



Addressing the Multi-Use Nature of Generative AI in Medical Devices

November 21, 2024

Traditional AI/ML and Generative AI (GenAI) pose some of the same challenges for monitoring performance, such as providing transparency and managing algorithmic bias. However, Generative AI, presents unique challenges to be considered in the context of medical devices due to output variability, complex evaluation metrics, and the risk of biased/harmful content.

Generative AI vs. Traditional AI



- More difficult to evaluate compared to purely quantitative models
- Errors can be harder to detect
- Difficult to test all potential input that could be fed into a GenAI model
- Lack of standard model metrics to assess performance
- Resource-intensive nature of monitoring due to computational power and significant human oversight needed

Potential Risks Posed by GenAI Outputs

Medical devices with patient/caregiver-facing GenAI outputs may require additional performance management approaches to mitigate the below risks.



Misinterpretation of Outputs

GenAI may generate incorrect medical advice due to data misinterpretation, leading to adverse health outcomes.



Jailbreaking Intended Use Cases

If not constrained with prompt engineering or rules-based logic, a large language model (LLM) might jailbreak its context and have a broader scope/search space than intended. Patients/caregivers are unlikely to be aware if this happens.



Potential Bias and Overreliance on Outputs

GenAI may exhibit biases, leading to unequal treatment recommendations, while overreliance on AI advice may reduce professional consultations, increasing mismanagement risks.

Approach to Creating Trustworthy Gen AI

Creating Trustworthy AI™ requires mitigating risk across multiple dimensions to create ethical safeguards while realizing the benefits associated with AI.

SAFE & SECURE

GenAI systems can be protected from risks (including cyber) that may cause harm. Models are secure and cannot be used as backdoors or to provide harmful, inappropriate, dangerous information.

Security techniques include prevention of model tampering and monitoring system and user prompts for malicious activity

ROBUST & RELIABLE

Models consistently produce accurate, coherent, and contextually appropriate responses across a wide range of queries and domains, do not hallucinate, and are resilient to erroneous inputs or adversarial accounts.

Ongoing performance monitoring includes validation of outputs when new data are incorporated into models

PRIVACY

Privacy is respected and personal data is not used beyond its intended and stated use. Prevents unauthorized access to or unintentional disclosure of sensitive information.

Patients are aware of how PHI may or may not be used in models

TRANSPARENT & EXPLAINABLE

Provides clear insights into GenAI decision-making processes and outputs, enabling users to understand and trust the information generated. Algorithms are open to inspection and can be explained by the responsible owner.

For devices in the home setting, outputs are available to and easily understood by patients

FAIR & IMPARTIAL

GenAI applications produce fair and unbiased results to all users. They are not influenced by biased data and do not favor certain groups over others.

Training data can be evaluated for bias, privacy, and representativeness for specified use cases

THANK YOU
QUESTIONS ?



About Deloitte

Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee (“DTTL”), its network of member firms, and their related entities. DTTL and each of its member firms are legally separate and independent entities. DTTL (also referred to as “Deloitte Global”) does not provide services to clients. In the United States, Deloitte refers to one or more of the US member firms of DTTL, their related entities that operate using the “Deloitte” name in the United States and their respective affiliates. Certain services may not be available to attest clients under the rules and regulations of public accounting. Please see www.deloitte.com/about to learn more about our global network of member firms.

This publication contains general information only and Deloitte is not, by means of this publication, rendering accounting, business, financial, investment, legal, tax, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte shall not be responsible for any loss sustained by any person who relies on this publication.