is assessed using different types of measurement approaches: quantitative metrics (e.g., accuracy and S&S measured numerically), qualitative assessments (e.g., safety evaluated on an ordinal scale, model transparency via expert opinion). Or a combination of both (performance being assessed using both quantitative and qualitative methods). iii) The current literature reveals two critical research gaps: the absence of a universal metric for comprehensive evaluation of AI-SaMD quality, and a lack of alignment between existing quality attributes and clinical validation standards, such as ISO 13485 and related regulatory guidelines.

ACKNOWLEDGMENT

The authors extend their appreciation to Northern Border University, Saudi Arabia, for supporting this work through project number (NBU-CRP-2025-1564).

REFERENCES

- [1] IMDRF SaMD Working Group. "Software as a Medical Device (SaMD): Key definitions." International Medical Device Regulators Forum (IMDRF), Dec 2013. [Online]. Available: <https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf>.
- [2] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Fourth edition, Global edition. Boston: Pearson, 2022.
- [3] H. A. Al-Jamimi and M. Ahmed, "Machine Learning-Based Software Quality Prediction Models: State of the Art," in 2013 International Conference on Information Science and Applications (ICISA), Suwon, Korea (South), Jun. 2013, pp. 1–4, <https://doi.org/10.1109/ICISA.2013.6579473>.
- [4] The Konverge Team, "From code to cure: the potential of software as a medical device", Konverge, Nov. 2023. [Online]. Available: <https://www.konverge.com/blog/general/code-to-cure-potential-software-medical-device-samd/>.
- [5] International Organization for Standardization and International Electrotechnical Commission, *ISO/IEC 25010:2011: Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models*, First edition, 2011.
- [6] B. Kitchenham and P. Brereton, "A systematic review of systematic review process research in software engineering," *Information and Software Technology*, vol. 55, no. 12, pp. 2049–2075, Dec. 2013, <https://doi.org/10.1016/j.infsof.2013.07.010>.
- [7] S. A. Ebad, A. Alhashmi, M. Amara, A. B. Miled, and M. Saqib, "Artificial Intelligence-Based Software as a Medical Device (AI-SaMD): A Systematic Review," *Healthcare*, vol. 13, no. 7, Apr. 2025, Art. no. 817, <https://doi.org/10.3390/healthcare13070817>.
- [8] Y. Wang, M. M. Rosli, N. Musa, and F. Li, "Multi-Class Imbalanced Data Classification: A Systematic Mapping Study," *Engineering, Technology & Applied Science Research*, vol. 14, no. 3, pp. 14183–14190, Jun. 2024, <https://doi.org/10.48084/etasr.7206>.
- [9] N. Ishizuka, "Statistical Review of Regulatory Requirements for AI Diagnosis," in *2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC)*, Osaka, Japan, Jul. 2024, pp. 1975–1976, <https://doi.org/10.1109/COMPSAC61105.2024.00315>.
- [10] A. Ramachandran, P. Malhotra, and D. Soni, "Current Regulatory Landscape of Software as Medical Device in India: Framework for Way Forward," in *2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, Mar. 2023, pp. 305–310.
- [11] S. Gerke, B. Babic, T. Evgeniou, and I. G. Cohen, "The need for a system view to regulate artificial intelligence/machine learning-based software as medical device," *npj Digital Medicine*, vol. 3, no. 1, p. 53, Apr. 2020, <https://doi.org/10.1038/s41746-020-0262-2>.
- [12] D. V. Gunasekaran *et al.*, "Acceptance and Perception of Artificial Intelligence Usability in Eye Care (APPRAISE) for Ophthalmologists: A Multinational Perspective," *Frontiers in Medicine*, vol. 9, Oct. 2022, Art. no. 875242, <https://doi.org/10.3389/fmed.2022.875242>.
- [13] J. P. Campbell *et al.*, "Artificial Intelligence for Retinopathy of Prematurity," *Ophthalmology*, vol. 129, no. 7, pp. e69–e76, Jul. 2022, <https://doi.org/10.1016/j.ophtha.2022.02.008>.
- [14] D. M. Hedderich *et al.*, "Artificial intelligence tools in clinical neuroradiology: essential medico-legal aspects," *Neuroradiology*, vol. 65, no. 7, pp. 1091–1099, Jul. 2023, <https://doi.org/10.1007/s00234-023-03152-7>.
- [15] A. Moglia, K. Georgiou, L. Morelli, K. Toutouzas, R. M. Satava, and A. Cuschieri, "Breaking down the silos of artificial intelligence in surgery: glossary of terms," *Surgical Endoscopy*, vol. 36, no. 11, pp. 7986–7997, Nov. 2022, <https://doi.org/10.1007/s00464-022-09371-y>.
- [16] V. Zinchenko *et al.*, "Changes in software as a medical device based on artificial intelligence technologies," *International Journal of Computer Assisted Radiology and Surgery*, vol. 17, no. 10, pp. 1969–1977, Jun. 2022, <https://doi.org/10.1007/s11548-022-02669-1>.
- [17] F. O'Driscoll, N. O'Brien, C. Guo, M. Prime, A. Darzi, and S. Ghafur, "Clinical Simulation in the Regulation of Software as a Medical Device: An eDelphi Study," *JMIR Formative Research*, vol. 8, Jun. 2024, Art. no. e56241, <https://doi.org/10.2196/56241>.
- [18] D. V. Parums, "Editorial: Global Initiatives Support the Use and Regulation of Digital Health Technology During the COVID-19 Pandemic," *Medical Science Monitor*, vol. 27, Oct. 2021, <https://doi.org/10.12659/MSM.935123>.
- [19] K. Palaniappan, E. Y. T. Lin, and S. Vogel, "Global Regulatory Frameworks for the Use of Artificial Intelligence (AI) in the Healthcare Services Sector," *Healthcare*, vol. 12, no. 5, Feb. 2024, Art. no. 562, <https://doi.org/10.3390/healthcare12050562>.
- [20] C. W. L. Ho, "Implementing the human right to science in the regulatory governance of artificial intelligence in healthcare," *Journal of Law and the Biosciences*, vol. 10, no. 2, Jul. 2023, Art. no. lsad026, <https://doi.org/10.1093/jlb/lsad026>.
- [21] J. A. Smith, R. E. Abhari, Z. Hussain, C. Heneghan, G. S. Collins, and A. J. Carr, "Industry ties and evidence in public comments on the FDA framework for modifications to artificial intelligence/machine learning-based medical devices: a cross sectional study," *BMJ Open*, vol. 10, no. 10, Oct. 2020, Art. no. e039969, <https://doi.org/10.1136/bmjopen-2020-039969>.
- [22] V. V. Zinchenko *et al.*, "Methodology for Conducting Post-Marketing Surveillance of Software as a Medical Device Based on Artificial Intelligence Technologies," *Sovremennyye tehnologii v medicine*, vol. 14, no. 5, Oct. 2022, Art. no. 15, <https://doi.org/10.17691/stm2022.14.5.02>.
- [23] C. J. Lee *et al.*, "Pivotal trial of a deep-learning-based retinal biomarker (Reti-CVD) in the prediction of cardiovascular disease: data from CMERC-HI," *Journal of the American Medical Informatics Association*,

- vol. 31, no. 1, pp. 130–138, Dec. 2023, <https://doi.org/10.1093/jamia/ocad199>.
- [24] J. E. Carolan, J. McGonigle, A. Dennis, P. Lorgelly, and A. Banerjee, "Technology-Enabled, Evidence-Driven, and Patient-Centered: The Way Forward for Regulating Software as a Medical Device," *JMIR Medical Informatics*, vol. 10, no. 1, Jan. 2022, Art. no. e34038, <https://doi.org/10.2196/34038>.
- [25] D. Giansanti, "The Regulation of Artificial Intelligence in Digital Radiology in the Scientific Literature: A Narrative Review of Reviews," *Healthcare*, vol. 10, no. 10, Sep. 2022, Art. no. 1824, <https://doi.org/10.3390/healthcare10101824>.
- [26] Z. Angehrn *et al.*, "Artificial Intelligence and Machine Learning Applied at the Point of Care," *Frontiers in Pharmacology*, vol. 11, Jun. 2020, Art. no. 759, <https://doi.org/10.3389/fphar.2020.00759>.
- [27] S. Muthamilselvan, P. Ramasami Sundhar Baabu, and A. Palaniappan, "Microfluidics for Profiling miRNA Biomarker Panels in AI-Assisted Cancer Diagnosis and Prognosis," *Technology in Cancer Research & Treatment*, vol. 22, Jan. 2023, Art. no. 15330338231185284, <https://doi.org/10.1177/15330338231185284>.
- [28] A. Arora, "Conceptualising Artificial Intelligence as a Digital Healthcare Innovation: An Introductory Review," *Medical Devices: Evidence and Research*, vol. Volume 13, pp. 223–230, Aug. 2020, <https://doi.org/10.2147/MDER.S262590>.
- [29] E. Shimron and O. Perlman, "AI in MRI: Computational Frameworks for a Faster, Optimized, and Automated Imaging Workflow," *Bioengineering*, vol. 10, no. 4, Apr. 2023, Art. no. 492, <https://doi.org/10.3390/bioengineering10040492>.
- [30] K. A. C. Zapata, R. Patil, T. E. Ward, R. Loughran, and F. McCaffery, "Analysis of the Classification of Medical Device Software in the AI Act Proposal," in *HHAI Workshops*, 2023.
- [31] B. Waissengrin *et al.*, "Artificial intelligence (AI) Molecular Analysis Tool Assists in Rapid Treatment Decision in Lung Cancer: A Case Report," *Journal of Clinical Pathology*, vol. 76, no. 11, pp. 790–792, Nov. 2023, <https://doi.org/10.1136/jcp-2023-208991>.
- [32] K. Ganapathy, "Artificial Intelligence and Healthcare Regulatory and Legal Concerns," *Telehealth and Medicine Today*, Apr. 2021, <https://doi.org/10.30953/tmt.v6.252>.
- [33] J. M. Woodside, "Decision-Making Model Transparency for SaMD Machine Learning Algorithms," in *2020 Proceedings of the Conference on Information Systems Applied Research Virtual Conference*, 2020, vol. 13, Art. no. 5351.
- [34] R. Haneef *et al.*, "Innovative use of data sources: a cross-sectional study of data linkage and artificial intelligence practices across European countries," *Archives of Public Health*, vol. 78, no. 1, Dec. 2020, Art. no. 55, <https://doi.org/10.1186/s13690-020-00436-9>.
- [35] Å. Ingvar *et al.*, "Minimum Labelling Requirements for Dermatology Artificial Intelligence - Based Software as Medical Device (SaMD): A consensus statement," *Australasian Journal of Dermatology*, vol. 65, no. 3, May 2024, <https://doi.org/10.1111/ajd.14222>.
- [36] K. Dortche, G. McCarthy, S. Banbury, and I. Yannatos, "Promoting Health Equity through Improved Regulation of Artificial Intelligence Medical Devices," *Journal of Science Policy & Governance*, vol. 21, no. 03, Jan. 2023, <https://doi.org/10.38126/JSPG210302>.
- [37] S. H. Park, J.-I. Choi, L. Fournier, and B. Vasey, "Randomized Clinical Trials of Artificial Intelligence in Medicine: Why, When, and How?," *Korean Journal of Radiology*, vol. 23, no. 12, 2022, Art. no. 1119, <https://doi.org/10.3348/kjr.2022.0834>.
- [38] A. Lal, J. Dang, C. Nabzdyk, O. Gajic, and V. Herasevich, "Regulatory Oversight and Ethical Concerns Surrounding Software as Medical Device (Samd) And Digital Twin Technology In Healthcare," *Annals of Translational Medicine*, vol. 10, no. 18, pp. 950–950, Sep. 2022, <https://doi.org/10.21037/atm-22-4203>.
- [39] S. T. Suravarapu and K. Yetukuri, "Software Programmes Employed as Medical Devices," *World Journal of Current Medical and Pharmaceutical Research*, pp. 67–74, Jun. 2023, <https://doi.org/10.37022/wjcmpr.v5i3.268>.
- [40] E. J. Hwang *et al.*, "Use of Artificial Intelligence-Based Software as Medical Devices for Chest Radiography: A Position Paper from the Korean Society of Thoracic Radiology," *Korean Journal of Radiology*, vol. 22, no. 11, 2021, Art. no. 1743, <https://doi.org/10.3348/kjr.2021.0544>.
- [41] C. W. L. Ho, "When Learning Is Continuous: Bridging the Research–Therapy Divide in the Regulatory Governance of Artificial Intelligence as Medical Devices," in *The Cambridge Handbook of Health Research Regulation*, 1st ed., G. Laurie, E. Dove, A. Ganguli-Mitra, C. McMillan, E. Postan, N. Sethi, and A. Sorbie, Eds. Cambridge University Press, 2021, pp. 277–286.
- [42] F. Chothani, V. Movaliya, K. Vaghela, M. Zaveri, S. Deshpande, and N. Kanki, "Regulatory Prospective on Software as a Medical Device," *International Journal of Drug Regulatory Affairs*, vol. 10, no. 4, pp. 13–17, Dec. 2022, <https://doi.org/10.22270/ijdra.v10i4.545>.
- [43] J. K. A. Rathod, N. Bhavani, P. P. Saldanha, P. M. Rao, and P. Patil, "A Review of Current Applications of Artificial Intelligence and Machine Learning in Medical Science," *International Journal of Advanced Research in Science, Communication and Technology*, pp. 316–325, Nov. 2021, <https://doi.org/10.48175/IJARSC-2125>.
- [44] A. Petzold *et al.*, "Artificial intelligence extension of the OSCAR-IB criteria," *Annals of Clinical and Translational Neurology*, vol. 8, no. 7, pp. 1528–1542, Jul. 2021, <https://doi.org/10.1002/acn3.51320>.
- [45] P. Battista *et al.*, "Alzheimer's Dementia Early Diagnosis, Characterization, Prognosis and Treatment Decision Via a Software-as-Medical Device with an Artificial Intelligent Agent," *Alzheimer's & Dementia*, vol. 19, no. S18, Dec. 2023, Art. no. e075674, <https://doi.org/10.1002/alz.075674>.
- [46] Y. L. E. Thompson *et al.*, "Applying queueing theory to evaluate wait-time-savings of triage algorithms," *Queueing Systems*, vol. 108, no. 3–4, pp. 579–610, Dec. 2024, <https://doi.org/10.1007/s11134-024-09927-w>.
- [47] J. T. Megerian *et al.*, "Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder," *npj Digital Medicine*, vol. 5, no. 1, May 2022, Art. no. 57, <https://doi.org/10.1038/s41746-022-00598-6>.
- [48] S. Askin, D. Burkhalter, G. Calado, and S. El Dakrouni, "Artificial Intelligence Applied to clinical trials: opportunities and challenges," *Health and Technology*, vol. 13, no. 2, pp. 203–213, Mar. 2023, <https://doi.org/10.1007/s12553-023-00738-2>.
- [49] T. J. Loftus *et al.*, "Artificial intelligence-enabled decision support in nephrology," *Nature Reviews Nephrology*, vol. 18, no. 7, pp. 452–465, Jul. 2022, <https://doi.org/10.1038/s41581-022-00562-3>.
- [50] M. Rassi-Cruz, F. Valente, and M. V. Caniza, "Digital therapeutics and the need for regulation: how to develop products that are innovative, patient-centric and safe," *Diabetology & Metabolic Syndrome*, vol. 14, no. 1, Dec. 2022, Art. no. 48, <https://doi.org/10.1186/s13098-022-00818-9>.
- [51] H. Rakugi, "Healthy 100-year life in hypertensive patients: messages from the 45th Annual Meeting of the Japanese Society of Hypertension," *Hypertension Research*, vol. 47, no. 1, pp. 1–3, Jan. 2024, <https://doi.org/10.1038/s41440-023-01533-2>.
- [52] M. I. Miller, L. C. Shih, and V. B. Kolachalama, "Machine Learning in Clinical Trials: A Primer with Applications to Neurology," *Neurotherapeutics*, vol. 20, no. 4, pp. 1066–1080, Jul. 2023, <https://doi.org/10.1007/s13311-023-01384-2>.
- [53] L. Berliner, "Minimizing possible negative effects of artificial intelligence," *International Journal of Computer Assisted Radiology and Surgery*, vol. 19, no. 8, pp. 1473–1476, Apr. 2024, <https://doi.org/10.1007/s11548-024-03105-2>.
- [54] K. Zhou and G. Gattinger, "The Evolving Regulatory Paradigm of AI in MedTech: A Review of Perspectives and Where We Are Today," *Therapeutic Innovation & Regulatory Science*, vol. 58, no. 3, pp. 456–464, May 2024, <https://doi.org/10.1007/s43441-024-00628-3>.
- [55] J. He, S. L. Baxter, J. Xu, J. Xu, X. Zhou, and K. Zhang, "The practical implementation of artificial intelligence technologies in medicine," *Nature Medicine*, vol. 25, no. 1, pp. 30–36, Jan. 2019, <https://doi.org/10.1038/s41591-018-0307-0>.
- [56] L. Cerdá-Alberich *et al.*, "MAIC–10 brief quality checklist for publications using artificial intelligence and medical images," *Insights*

- into *Imaging*, vol. 14, no. 1, Jan. 2023, Art. no. 11, <https://doi.org/10.1186/s13244-022-01355-9>.
- [57] A. Agrawal, G. D. Khatri, B. Khurana, A. D. Sodickson, Y. Liang, and D. Dreizin, "A survey of ASER members on artificial intelligence in emergency radiology: trends, perceptions, and expectations," *Emergency Radiology*, vol. 30, no. 3, pp. 267–277, Mar. 2023, <https://doi.org/10.1007/s10140-023-02121-0>.
- [58] D. Dreizin *et al.*, "Artificial intelligence CAD tools in trauma imaging: a scoping review from the American Society of Emergency Radiology (ASER) AI/ML Expert Panel," *Emergency Radiology*, vol. 30, no. 3, pp. 251–265, Mar. 2023, <https://doi.org/10.1007/s10140-023-02120-1>.
- [59] R. J. Woodman and A. A. Mangoni, "A comprehensive review of machine learning algorithms and their application in geriatric medicine: present and future," *Aging Clinical and Experimental Research*, vol. 35, no. 11, pp. 2363–2397, Sep. 2023, <https://doi.org/10.1007/s40520-023-02552-2>.
- [60] F. Tettey, S. K. Parupelli, and S. Desai, "A Review of Biomedical Devices: Classification, Regulatory Guidelines, Human Factors, Software as a Medical Device, and Cybersecurity," *Biomedical Materials & Devices*, vol. 2, no. 1, pp. 316–341, Mar. 2024, <https://doi.org/10.1007/s44174-023-00113-9>.
- [61] E. J. Hwang *et al.*, "2023 Survey on User Experience of Artificial Intelligence Software in Radiology by the Korean Society of Radiology," *Korean Journal of Radiology*, vol. 25, no. 7, 2024, Art. no. 613, <https://doi.org/10.3348/kjr.2023.1246>.
- [62] J. Feng, S. Emerson, and N. Simon, "Approval policies for modifications to machine learning-based software as a medical device: A study of bio-creep," *Biometrics*, vol. 77, no. 1, pp. 31–44, Mar. 2021, <https://doi.org/10.1111/biom.13379>.
- [63] J. Herington *et al.*, "Ethical Considerations for Artificial Intelligence in Medical Imaging: Deployment and Governance," *Journal of Nuclear Medicine*, vol. 64, no. 10, pp. 1509–1515, Oct. 2023, <https://doi.org/10.2967/jnumed.123.266110>.
- [64] H. Dickinson, J. Feifel, K. Muelle, T. Ochi, and E. Vallejo-Yaguie, "Learning with an evolving medicine label: how artificial intelligence-based medication recommendation systems must adapt to changing medication labels," *Expert Opinion on Drug Safety*, vol. 23, no. 5, pp. 547–552, May 2024, <https://doi.org/10.1080/14740338.2024.2338252>.
- [65] P. Li, R. Williams, S. Gilbert, and S. Anderson, "Regulating Artificial Intelligence and Machine Learning-Enabled Medical Devices in Europe and the United Kingdom," *Law, Technology and Humans*, vol. 5, no. 2, pp. 94–113, Nov. 2023, <https://doi.org/10.5204/lthj.3073>.
- [66] A. Olorunfoba *et al.*, "Examining labelling guidelines for AI -based software as a medical device: A review and analysis of dermatology mobile applications in Australia," *Australasian Journal of Dermatology*, vol. 65, no. 5, pp. 409–422, Aug. 2024, <https://doi.org/10.1111/ajd.14269>.
- [67] D. D. Miller, "Machine Intelligence in Cardiovascular Medicine," *Cardiology in Review*, vol. 28, no. 2, pp. 53–64, Mar. 2020, <https://doi.org/10.1097/CRD.0000000000000294>.
- [68] D. Hua, N. Petrina, N. Young, J.-G. Cho, and S. K. Poon, "Understanding the factors influencing acceptability of AI in medical imaging domains among healthcare professionals: A scoping review," *Artificial Intelligence in Medicine*, vol. 147, Jan. 2024, Art. no. 102698, <https://doi.org/10.1016/j.artmed.2023.102698>.
- [69] K. Lim, T.-Y. Heo, and J. Yun, "Trends in the Approval and Quality Management of Artificial Intelligence Medical Devices in the Republic of Korea," *Diagnostics*, vol. 12, no. 2, Jan. 2022, Art. no. 355, <https://doi.org/10.3390/diagnostics12020355>.
- [70] M. Komatsu, N. Teraya, T. Natsume, N. Harada, K. Takeda, and Hamamoto, "Clinical Application of Artificial Intelligence in Ultrasound Imaging for Oncology," *JMA Journal*, vol. 8, no. 1, pp. 18–25, 2025, <https://doi.org/10.31662/jmaj.2024-0203>.
- [71] K. M. Adams, *Non-functional Requirements in Systems Analysis and Design*, First edition. Cham: Springer International Publishing, 2015.
- [72] R. B. Grady, *Practical software metrics for project management and process improvement*. Englewood Cliffs, NJ: Prentice Hall P T R, 1992.
- [73] I. Sommerville, *Software engineering*, Tenth edition. Boston: Pearson, 2016.
- [74] S. A. Ebad, "The influencing causes of software unavailability: A case study from industry," *Software: Practice and Experience*, vol. 48, no. 5, pp. 1056–1076, May 2018, <https://doi.org/10.1002/spe.2569>.