Transparency for Machine Learning-Enabled Medical Devices: Guiding Principles

June 2024

In 2021, the U.S. Food and Drug Administration (FDA), Health Canada and the United Kingdom's Medicines and Healthcare products Regulatory Agency (MHRA) jointly identified 10 guiding principles for good machine learning practice (GMLP). GMLP supports the development of safe, effective and high-quality artificial intelligence/machine learning technologies that can learn from real-world use and, in some cases, improve device performance.

Learn more: Good machine learning practice for medical device development: Guiding principles

The FDA, Health Canada and MHRA have further identified guiding principles for transparency for machine learning-enabled medical devices (MLMDs). These principles build upon the GMLP principles, especially:
- principle 7: Focus is placed on the performance of the human-AI team.
- principle 9: Users are provided clear, essential information.

While the guiding principles presented here promote transparency for MLMDs, transparency is a good practice to consider for all medical devices.

In this document, "transparency" describes the degree to which appropriate information about a MLMD (including its intended use, development, performance and, when available, logic) is clearly communicated to relevant audiences. "Logic" refers to information about how an output or result was reached or the basis for a decision or action. The degree to which this logic can be explained in a way that a person can understand is known as "explainability". Logic and explainability are aspects of transparency.

Effective transparency:
- ensures that information that could impact risks and patient outcomes is communicated.
- considers the information that the intended user or audience needs and the context in which it's used.
- uses the best media, timing and strategies for successful communication.
- relies on a holistic understanding of users, environments and workflows.

Another important concept related to transparency is "human-centered design". This is an iterative process that addresses the whole user experience and involves relevant parties throughout design and development. This approach can be used to help:
- develop MLMDs with a high degree of transparency.
- help validate transparency.
- ensure that users have all of the device-related information they need.

Learn more about human-centered design.

These guiding principles are intended as considerations when adopting and advancing good transparency practices. Continued engagement on this topic can help inform the collaborative development, implementation and iteration of good transparency practices and consensus standards in this rapidly evolving field.

We welcome your continued feedback through the FDA public docket (FDA-2019-N-1185) at Regulations.gov, and we look forward to engaging with you on these efforts. Contact us directly at Digitalhealth@fda.hhs.gov, mddpolicypolitiquesdim@hc-sc.gc.ca, and software@mhra.gov.uk.

The Guiding Principles

The guiding principles for transparency of MLMDs consider the following:

- who (relevant audiences)
- why (motivation)
- what (relevant information)
- where (placement of information)
- when (timing)
- how (methods used to support transparency)

Who: Relevant audiences

Transparency is relevant to:

- those who use the device.
    - such as health care professionals, patients, caregivers.
- those who receive health care with the device.
    - such as patients.
- additional parties, including those who make decisions about the device to support patient outcomes.
    - such as support staff, administrators, payors, governing bodies.

Why: Motivation

Transparency is essential to patient-centered care and for the safety and effectiveness of a device.

The use of MLMDs can involve understanding complex and context-dependent information. Transparent information can be used to identify and evaluate the device's risks and benefits and to help ensure the device is used safely and effectively.

Effective transparency that helps parties make informed decisions can help control risks. For example, a clear intended use supports an understanding of whether the device is intended to inform or replace the judgment of a health care provider.

Transparency helps parties detect and investigate errors or a decline in performance. It can also promote health equity, as knowledge of how a device works and how it was developed can help identify bias and assess whether a system or output is justifiable.

Upholding transparency over time can help support the maintenance and continued safety of a device.

The transparent and consistent presentation of information, including known gaps in information, can have many benefits. It builds fluency and efficiency in the use of MLMDs. It can also foster trust and confidence in the technology and encourages the adoption and access to beneficial technologies.

What: Relevant information

The type of information that is appropriate to share will vary across the range of MLMDs. It will depend on the benefits and risks of each MLMD and the needs of their intended users.

It's good practice to provide information that enhances understanding of the device and its intended use. A clear and accurate description of a device generally includes:

- information about its medical purpose and function.
- the diseases or conditions it's meant to address.
- intended users, use environments and target populations.

It's also good practice to explain how the device fits in the health care workflow. This would include a description of the intended inputs and outputs. It would also describe how the output is intended to impact health care decisions or actions, or the judgment of a health care professional. Details about device performance, benefits and risks can help one decide how best to use the device. Information about risk management activities across its lifecycle, such as bias management strategies, can also be helpful.

It's valuable to provide the basis of a device output or other information that explains how the MLMD reaches

its output (the "logic" of the model), when this information is available and easily understood. Access to this information helps one critically assess the device and its output when patient care decisions are being made. Another good practice is to provide information on product development and risk management activities across the total product lifecycle. Relevant information:

- explains the underlying technology and machine learning approaches.
- characterizes the training and testing data.
- provides summaries of clinical studies.
- provides ongoing updates on:
  - model and dataset characteristics.
  - performance monitoring.
  - detection and investigation of issues and risks across the total product lifecycle.

Clinically relevant limitations, gaps in information or specific contraindication statements are also good practice to communicate. This can include:

- known biases or failure modes.
- confidence intervals associated with the provided outputs.
- known gaps in the data characterization.
  - such as identifying any patient populations that are not well represented in training or clinical datasets and, therefore, may be at risk of bias.
- limitations in developing the model or evaluating model performance.
- known circumstances or use cases where the device input will not align with the data used to develop and validate the device.

It can also be helpful to provide information about how the safety and effectiveness of the device is maintained throughout the total product lifecycle, such as:

- how to conduct local site-specific acceptance testing or validation.
- plans for ongoing performance monitoring.
- transparent reporting of successes and failures.
- change management strategies.
- proactive approaches to address vulnerabilities.

Where: Placement of information

Device information can be accessed through the user interface. This interface includes all elements of the device with which the user interacts (those the user sees, hears and touches). For example, training, physical controls, display elements, packaging, labeling, and alarms are part of the user interface.

A good practice is to optimize use of the software user interface so that the information it conveys is responsive to the user. This software user interface can allow information to be personalized, adaptive and reciprocal. User needs may be addressed with a variety of modalities, including audio, video, on-screen text, alerts, diagrams, software safeguards and document libraries.

When: Timing of communication

Considering the information needs throughout each stage of the total product lifecycle can support successful transparency. Detailed device information may be needed when considering whether to acquire or implement a device, and whether and how to use it.

It may also be helpful to provide timely notifications when the device is updated or modified or when new information about the device is discovered. It may be appropriate to provide targeted information, such as on-screen instructions or warnings:

- at a specific stage in the workflow.
  - such as during high-risk steps.
- upon specific triggers.
  - such as when certain input or output features are present.

How: Methods to support transparency

Communicating information about MLMDs requires a holistic understanding of users, environments and workflows. These can be addressed by applying human-centered design principles.

Human-centered design means taking into account a human-centered perspective that is iterative, addresses the whole user experience and involves relevant parties throughout design and development. It involves responsive and iterative design, validation, monitoring and communications. For example, information is more accessible and usable if the following are considered:

- providing the appropriate level of detail for the intended audience.
- arranging the content in order of importance or to best support the ability of users to process the information and make informed decisions.

Plain language may be appropriate in some cases for understanding and usability. In other cases, technical language may be relevant for specialized clinical users.

These guiding principles for transparency are summarized in Table 1.

Table 1: Summary of transparency guiding principles

Guiding Principles	Description
Who: Relevant audiences for transparency	Transparency is relevant to all parties involved in a patient's health care, including those intended to: • use or receive health care with the device. make decisions about the device to support patient outcomes.
Why: Motivation for transparency	Transparency supports: • safe and effective use. • patient-centered care. • identification and evaluation of risks and benefits of a device. • informed decision-making and risk management. • device maintenance and detection of errors or performance degradation. • health equity through identification of bias. • increased fluency and confidence in MLMD use, increased adoption of the technology.
What: Relevant information	Enabling an understanding of the MLMD includes sharing relevant information on: • device characterization and intended use. • how the device fits into health care workflow, including the intended impact on the judgment of a health care professional. • device performance. • device benefits and risks. • product development and risk management activities across the lifecycle. • logic of the model, when available. • device limitations, including biases, confidence intervals and data characterization gaps. • how safety and effectiveness are maintained across the lifecycle.
Where: Placement of information	Maximizing the utility of the software user interface can: • make information more responsive. • allow information to be personalized, adaptive and reciprocal. • address user needs through a variety of modalities.
When: Timing of communication	Timely communication can support successful transparency, such as: • considering information needs at different stages of the total product lifecycle. • providing notifications of device updates. • providing targeted information when it's needed in the workflow.
How: Methods to support transparency	Human-centered design principles can support transparency.