April 23, 2025

SAFE DEVELOPMENT OF MEDICAL DEVICES WITH GENERATIVE AI

I. INTRODUCTION

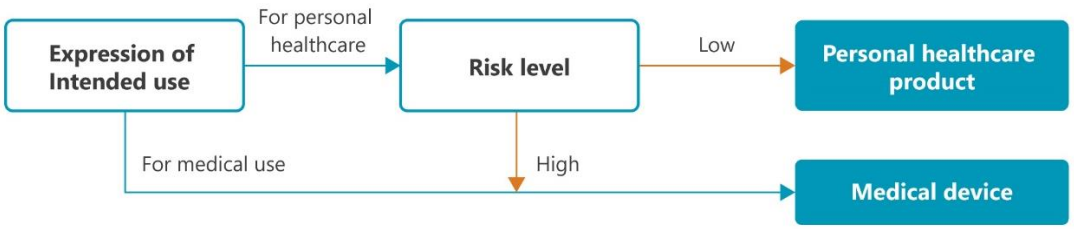
As artificial intelligence ("**AI**") technology rapidly evolves, its use in the medical field has expanded significantly. Previously, AI was primarily used for medical image analysis and diagnostic assistance in radiology, particularly with CT and MRI scans. However, recent advancements suggest that AI could extend its capabilities to mental health treatment through the use of generative AI. Clinical trials[1] involving mental health therapy, often referred to as "**Therabot,**" have shown that AI-based chatbots can have therapeutic effects on conditions such as major depressive disorder and anxiety disorders. Despite these promising developments, the use of generative AI is not without its challenges, including issues like hallucinations and bias. Such errors can pose serious risks patient's well-being and safety. Therefore, it is crucial to implement systemic controls to mitigate these hazards.

The Ministry of Food and Drug Safety (the "**MFDS**") is actively engaged in establishing international standards for the development of AI-based medical devices. In January 2025, it published the world's first "Guidelines on the Review and Approval of Generative AI-based Medical Devices", followed by the release of the 10 Principles of Good Machine Learning Practice for Medical Device Development in March 2025. This newsletter aims to provide a brief overview of the potential risks associated with AI medical devices and to summarize the key points of the MFDS guidelines designed to address these issues.

II. CRITERIA FOR MEDICAL DEVICES – INTENDED USE AND RISK LEVEL

The MFDS assesses whether AI software qualifies as a medical device by evaluating its **(a) intended use and (b) risk level**[2].

<Criteria for Medical Devices>[3]



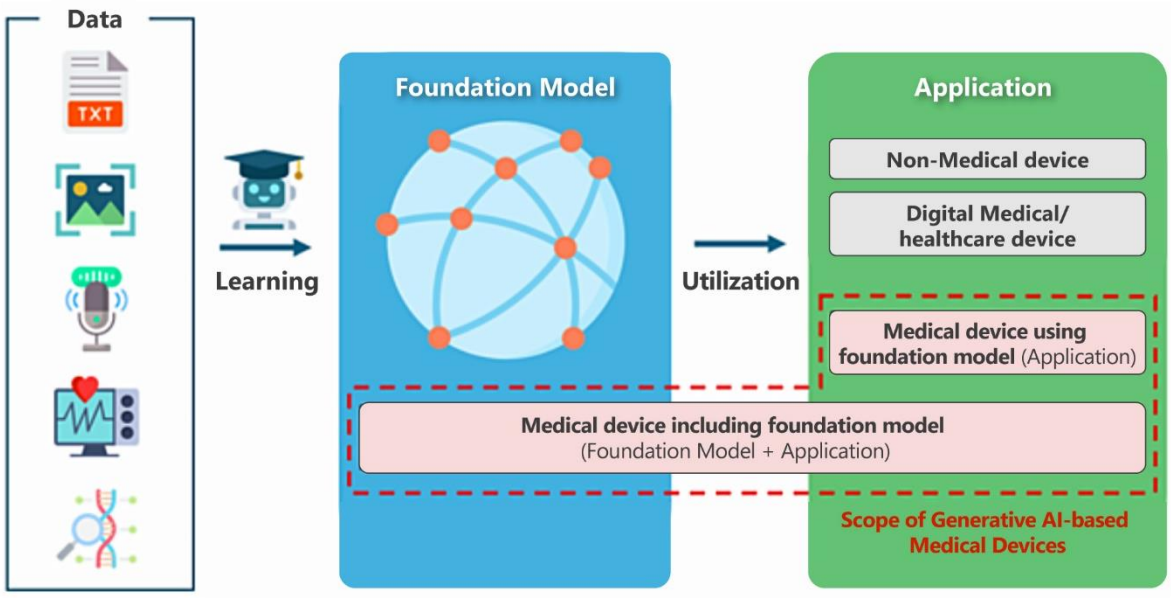
1 Kumar, S., Zhang, J., Kim, J., et al. (2024). Efficacy of a Generative AI Therapy Chatbot for Mental Health: A Randomized Clinical Trial. NEJM AI, 1(4). https://doi.org/10.1056/AIoa2400802

2 The MFDS, Guidelines on the Review and Approval of Artificial Intelligence (AI)-based Medical Devices, May 2022, See pp. 9-13

3 The MFDS, Criteria for Determining Medical Devices and Personal Healthcare (Wellness) Products, November 2020, See p. 7

(a) Intended use refers to the functions and purposes that the manufacturer intends for their product[4]. If the manufacturer claims that the AI software is designed for diagnosing, treating, and preventing diseases, then it is likely classified as a "digital medical device" based on intended use. However, even if generative AI technology is utilized, the product should not regarded as a 'generative AI medical device' if it does not qualify as a medical device or falls under the category of a digital medical or healthcare device[5].

<Scope of Management of Generative AI-based Medical Devices>[6]



Even if a product does not qualify as a medical device based on its intended use, it may still require approval by the MFDS as a medical device if it has a high **(b) risk level**. Risk level is assessed based on the following factors: (i) whether it poses biocompatibility issues, (ii) whether it is invasive, (iii) whether it could cause injury or illness to the user if it fails to operate as intended, (iv) whether it has the capability to detect emergencies, and (v) whether it can control or alter the function or characteristics of another device[7].

---

4 Intended use is determined by comprehensively considering the manufacturer's objectives, which are expressed in product labels, specifications, product descriptions, attachments, and advertisements. {The MFDS, Criteria for Determining Medical Devices and Personal Healthcare (Wellness) Products, November 2020, See p. 7}.

5 A 'digital medical/healthcare device' refers to any device, machine, apparatus, software, or similar product that does not qualify as a digital medical device but utilizes digital technology to monitor, measure, collect, and analyze bio-signals. These devices are designed to assist in medical practices or to maintain and improve health. The may also record and analyze lifestyle habits to provide healthcare information, such as guidance on diet and exercise. This definition is designated by the MFDS (Article 2(4) of the Digital Medical Products Act). Additionally, regulations governing the manufacturing, import, performance management, and distribution of digital medical/healthcare devices will come into effect on January 24, 2026.

6 The MFDS, Guidelines on the Review and Approval of Generative AI-based Medical Devices, January 2025, See p. 2

7 The MFDS, Criteria for Determining Medical Devices and Personal Healthcare (Wellness) Products, November 2020, See p. 8

III. HAZARDS OF AI MEDICAL DEVICES

Risk is a key criterion for distinguishing between medical devices and non-medical (or digital health) products. When developing software products that use AI, such as healthcare apps on smartphones, particularly those that do not include elements that directly interact with the patient's body (direct interaction is deemed absent if there are no issues with biocompatibility under (i) above and invasiveness (ii) above), emergency detection functions ((iv) above), and hardware control functions ((v) above), the key consideration is whether the software **could cause harm or illness to the user if it fails to function as intended** ((iii) above).

The Guidelines for the Review and Approval of Artificial Intelligence (AI)-Based Medical Devices (1st ed., MFDS, November 2017) outline several factors for determining the risk level of AI-based software products. These factors include errors in data or algorithms, as well as the explainability of clinical evidence—such as the source of the training datasets and the correlation between those datasets and the output.

<Criteria for Medical Devices – Factors for Determining Risk Level>[8]

(1) Whether the software is likely to cause harm to patients if it does not work as intended

"Medical software may cause harm to public health if accuracy is not guaranteed. For example, **inaccurate or inappropriate data input and learning**, or **errors in algorithms** may lead to incorrect predictions of the likelihood of disease, and incorrect detection or display of abnormal parts can directly affect diagnosis and treatment outcomes."

(2) Whether the software guarantees the clinical judgment of healthcare professionals

"When healthcare professionals decide on clinical diagnosis or treatment methods for patients, they should be informed that they do not make major judgments solely based on the recommendations (provided by the software), and **sufficient explanations should be provided on the source of the training dataset and correlation between the training dataset and output so that the healthcare professionals can understand the clinical basis for the information provided**."

Meanwhile, in January 2025, the MFDS published the world's first "Guidelines on the Review and Approval of Generative AI-based Medical Devices" in January 2025," which specifies hazards to be considered for generative AI medical devices by dividing them into six categories: performance, data quality, bias, user, adaptive system, and others. This was based on the "Technical Report on Artificial Intelligence and Machine Learning Based on the International Standard for Risk Management of Medical Devices (ISO 14971),[9]" published by the Association for the Advancement of Medical Instrumentation (AAMI[10]).

8 The MFDS, Guidelines on the Review and Approval of Artificial Intelligence (AI)-based Medical Devices, May 2022, See p. 9

9 AAMI TIR34971:2023. Application of ISO 14971 to Machine Learning in Artificial Intelligence – Guide. Association for the Advancement of Medical Instrumentation (AAMI), 2023.

10 The Association for the Advancement of Medical Instrumentation

<Examples of Hazards in Generative AI-based Medical Devices>[11]

Category	Examples of Typical Hazards
Performance	Hallucination, inconsistency, irrelevancy, no uncertainty indicator, etc.
Data Quality	Incorrect data, incorrect handling of outliers, data drift, domain shifted data, etc.
Bias	Selection bias, confounding variables, implicit bias, group attribution bias, etc.
User	Overconfidence, perceived risk, variation in social trust, etc.
Adaptive System	Continuous learning, etc.
Others	Lack of knowledge, etc.

IV. MAIN CONTENTS OF GOOD MACHINE LEARNING PRACTICE

On March 26, 2025, the MFDS published a set of guidelines titled "Good Machine Learning Practice ("**GMLP**") for Medical Device Development: Guiding Principles." These guidelines are a collaborative international document that was developed by the MFDS through its participation in the AI and machine learning working group of the International Medical Device Regulators Forum (IMDRF) over a two-year period. The IMDRF officially listed these guidelines in January 2025.

The guidelines outline 10 core principles that should be taken into account throughout the entire process of developing, training, verifying, and clinically applying AI-based medical devices. They provide specific criteria to ensure reliability, including the independence of data, representativeness of clinical trial data, appropriateness of model design, and ongoing risk monitoring. The MFDS hopes that these support the development of products that meet international standards and facilitate the advancement of Korean companies into the global market.

11 The MFDS, Guidelines on the Review and Approval of Generative AI-based Medical Devices, January 2025, See excerpt on pp. 8-11

<Good Machine Learning Practice for Medical Device Development: Guiding Principles> [12]

Main contents: 10 key principles

(1) Clearly define the intended use of medical devices and utilize multidisciplinary expertise throughout the entire product lifecycle.
(2) Apply best practices in software engineering, medical device design, and security principles throughout the entire lifecycle.
(3) Include datasets that accurately represent the intended patient population during clinical evaluations.
(4) Ensure that training and test data are set independently.
(5) Choose reference standards that are suitable for the intended use.
(6) Select and design models that align with the available data and the intended use of the medical devices.
(7) Evaluate medical devices by focusing on the interaction between humans and AI, emphasizing the performance of the Human-AI Team* in the intended user environment, rather than only assessing the devices' performance.
* Human-AI Team: A collaborative system where healthcare professionals make final diagnoses based on AI analysis results
(8) Test the performance of medical devices under clinically relevant conditions.
(9) Provide clear and essential information to users.
(10) Continuously monitor the performance of deployed models and manage the risk associated with retraining.

V. IMPLICATIONS

In alignment with the guidelines set by the MFDS, developers of generative AI-based medical devices must carefully consider the impact of these devices on patients' lives and bodies. Clearly defining the intended use is crucial to ensuring that potential hazards remain manageable. To achieve this, developers must secure appropriate and representative training data for clinical evaluation. They should also design a system architecture that supports healthcare professionals' judgment and interaction within real-world healthcare environments.

Korea is notable as the first country in the world to enact and implement a law—the Digital Medical Products Act—specifically to regulate AI-based medical devices. It has gained substantial experience in approving existing medical technologies, including AI-powered diagnostic assistive tools. The MFDS is also actively contributing to the establishment of international standards for AI medical devices through leading consultative bodies such as the International Standardization Organization (ISO) and the International Medical Device Regulators Forum (IMDRF). Given Korea's leadership and industry experience, the country is expected to serve as a strategic testbed for advancing generative AI-based medical devices in the global market.

12 The MFDS's press release dated March 26, 2025, <Publication of International Common Guidelines on Machine Learning for Medical Device Development>

In this context, companies developing generative AI-based medical devices must incorporate MFDS regulatory requirements from the earliest stages of development. They must also establish data strategies that ensure access to high-quality, clinically relevant data and proactively address potential biases. In addition, systems should be designed to support informed judgment and intervention by healthcare professionals, while considering user experiences within real clinical environments. Moreover, companies must stay current with relevant laws and guidelines, maintain comprehensive technical documentation and quality management systems, and remain responsive to regulatory changes both domestically and internationally.

Finally, companies should implement internal ethical standards and operational systems to protect patient data and mitigate risks inherent in generative AI, such as hallucinations and bias.

BKL's AI team is available to support you with a more effective and strategic roadmap for developing generative AI-based medical devices.

Related Professionals

Juho Yoon
Partner
T 82.2.3404.6542
E juho.yoon@bkl.co.kr

Jeonghee Kang
Partner
T 82.2.3404.6480
E jeonghee.kang@bkl.co.kr

Jiyoung Sohn
Senior Foreign Attorney (New York)
T 82.2.3404.0241
E jiyoung.sohn@bkl.co.kr

Seulki Lim contributed to this update.