IMDRF AIML WG
Comment Form

Date: July 9, 2026	Document: N93 Technical Framework for Artificial Intelligence Life Cycle Management

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	125-132	2.2	The proposed changes are intended to better align with the descriptions in the NIST Artificial Intelligence (AI) Risk Management Framework that is referenced in footnote 3. Specifically, the current text says that explainability includes both how and why decisions are made, which is contradictory to the NIST document, which says explainability relates to the "how" and interpretability relates to the "why."	Please revise as follows: "Explainability and Interpretability are also discussed throughout this document. Explainability "refers to a representation of mechanisms underlying AI systems' operation". It includes ~~not only~~ understanding how the AI system functions ~~decisions are made, but also why they are appropriate and trustworthy in the clinical context~~. Interpretability "refers to the meaning of AI systems' output in the context of their designed functional purposes". In other words, the ability for clinicians and healthcare professionals, patients, and other users to comprehend why ~~how~~ the device arrives at its outputs or recommendations. Understanding the difference between these concepts is important when applying considerations outlined throughout this document."	
GMTA	133-135	2.2	This document acknowledges that AI-enabled medical devices have a variety of users and can be deployed in a variety of environments, including at medical institutions, hospitals, or in other healthcare settings (collectively referred to as "sites").	This document acknowledges that AI-enabled medical devices have a variety of users and can be deployed in a variety of environments, including at medical institutions, hospitals, or in other healthcare settings (collectively referred to as "use environments")	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			The term “sites” in the context of AI or cyber could be misconstrued. It is therefore suggested to call these various environments as “use environments” to make this more intuitive and consistent with commonly used terminology.		
GMTA	141	3	• IEC 8100-5-1 defines the lifecycle for development and maintenance of health software, including medical devices. Additionally, this standard complements other standards that are referenced in the N93 draft document. As such, we recommend its inclusion as a cited reference within the IMDRF N93 document as it is pertinent to the lifecycle of AI-enabled medical devices. • The N93 draft document includes a discussion on cybersecurity considerations, and we recommend including the AAMI CR 515 standard as a suggested reference as it is recognized by some jurisdictions and contains	Recommend adding reference to the following standards that are pertinent to lifecycle management of AI-enabled devices: • IEC 81001-5-1:2021 Health software and health IT systems safety, effectiveness and security - Part 5-1: Security – Activities in the product life cycle • AAMI CR 515 - Cybersecurity considerations unique to machine learning – enabled medical devices. Recommend adding the following to the References section: “The list of references is not intended to be exhaustive; there may be additional references that are applicable to lifecycle management of AI-enabled medical devices. Inclusion of a reference in the list should not be interpreted as	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			information pertinent to cybersecurity considerations specific to medical devices. We appreciate the list of References included in the document; however we think it is important the document clarify that not all the cited references are applicable or recognized by all IMDRF jurisdictions and that not all references are applicable to medical devices, specifically. Additionally, we think it important for the document to recognize that standards are not static, and the applicability of the references may change in time. We recommend the References section be presented in a tabular format indicating which section(s) of the document the respective reference is applicable to. Given the volume of included references, the suggested presentation would help facilitate consistent understanding of the applicability of each listed reference.	indicating that it is recognized, accepted, or applicable in all jurisdictions. Similarly, not all of the listed references are specific to medical devices and may not be applicable to all AI-enabled medical devices. Consensus standards may be revised or amended over time. As such, it is recommended that stakeholders assess the applicability of the cited reference in the context of the respective jurisdiction and in the context of the individual device. If references are updated after the issuance of this document, stakeholders should compare the newer version to that included in this list to understand any changes that may impact their applicability." Recommend including the list of references in a tabular format indicating which section(s) of the document the respective reference is applicable to.	
GMTA	163 551 (footnote 14) 574	3 5.2	ISO/IEC 5259-4 is quite problematic and not relevant for machine learning. It is first referenced on line 551 but isn't needed there.	Remove reference to ISO/IEC 5259-4 (this also includes removing it from the "Additional Standards and References to be Considered" and the "References" sections)	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
	588				
GMTA	208 & throughout (below are some other line numbers where the proposed change is recommended) 385 733 745 752 1092 1094	multiple	As per IMDRF N81, the term SaMD is just one of many types of medical device software (MDSW). The N93 document is intended to apply to AI-enabled devices, broadly, and is not limited to SaMD. The proposed change is intended to make clear that the document is applicable to MDSW, not just SaMD.	Recommend replacing "SaMD" with "MDSW" throughout the document.	
GMTA	210-211	4.1	Risk management for AI-enabled medical devices should address both	Recommend revising as follows:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			safety and cybersecurity risks, including risks associated with model integrity, adversarial manipulation, unauthorized model modification, data poisoning, and compromised AI outputs. The risk management on the security aspect is addressed in the ISO 14971:2019, ISO/TR 24971:2020, and IEC 81001-5-1:2021	"AI-enabled medical devices benefit from implementation of scalable life cycle support processes that emphasize security- and safety-focused risk management throughout all life cycle steps."	
GMTA	210-211	4.1	Risk management for AI-enabled medical devices should address both safety and cybersecurity risks, including risks associated with model integrity, adversarial manipulation, unauthorized model modification, data poisoning, and compromised AI outputs. The risk management on the security aspect is addressed in the ISO 14971:2019, ISO/TR 24971:2020, and IEC 81001-5-1:2021	AI-enabled medical devices benefit from implementation of scalable lifecycle support processes that emphasize security- and safety-focused risk management throughout all lifecycle steps.	
GMTA	229	4.2	The document should highlight and align with the risk-based proportionality principles of existing regulatory frameworks. Please also refer to AdvaMed's General comment above.	Recommend adding the following sentence to the end of line 229: "Application of risk management mitigation efforts should be commensurate with the device's risk, including intended use, and technological characteristics."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	235	4.2.1	The risks identified in the bullet points are valid; however, attributing them directly to the “black box” nature of AI models is somewhat misleading.	Suggested to clarify that although the “black box” nature of AI models is a contributing factor, the root cause of the risks related to information is that the AI is producing outputs that are unknown.	
GMTA	240	4.2.1	It is recommended replacing “including” with “such as”. An explanation of the probability of error will not be necessary or appropriate for every product, so using “including” could imply that this explanation is expected in all cases. Using “such as” more clearly positions it as one possible example of incomplete information.	Please revise as follows: Incomplete information presentation, ~~including~~ such as lack of explanation on probability of error	
GMTA	235-241	4.2.1	This section of the document anchors too much on the “black box” nature of AI. While the “black box” nature of AI may be a contributing factor of risks related to information, the root cause is related to the transparency and explainability of the model. The suggested revisions to lines 235-238 are intended to address this issue. Also “probability of error” is ambiguous and hard to define. Suggested wording to enhance clarity	Recommend replacing lines 235 – 238 with: “Complex models where the underlying mechanism cannot be fully explained create unique transparency, explainability and interpretability concerns. Risks related to information can include:” Additionally, recommend revising line 240 as follows: “incomplete information presentation, including lack of transparency on the basis for the recommendation (e.g., Algorithm Confidence scores, known limitations and performance	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				characteristics) ~~explanation on probability of error~~"	
GMTA	254	4.2.2	De-learning of clinical knowledge is a potential risk associated with many innovations that simplify or automate aspects of clinical work and should not be characterized as an issue unique to AI-enabled technologies. In addition, this is not a risk that manufacturers are positioned to control or meaningfully mitigate through product design or labeling. Maintaining clinical competency and ensuring that healthcare professionals retain appropriate knowledge and skills are responsibilities that fall within the purview of professional licensing, accreditation, and continuing education bodies. Given that this issue is broader than AI and lies outside the scope of manufacturer responsibility, we recommend removing this risk from the document.	Recommend deleting the bullet in line 254 "~~• De-learning of clinical knowledge (due to over-reliance over time)~~"	
GMTA	271-272	4.2.4	The current title, "Risks Related to Deployment and Post-Market Monitoring and Performance," appears to place "post-market monitoring" in the same category as the risk source itself. Conceptually, the focus of this section appears to be on risks associated with deployment and post-market device	Recommend revising the title as follows: "Risks Related to Deployment and Post-Market ~~Monitoring and~~ Performance"	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			performance, whereas post-market monitoring primarily functions as a surveillance and risk control activity intended to identify, assess, and mitigate such risks. The current phrasing may therefore introduce ambiguity regarding the intended scope of the section.		
GMTA	279-280 1150	4.2.4 6	General purpose models are likely to be *integrated* into a device (i.e., connected, but not part of the device itself), rather than *incorporated* into a device.	Lines 279-280 Recommend revising as follows: "Performance degradation due to changes in third-party, general-purpose models incorporated or integrated into a device (See Section 5.3 for more information)." Line 1150: Recommend revising as follows: "When general-purpose, off-the-shelf models are incorporated or integrated, …"	
GMTA	281	4.2.4 Risks Related to Deployment…	Please provide an example of this so readers can understand the concern here and take appropriate action.	Provide an example of computational/infrastructure constraints.	
GMTA	284	4.2.4	Although "misaligned performance and calibration if context-of-use differs from validation" is valid, there is a distinction with unanticipated input data that the model was not trained or tuned to handle properly.	Recommend adding a bullet after line 284: "unanticipated input data that the model was not trained or tuned to handle properly."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	292-296	4.3	It should be recognized that manufacturers and developers of AI-enabled medical devices also play an important role in providing human oversight throughout the device lifecycle. In addition to the oversight exercised by end users in clinical settings, manufacturers contribute oversight through activities such as system design, validation and verification, performance monitoring, post-market surveillance, and the implementation of updates or corrective actions when issues are identified.	Recommend revising as follows: “Human oversight, including that of manufacturers, clinicians, health care providers, patients, and lay users, is essential throughout the entire life cycle of AI-enabled medical devices…”	
GMTA	292-296	4.3	Human oversight shouldn't be universal. Section 4.3 treats human oversight as "essential throughout the entire life cycle." This constrains autonomous and semi-autonomous devices. Oversight expectations should be context- and risk-dependent, and at the appropriate parts of the lifecycle based on the context of use.	Recommend revising as follows: “Human oversight, including that of clinicians, health care providers, patients, and lay users, should be considered in the context of device risk and at the relevant point(s) in the device ~~is essential throughout the entire~~ life cycle of AI-enabled medical devices to ensure that human and clinical expertise informs model development …”	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	297-301	4.3	The suggested edit is intended to more clearly establish the connection between transparency and effective human oversight. Specifically, highlighting the importance of transparency helps reinforce that users must understand when and how AI is used within the device, as well as how to interpret and appropriately act on its outputs, in order to support safe and effective human oversight.	Recommend revising as follows: "For example, human involvement in identifying user needs for an AI-enabled medical device can help to ensure that the device design accounts for the user's ability to interpret AI outputs, override or reverse automated recommendations, and intervene or interrupt automation. Transparency about the role of AI in the device and clear information on how it should be used safely and effectively are critical to enabling appropriate human oversight (see Section 6). As with all medical devices…"	
GMTA	302-305	4.3	Clinician involvement or input may not be relevant for all devices. Furthermore, there may be other subject matter experts, other than clinicians, who can play a role in supporting human oversight.	Recommend replacing lines 302-305 with the following: "Subject matter expertise and input, such as from clinicians or other health experts, can be leveraged in activities such as feature selection, data set selection, validating the clinical relevance of models, data labeling and annotation and identifying potential biases."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	309-310	4.3	The current wording is conceptually valid but lacks a practical implementation framework and introduces ambiguity regarding how manufacturers would realistically obtain information about automation bias in post-market settings and what types of "adjustments" would be expected or feasible. The statement may unintentionally imply ongoing manufacturer responsibility for monitoring clinician behavior and correcting cognitive biases, without clarifying the role of established post-market surveillance, quality management, or human factors processes in supporting such activities.	Recommend revising as follows: "This oversight can also help manufacturers to remain aware of the possible tendency of automatically relying or over-relying on the output produced by the AI-enabled medical device (automation bias) where such information is available through post-market monitoring, user feedback, complaint handling, or other quality management processes.~~and make adjustments for such biases~~."	
GMTA	311	4.3	Recommend the document acknowledge the various categories/degrees of autonomy to avoid unintended interpretations that equate all AI-enabled functionality with autonomous decision-making and reinforce the importance of human-in-the-loop designs in reducing risk and supporting safe clinical use.	Recommend adding the following after the sentence that ends on line 310: "It is important to note that there are varying degrees of autonomy in AI-enabled medical devices, from fully autonomous to non-autonomous. Non-autonomous AI-enabled medical devices provide recommendations, alerts, predictions, or decision-support outputs that are reviewable, modifiable, and overridable by a qualified human user, represent a distinct and common category of AI-enabled medical devices. The level of human oversight required for a medical device should be	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				commensurate with its degree of autonomy."	
GMTA	314	4.4	We recommend revising the name of this section to be more specific to the focus of the cyber content as it relates to integrity and data poisoning. This will help affirm that this section is not intended to provide general cybersecurity considerations that would be more appropriately found in IMDRF/CYBER WG/N60 FINAL:2020 Principles and Practices for Medical Device Cybersecurity.	Recommend revising the section title as follows: "Cybersecurity and Integrity Across the AI-Enabled Medical Device Lifecycle"	
GMTA	322-325	4.4	Typically, the goal is to protect the data and control access to it. The phrase "protect access to the data" is ambiguous.	"Mechanisms to ~~help~~ protect data and control access to ~~the~~ data used for the development and validation…"	
GMTA	325-327	4.4	For AI-enabled medical devices specifically, it is essential to follow best practices for mapping data sources and controlling data suppliers and labelers as part of QMS processes. As an addition to this section, we recommend addressing data provenance and chain-of-custody as it is important to enable manufacturers to identify and manage data supply-chain risks, including data poisoning that may enter through otherwise trusted sources	Recommend revising as follows: "For AI-enabled medical devices specifically, it is essential to follow best practices for mapping data sources and controlling data suppliers and labelers as part of QMS processes. This includes appropriate data provenance, dataset versioning, and chain-of-custody controls to support traceability and integrity of training, validation, and post-market data over time. This can help control…"	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			and remain difficult to detect without traceability controls.		
GMTA	326	4.4	Does not clearly indicate that manufacturers can have a shared responsibility with suppliers regarding control of data sources.	Add shared responsibility through supplier controls and agreements as clarification.	
GMTA	327-329	4.4	The reference to “unwanted bias” in this context may introduce conceptual ambiguity, as the surrounding discussion is focused primarily on cybersecurity, data integrity, supplier controls, and data poisoning risks. Bias-related concerns are more appropriately addressed in the context of data quality, representativeness, and model performance (e.g., Section 4.2.3). Removing the reference to “bias” would maintain alignment with the cybersecurity focus of this section. In addition, the phrase “post-market performance optimization” may be interpreted broadly or ambiguously in a regulatory context.	Recommend revising as follows: “This can help control ~~unwanted bias~~ ~~and~~ data poisoning and other risks stemming from ~~outside sources in~~ untrusted or compromised external data sources during both development and post-market ~~performance optimization (when carried out~~ model maintenance or improvement activities where conducted in accordance with appropriate regulatory requirements and change control processes ~~approval)~~.”	
GMTA	331	4.4	The phrase “cause-effect relationship between input data and AI-enabled medical device output” is somewhat imprecise. For many AI	Recommend revising as follows: “In some instances, due to the complex nature of AI-enabled medical device design, it might be difficult to recognize	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			models, the issue is not that there is no cause-effect relationship, but that the relationship may be difficult to interpret, trace, or explain. A better phrase would be “the relationship between changes in input data, training data, and model outputs may not be readily interpretable.”	when data poisoning occurs because the ~~cause-effect~~ relationship between changes in input data, training data, and model outputs may not be readily interpretable. ~~and~~ AI-enabled medical device output is not always transparent and explainable, such that malicious degradation may be thought to be within performance specifications of the model or mistaken for natural performance drift.”	
GMTA	337	4.4	The current cybersecurity discussion references several adversarial machine learning threats but does not fully address emerging risks. It is suggested to add broader language.	As outlined below, robust security controls and monitoring activities can help mitigate these vulnerabilities. AI-enabled medical devices may also be vulnerable to a broader range of adversarial machine learning attacks designed to intentionally manipulate model behavior or degrade performance.	
GMTA	339-342	4.4	This section would be improved with risk-based framing and reference to key aspects of considerations for mitigations that are omitted.	Recommend replacing lines 339-342 with the following: “Using device design processes to support secure product development and maintenance may include Threat Modelling, Cybersecurity Risk Assessments, mitigations appropriate and commensurate with device risk, as well as considerations of interoperability, third-party software components, unresolved cybersecurity anomalies, and other applicable cybersecurity risk	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				management activities throughout the product lifecycle."	
GMTA	343-353	4.4	The document recommends hosting the AI-enabled device locally to minimize cybersecurity threats. This implies that cloud connectivity introduces greater cyber risk. However, cloud architectures can provide equal or greater security. Recommendations in the document should be architecture-neutral. Cybersecurity expectations should be met through appropriate design regardless of deployment model.	Recommend lines 349-351 revising as follows: "Unwanted access may be mitigated via mechanisms such as; limiting the frequency of updates to batches, ~~hosting the AI-enabled medical device locally, if possible~~ and ensuring security controls are centralized to manage monitoring, auditing and validation more reliably."	
GMTA	348–351	4.4	In the context of cybersecurity concerns, unwanted data access may be a relevant risk. This section of the document, which discusses potential mitigations, would benefit from reference to data privacy, as indicated in the proposed redline text. The document recommends "limiting the frequency of updates to batches" as a mitigation strategy. However, this approach may not be a suitable mitigation in many instances. For example, immediate updates and patches may be required to address cyber issues.	Recommend revising as follows: "Unwanted access may be mitigated via mechanisms such as; considering data privacy expectations, ~~limiting the frequency of updates to batches~~, deploying the AI-enabled medical device within a secure, managed infrastructure (e.g., logically isolated cloud environments or secure local servers) with centralized security controls ~~hosting the AI-enabled medical device locally, if possible and ensuring security controls are centralized to manage monitoring, auditing and validation more reliably.~~"	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			The original phrasing implies cloud-based solutions are inherently less secure than locally deployed devices, however, this assumption may not be accurate. The suggested redline edits are intended to emphasize device deployment within secure infrastructure as a mitigation strategy, agnostic of the deployment environment.		
GMTA	356	4.4	While local may be a means, we would recommend to highlight that e.g., cloud-based solutions are likely much higher in terms of security depending on the nature of the AI solution and the medical device being utilized.	Revise to: "Unwanted access may be mitigated via mechanisms such as deploying the AI-enabled medical device within a secure, managed infrastructure (e.g., logically isolated cloud environments or secure local servers) with centralized security controls."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	357	4.4	The statement “The primary responsibility for validation of input data resides with the legal manufacturer” may be too broad depending on jurisdiction and deployment model. The manufacturer is generally responsible for ensuring that the device is safe and effective for its intended use, including defining data requirements, validation methods, controls, and instructions for use. However, healthcare organizations or users may be responsible for following those instructions and maintaining local data quality or IT controls. EU AI Act Art. 26 imposes this obligation on the deployer, rather than the provider. The paragraph should avoid implying that the manufacturer directly validates every input data item used in clinical operation unless that is actually the intended control model.	Recommend revising as follows: “~~The primary responsibility for validation of input data resides with the legal manufacturer; in practice, additional parties such as end users and sites at which the device is being deployed may also be involved.~~ The legal manufacturer is responsible for establishing and validating controls for data used by the device, including data requirements, provenance, preprocessing, access, validation, and defined responsibilities for all parties handling or supplying data. The legal manufacturer is responsible for device safety and performance, while end users and deployment organizations are responsible for complying with instructions for use, maintaining data quality, and protecting connected systems. If the model is autonomously adapting or updating, ~~and improving, it may warrant further scrutiny of interconnected systems that deliver this data and their attack surfaces and it remains critical that updates are verified for safety prior to implementation.~~ the manufacturer should evaluate the interconnected systems and data pipelines that supply the model, paying close attention to their vulnerabilities and potential risks for data poisoning or other disruptive attacks.”	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	389	5. AI-enabled Medical Device Life Cycle Steps. Figure 1. Appendix A	The lifecycle diagram is useful, but some steps are described as general considerations rather than auditable lifecycle outputs. This may limit practical implementation by manufacturers and regulators.	For each lifecycle stage, add expected outputs such as AI lifecycle plan, data management plan, dataset provenance record, model card or model description, verification and validation (V&V) protocol, clinical evaluation report, deployment verification record, monitoring plan, change-control plan, and sunsetting plan.	
GMTA	366	4.4	This standard defines the lifecycle for development and maintenance of health software, which includes Software in a Medical Device (SimD)and Software as a Medical Device (SaMD). Security aspect also needs to be considered in the AI Lifecycle management	Add IEC 81001-5-1:2021 Health software and health IT systems safety, effectiveness and security - Part 5-1: Security – Activities in the product life cycle	
GMTA	381-383	5	Typo in the text. Use the latest consolidated version (Edition 1.1) instead of only the amendment.	and the International Electrotechnical Commission (IEC) ~~62304 A1:152015~~ 62304:2006+A1:2015 Medical device software – Software life cycle processes	
GMTA	389	Figure 1	The purpose of the lines underneath the figure is unclear (e.g., the blue lines on either side of the words “Quality Management”, “Risk Management”, and “Human Oversight”).	Suggest updating the bottom lines to be dual-sided arrows to illustrate that they span the entire lifecycle.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	398	Footnote 4	IEC 62304:2006+A1:2015 is suitable as a foundational software life-cycle process standard for AI-enabled medical devices, particularly for software development, maintenance, configuration management, problem resolution, and software risk management. However, because it was not developed specifically for AI/ML (Artificial Intelligence/Machine Learning) systems, it should be supplemented with AI-specific life-cycle controls for data governance, model training and validation, model change control, monitoring for drift or degradation, cybersecurity, transparency, human oversight, and post-market performance monitoring.	Recommend editing the footnote as follows: “IEC 62304:2006+A1:2015”	
GMTA	405-408	5.1	While the current language specifically references “potential risks and controls,” reiterating the need for overall risk management planning and quality management would improve clarity. This addition emphasizes that product definition and planning activities should be supported by and integrated within, formal quality and risk management	Please revise as follows:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			processes rather than occurring as ad hoc or isolated activities.	This can be achieved through the application of robust quality and risk management practices and ~~with~~ a comprehensive product definition, which includes understanding the clinical objectives, existing standards of care, use environment, potential risks and controls, relevant clinical workflows and user needs and constraints as well as team resource planning.	
GMTA	411-413	5.1	While lines 458–459 of the draft document highlight the importance of understanding regulatory requirements in the jurisdictions where the device will be deployed in the specific context of post-market monitoring and surveillance, we believe this is an important consideration for the Planning and Design life cycle stage more broadly. We therefore propose including a similar statement earlier in Section 5.1 to reflect its broader relevance at this stage of development.	Please revise as follows: Furthermore, this section points out considerations for both the AI-enabled medical device and its model(s), which both have their own unique planning and design considerations. Regulatory requirements in the particular jurisdiction(s) in which the AI-enabled medical device will be deployed may help to scope what might be needed for this life cycle step.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	423-424	5.1	This revision clarifies that model selection often involves balancing interpretability with performance. While simpler models may be easier to interpret, they may not always provide the level of performance needed for a given intended use. Conversely, more complex models may offer improved performance but may require additional approaches to support explainability. Noting this tradeoff helps reflect how manufacturers typically consider multiple factors when selecting a model architecture.	Please revise as follows: Manufacturers may also want to consider opting for simplicity if explainability and interpretability is critical to the intended use environment. However, this decision should be balanced against model performance, as more complex models may achieve higher accuracy or robustness. Model selection should therefore consider the appropriate trade-off between interpretability and performance for the specific intended use.	
GMTA	430-431	5.1	This revision is intended to align with the text in Good Machine Learning Practice (GMLP) Principle 4 and reflect that training and test datasets should be appropriately independent. This wording allows flexibility in situations where absolute independence is not feasible, while still emphasizing that the datasets should be determined to be appropriate and fit for purpose, with a sufficient level of independence maintained.	Please revise as follows: Furthermore, it is important for manufacturers to consider quantity of data needed such that training datasets can be appropriately independent from test data sets.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	435-443	5.1	We recommend specifically referencing Section 5.5 here to clarify that, although the recommendations appear in the Deployment section, the considerations described should be addressed earlier in the product lifecycle.	Please revise as follows: For AI-enabled medical devices specifically, deployment may include site-specific localization or customization (with appropriate regulatory approval), and it is important that the infrastructure account for version control and resets or rollbacks in the event that recalls or issues occur. Section 5.5 of this document outlines additional planning considerations related to deployment that should be considered at this stage. It is beneficial for manufacturers to not only define requirements…"	
GMTA	452-459	5.1	The current wording may unintentionally imply an expectation for comprehensive or continuous real-world monitoring infrastructure at the earliest planning and design stage, regardless of device risk, deployment model, or manufacturer access to post-market data. For many manufacturers, access to ongoing clinical workflow data may be limited or outside the manufacturer's direct control. A more risk-based and proportionate framing would better align with least burdensome regulatory principles and acknowledge that post-market	Recommend replacing lines 452-259 as follows: "Post-Market Monitoring and Surveillance: Manufacturers should consider, as appropriate and proportionate to the device risk and intended use, preliminary planning for post-market monitoring and surveillance activities to support ongoing assessment of AI-enabled medical device performance in the intended use environment. The scope and nature of such activities may depend on factors such as deployment context, manufacturer access to post-market	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			monitoring approaches may vary depending on device risk, intended use, technical architecture, deployment environment, and applicable jurisdictional requirements.	data, and applicable jurisdictional requirements."	
GMTA	463	5.2 Data collection and management	The current section title and structure may not fully reflect the breadth of data-related lifecycle activities discussed in this section. In addition to data collection and management, activities such as data labeling/annotation, dataset curation, and dataset construction are also critical components of AI model development and validation and may warrant more explicit recognition within this section.	Consider revising the section title to better reflect the broader scope of activities discussed, for example: "Data Collection, Curation, and Management"	
GMTA	488-490	5.2	We recommend removing the specific example related to age and sex data, as it may unintentionally imply that such data are always required when evaluating model performance. Instead, a more general statement should be included indicating that relevant parameters should be considered when sourcing data for training, tuning, and testing.	Please revise as follows: ~~For example, for some use cases, age and sex data may not be needed to train the model but can help users understand performance across these demographics.~~ Manufacturers should consider the clinical and non-clinical parameters relevant to the model and its performance, and ensure that the data used for training, tuning, or testing capture the appropriate parameters.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	505-506	5.2	This addition is intended to clarify that synthetic data encompasses both modified versions of existing data (e.g., through augmentation or transformation techniques) and entirely new data generated through synthetic or simulation-based methods.	Please revise as follows: Manufacturers may also seek to use newly generated synthetic or simulated data to supplement actual, human data for model development and validation (e.g., data generated through statistical, generative AI, or simulation-based methods).	
GMTA	506-508	5.2	The revision clarifies that the use of synthetic or simulated data may also help reduce potential risks associated with collecting certain real-world data. While the original sentence referenced ethical and privacy considerations, synthetic and simulated data can also support safer exploration of scenarios that may be difficult, rare, or potentially risky to capture in clinical settings.	Please revise as follows: Use of synthetic or simulated data can help address ethical and privacy concerns, reduce potential risks associated with collecting real-world data, and accelerate data generation, particularly for underrepresented populations such as patients with rare diseases.	
GMTA	520-525	5.2	The revisions are intended to allow for a more flexible, risk-based approach. In practice, it may not be feasible or realistic for manufacturers to capture the “full spectrum” of data the model could encounter in clinical use. The revised wording (“sufficiently representative of the types of data the model is intended to encounter”) maintains the intent of ensuring that datasets are fit for purpose while allowing manufacturers to take a proportionate approach based on the	Recommend revising as follows: “It is important that data collected and used for training, tuning and validation is sufficiently representative of the intended use, target population, and risk profile ~~represents the full spectrum of data that the model is intended to encounter in clinical use in order~~ to ensure that the data is fit-for-purpose. ~~Furthermore, it is important that this data represents not only~~. Where relevant, manufacturers	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			device's intended use and available data. The revision also clarifies that factors that may influence model performance should be considered, while adding "where relevant" to avoid implying that all factors must be evaluated in every case. If interpreted strictly, the original language may unintentionally result in fewer devices for rare or emerging use cases where "full spectrum" representativeness is not feasible.	should consider representation across both clinical subtypes, such as demographics of the intended use population and disease conditions (both cases and controls) ~~but also~~ and nonclinical subtypes, such as data acquisition equipment and their parameters and collection sites, that are relevant to the device's intended purpose and may influence model performance."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	526-529	5.2	This revision is intended to introduce greater flexibility and reinforce a risk-based approach to dataset representativeness. While considering potential under-representation within the intended use population is important, it may not always be feasible or appropriate to ensure representation of all specific population groups in every dataset. The revised language clarifies that such considerations should be made where relevant to the device's intended use. We also recommend removing the term "priority populations," as this sentence specifically addresses representation within datasets used for training, tuning, and validation, for which "under-represented populations" is the more appropriate term. "Priority populations" is a broader public health or policy term that can refer to groups identified for health equity or programmatic focus regardless of their representation in the data. As a result, its inclusion may introduce confusion by implying considerations beyond dataset representativeness.	Please revise as follows: In addition to clinical and non-clinical subtypes, ~~it is important for manufacturers to consider how under-represented and priority populations within the intended use population (e.g., specific ethnic groups, Indigenous populations, or rural and remote communities) are reflected in the data used~~ manufacturers should consider relevant under-represented populations within the intended use population, as appropriate, when developing datasets for training, tuning, and validation. This may include consideration of populations that could experience different performance outcomes, (e.g., specific ethnic groups, Indigenous populations, or rural and remote communities), where relevant to the intended use.	
GMTA	545-549	5.2	Because bias cannot be completely eliminated, we recommend revising the language to state that the risk of biased outputs should be considered and appropriately addressed.	"When developing adaptive AI models that continue to learn after their initial deployment, it is important for manufacturers to ~~eliminate or reduce as far as possible~~ consider and appropriately address the risk of biased	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by ******)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				outputs influencing input for future operations..."	
GMTA	567-574	5.2	Full lifecycle traceability expectations exceed ISO 13485 and MDR documentation requirements and may be infeasible for federated or real-world datasets.	Recommend revising as follows: "Building upon the documentation recommendations mentioned for data cleaning, it is important for manufacturers to establish ~~comprehensive data lineage and provenance tracking~~ lineage and provenance tracking, which should be maintained to the extent necessary to support safety, performance, and regulatory compliance.'	
GMTA	587		IEC PAS 63621:2025 provides a framework for the data lifecycle processes for management of data used to train, test, or validate an AI model that is part of a medical device.	Add IEC PAS 63621:2026	
GMTA	606-612	5.3	The current wording, together with the subsequent example describing selection of one model architecture over another, may imply that detailed justification of model architecture choices is expected for regulatory review. While documentation supporting intended use, performance, and risk	Recommend replacing both sentences in lines 606-612 (i.e., the sentence that begins with "Justification" and the following sentence that begins with "For instance") with the following:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			management is important, consideration should be given to avoiding language that may be interpreted as creating unnecessary documentation expectations regarding internal model development decisions that may not be directly relevant to demonstrating device safety, effectiveness, or clinical performance. This may be particularly important for a high-level technical framework document intended to describe general considerations rather than specific regulatory submission expectations.	Documentation of model design considerations relevant to the intended use, performance, and risk profile of the device may support transparency, risk management, and regulatory communication where appropriate.	
GMTA	616-622	5.3	The current text implies a binary characterization of models as "explainable" or "unexplainable." The suggested revision is intended to convey that explainability exists along a spectrum, with different models demonstrating varying degrees of explainability.	Recommend revising as follows: "Models that have a higher degree of explainability ~~are explainable~~ can enable users to understand their underlying reasoning and limitations and can also support the ability of the manufacturer to create an appropriate evaluation plan. …As such, manufacturers can consider the trade-offs between the potential risks and benefits of ~~unexplainable~~ models with more limited explainability in the context of the clinical use..."	
GMTA	622-623	5.3	The proposed revision is intended to better align the discussion with a risk-based regulatory approach.	Please revise as follows:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			Framing the relationship between explainability and clinical evaluation as a direct expectation may unintentionally imply that lower model explainability inherently requires increased clinical evidence. In practice, the risks associated with limited explainability can vary depending on the device's intended use, clinical context, and the availability of other risk controls. This approach provides greater flexibility while still ensuring that any risks associated with reduced explainability are appropriately evaluated and controlled.	There may also be situations where models with lower levels of explainability introduce additional uncertainty or risk, which may require further evaluation to support safe and effective use in clinical practice. In some cases, generating additional clinical evidence may be an appropriate way to address these uncertainties and support ~~a higher expected level of clinical evaluation for less explainable models to gain~~ clinician trust for real-world adoption .	
GMTA	646-648	5.3	The current wording may be interpreted as creating broad documentation expectations regarding internal feature preprocessing and feature engineering decisions, including expectations for detailed justification of iterative development activities that may not be directly relevant to demonstrating device safety, effectiveness, or clinical performance. Consider using more focused and risk-based language consistent with the high-level nature	Original text: "Finally, regardless of the approach to feature preprocessing used, it is important that manufacturers ensure the decision-making process of choosing the approach is transparent and justified, which can inform evidence for future regulatory inquiries and transparency for users."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			of this technical framework document.	Suggested edit: "Manufacturers may consider documenting feature preprocessing approaches relevant to the intended use, performance, explainability, or risk profile of the device, where appropriate, to support transparency, risk management, and regulatory communication."	
GMTA	649-655	5.3	We appreciate the inclusion of the discussion for "loss function" in this document, however its inclusion in a paragraph titled "Selection of Evaluation Metrics" appears misaligned. We recommend this paragraph remain focused on evaluation metrics used to assess evidence of model performance. The discussion on "loss function" may be better suited elsewhere in the document, such as in the context of model training/tuning discussion.	Recommend moving the sentences in this paragraph that discuss "loss function" to a different location in the document that discusses model training/tuning.	
GMTA	667-678	5.3	The heading of Section 5.3 is "Model building and Tuning, however the paragraph titled "Leveraging general-purpose and off-the-shelf models" discusses use of ~~3~~$^{\text{rd}}$ third-party general purpose models. In general, manufacturers using a ~~3~~$^{\text{rd}}$ third-party general purpose model are not building that model. Therefore, we think this paragraph is better suited to a different section in the document, particularly because the considerations discussed	Recommend moving the paragraph on "leveraging general-purpose and off-the-shelf models" to section 5.1. Planning and Design. More specifically, we recommend moving this content to the paragraph titled "Model Selection and Anticipating Risks" and adding the content in lines 667-678 to the end of line 424. Additionally, we recommend revising lines 672-674 as follows:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			in this paragraph are important for manufacturers to consider during the planning phase so they can better prepare for the types of documentations/requirements that are needed from the third-party. Suggested revisions to lines 672-674,: This section could be interpreted as requiring manufacturers to have access to all of this information about third-party models, which may not always be feasible and may not be necessary depending on how the model is used. The proposed revisions are intended to better reflect this reality. The subsequent paragraphs acknowledge that information gaps may exist and that manufacturers should evaluate and consider these gaps within their risk management processes, and the proposed revision is intended to align with that concept. Additionally, we proposed removing the term "model longevity" because its meaning and how it would inform a manufacturer's assessment of the suitability of a third-party model is unclear.	"While manufacturers may not always have access to all information about third-party models, they should consider, where available and appropriate, additional factors such as ~~Other key considerations include assessing~~ the third-party supplier credibility, model provenance, ~~model's longevity~~, and life cycle management processes, including the model's version history, update cadence, and support commitments, as appropriate and feasible. Additionally, we recommend adding the following to the end of the paragraph: "Manufacturers should assess third-party models based on information reasonably available and apply risk-based supplier controls commensurate with their level of control"	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	725, 727, 747, & 751	5.4	The term “external validation” may be interpreted inconsistently across jurisdictions and stakeholders. For example, “external validation” may be interpreted as specifically referring to use of data from *sites* that were not part of the training dataset. This is distinct from the recommendations in IMDRF’s GMLP document (N88) which emphasize “independent and representative data” for validation. The suggested revisions are intended to promote consistency with IMDRF’s GMLP document (N88).	Recommend replacing all instances of “external validation” with “validation of the model using independent and representative data.” (i.e., lines 725, 727, 751) Suggested revision of the paragraph heading for line 747: “Validation of the Model Using Independent Dataset”	
GMTA	741-744	5.4	Providing V&V outcomes to end uses may be too detailed and may confuse end users more than be beneficial. Additionally, this level of detail may compromise security or intellectual property. Public-facing documentation (e.g., device labeling) should contain sufficient documentation to support transparency. .	Recommend revising as follows: “This recordkeeping can help develop comprehensive and transparent documentation for the AI-enabled medical device and its model(s). ~~so that clinicians and patients can understand the V&V outcomes and make informed choices regarding AI-enabled medical device deployment and use.~~ See Section 6 for more information on transparency and labeling..”	
GMTA	750	5.4.1	Recommend the document clarify that the validation data should be separate from both training and *tuning* activities.	“This involves, for example, validating the model on a dataset that is separate from the data used to train and tune the	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				model ~~training data~~, and assessing the performance…"	
GMTA	763-765	5.4.1	We recommend removing the reference to "red teaming" from this section. While adversarial robustness evaluation and stress testing may be relevant considerations for certain AI-enabled medical devices, particularly higher-risk or generative AI systems, "red teaming" is currently an evolving and inconsistently defined concept that is not broadly established within medical device regulatory validation frameworks or harmonized standards. The term is also more commonly associated with cybersecurity and adversarial security assessment activities rather than established medical device verification and validation methodologies. Inclusion of the term in a high-level IMDRF lifecycle framework may therefore introduce ambiguity regarding regulatory expectations, scope, applicability, and required methodologies, and may unintentionally imply resource-intensive validation expectations beyond established risk-based software and cybersecurity testing practices.	Recommend revising as follows: "~~Additionally, red teaming, a process where a team intentionally probes the model for weaknesses and exploits potential vulnerabilities, is a valuable method for identifying issues and validating implemented risk controls and failure/error handling mechanisms.~~"	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	772	5.4.2	Clarifying this distinction may support risk-proportionate evaluation and appropriate clinical validation strategies.	Recommend adding the following to the end of line 772: "Semi-automated AI functionalities that are explicitly reviewed, edited and validated by trained operators prior to downstream clinical use, exemplify human oversight and may present lower risk profiles compared to fully automated outputs."	
GMTA	780-782	5.4.2	We propose using the phrase "clinically meaningful," which aligns with IMDRF N41, rather than "clinically useful" which is not as appropriate or clear.	Recommend revising as follows: "Diagnostic medical device evaluations may therefore focus specifically on determining whether the output of the model is accurate and clinically meaningful ~~useful~~ based on metrics like positive and negative percent agreement or predictive value."	
GMTA	804-806	5.4.2	The added text is intended to clarify that the usability and human factors considerations discussed here are not necessarily unique to AI-enabled medical devices.	Please revise as follows: Usability and Human Factors: As with all medical devices, to ~~To~~ ensure that the AI-enabled medical device is effective and usable in its intended environment, it is important for manufacturers to consider what is needed for human factors and usability validation.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	814-816	5.4.2	"Unintended consequences" is not a term frequently used in usability. We recommend replacing with "risk" which is more typically discussed in the context of use errors.	Recommend revising as follows: "It is also important for manufacturers to consider reasonably foreseeable misuse and whether any potential user errors could result in risk to the patient or user ~~lead to unintended consequences~~."	
GMTA	826-829	5.4.2	The term "usefulness" is subjective and not clearly defined in this context. We recommend more precise terminology, such as "clinical significance" or "clinical impact" to improve clarity and better align the consideration with the device's intended use, risk, and autonomy of the device output.	Recommend revising as follows: "the clinical significance ~~usefulness, importance,~~ and degree of autonomy of the AI-enabled medical device outputs..."	
GMTA	833-835	5.4.2	The current wording may blur the distinction between information intended for users and information intended for regulators. It is recommended separating this concept into two statements to distinguish between the justification intended for users and the justification intended for regulators. Communications directed to users should focus on clearly conveying the benefits and risks of the AI-enabled medical device to support appropriate	Please revise as follows: Manufacturers may benefit from crafting a clear justification to ~~both~~ articulate to users on the benefits and risks of an AI-enabled medical device. In addition~~, and~~ as part of regulatory approval process, manufacturers may ~~benefit from~~ provide a justification that describes the device's benefits, risks, and associated risk mitigation actions, and explains ~~on~~ why the V&V activities and outcomes are relevant to the device's intended use.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			understanding and use. In contrast, information provided as part of the regulatory approval process should more comprehensively address the device's benefits, risks, and relevant risk mitigation measures, as well as explain how the V&V activities and outcomes support the device's intended use. This distinction improves clarity regarding the different purposes and audiences for these justifications.		
	877-880	5.5	This revision introduces greater flexibility by clarifying that such mechanisms should be considered where relevant to the device and its intended use. As written, the original text could be interpreted as suggesting that all AI-enabled medical devices require infrastructure to actively monitor input data characteristics, data drift, or performance degradation across sites. In practice, the need for these activities depends on the specific characteristics of the device, its deployment environment, and its risk profile.	Please revise as follows: Additionally, where relevant to the device and its intended use, manufacturers may also consider mechanisms to assess ~~it is important for the infrastructure to have mechanisms in place to determine~~ whether the quality and quantity of inputs to and data requirements of the AI-enabled medical device meet specifications and account for variation across sites. This may help support the identification of issues such as ~~enabling the ability to monitor for~~ data drift and performance degradation, potentially through Human Oversight features (as described in Section 4.3).	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	888-897	5.5	The recommended revisions are intended to reinforce the statement that "deployment activities will vary in complexity" depending on the risk profile of the specific device. In the Deployment Verification and Validation section (lines 888-897), we propose rearranging the paragraph to first emphasize that "depending on the risk of the device, deployment V&V activities can vary in complexity and duration," and then provide verification of training, use of retrospective data, and longitudinal activities as examples of approaches that may be appropriate. In the Deployment Approach section (lines 898-905), we propose edits to clarify that not all AI-enabled devices require a phased deployment, but that this approach may be beneficial for certain devices depending on their risk profile.	"~~This can include ensuring that training for site staff and healthcare professionals was adequate and appropriate for the environment and users.~~ Depending on the risk of the device, deployment of V&V activities can range in complexity and timespan. For example, activities may include ensuring that training for site staff and healthcare professionals was adequate and appropriate for the environment and users. For some devices, it may be necessary to utilize ~~utilizing~~ a combination of site-specific retrospective data and control data ~~may be sufficient~~ to verify that the deployment process was successful." Recommend revising lines 898-901 as follows: *"Deployment Approach:* Defining the implementation approach in deployment plans is a critical step, and the specific deployment approach will depend on the specific device and its associated risk. For example, some devices may benefit from ~~whether that is~~ deployment in phases, such as a preview release followed by a final release, or through phased release across multiple sites. A phased release across multiple sites may be beneficial for certain AI-enabled	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				medical devices because it enables manufacturers to..."	
	958-961	5.6	The revision clarifies that the listed capabilities are examples that may be implemented depending on the device and deployment context, rather than implying that all elements are expected in every case. By introducing "may include, as appropriate," the text better reflects a risk-based and proportionate approach while maintaining the recommendation that manufacturers consider mechanisms to support monitoring, incident response, and troubleshooting where relevant.	Please revise as follows: As discussed in Section 5.5 above, the deployed AI-enabled medical device's infrastructure facilitates the collection of data needed to fulfill monitoring requirements, and may include, as appropriate~~, can incorporate~~ data logging and performance tracking capabilities, ~~as well as~~ anomaly detection algorithms, incident response procedures and troubleshooting workflows.	
	961-965	5.6	The revision introduces more flexible language to recognize that manufacturers may not always have remote access to deployment data due to technical limitations, institutional policies, or data privacy requirements.	Please revise as follows: For AI-enabled medical devices, it may be worthwhile for the infrastructure to have both local monitoring capabilities and, where feasible and appropriate, ~~as well as~~ the ability to contribute accumulated monitoring data to central repositories, ensuring a comprehensive and uninterrupted record of the AI-enabled medical device's operational performance across deployments for the manufacturer.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
	965-975	5.6	This revision introduces greater flexibility and a more product-specific, risk-based framing by clarifying that the appropriateness and feasibility of logging capabilities will depend on the device's intended use, technical design, deployment environment, and risk profile. This helps avoid implying that specific monitoring technologies or detailed logging of model predictions or explainability artifacts are expected for all AI-enabled medical devices.	Please revise as follows: Where possible and appropriate, the process may integrate advanced technologies that enable continual assessment of the AI-enabled medical device's accuracy and responsiveness, supporting rapid intervention, adaptive recalibration, and maintenance of the AI-enabled medical device's reliability and effectiveness across its operational lifecycle. In addition, logging and auditing can help support ~~are critical components of an AI-enabled medical device's~~ operational monitoring activities and maintenance. For AI-enabled devices specifically, depending on the device and its design, it may be beneficial for the logging mechanisms to capture detailed logs and audit trails (as allowed within jurisdictional regulations) of model predictions and any explainability elements (if available) such as heatmaps or top predictive features, data inputs, model outputs, and metadata such as timestamps, user interactions, model versions, and environmental context for traceability, accountability, and compliance purposes.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	970-975	5.6	Storing detailed input data, outputs, prompts, responses, heatmaps, and user interactions can create significant privacy, cybersecurity, storage, consent, and operational burdens. In some clinical contexts, such logging may be restricted, undesirable, or clinically impractical. Excessive logging recommendations could inhibit use in privacy-sensitive areas such as mental health, reproductive health, pediatrics, home monitoring, or emergency settings. It may also increase breach impact. The suggested edits are intended to support a more flexible approach that supports risk-based, privacy-preserving traceability. Logging personal data must be done in compliance with the respective jurisdiction. The term “as allowed” could be misunderstood as “it is allowed”, rather than “if allowed”.	Recommend revising as follows: “For AI-enabled devices specifically, it may be beneficial for the logging mechanisms to capture detailed logs and audit trails (as ~~allowed~~ appropriate and in compliance within jurisdictional regulations and data protection legislation) of model predictions and any explainability elements (if available). Aggregated metrics, local-only logs, sampling, de-identification, event-triggered logging, or privacy-preserving monitoring may be appropriate. ~~such as heatmaps or top predictive features, data inputs, model outputs, and metadata such as timestamps, user interactions, model versions, and environmental context for traceability, accountability, and compliance purposes.~~ Monitoring of user interactions should be limited to anonymized or aggregated data unless collection of identifiable information is strictly necessary for safety purposes.”	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	976-1003 & throughout	5.6	The “Performance Degradation and Drift Detection” and “Alerting and Reporting” sections discuss monitoring, however it is unclear which recommendations are applicable to model-level monitoring, versus device-level monitoring. For example, it is unclear if the discussion on drift detection and reporting are intended to apply at the model level or device level. Lines 116–124 explicitly state that “model” and “device” are not interchangeable terms. As such, we recommend the document more clearly indicate which considerations are applicable to the model level and which apply at the device level.	Recommend revising this section to more clearly indicate which considerations apply at the model level and which apply at the device level. Additionally, we recommend ensuring the distinction between recommendations for model level and device level are clear and consistent throughout the document.	
	987-988	5.6	The extent to which generative AI may “rapidly and unpredictably deviate from its original training objectives” depends on the specific model design and the degree to which the system is stochastic versus deterministic. Not all generative AI systems will necessarily present this risk, and the language should be revised to reflect this variability.	Please revise as follows: Monitoring drift in generative AI-enabled medical devices may be especially important ~~is especially critical~~ as some generative AI ~~these~~ systems can rapidly and unpredictably deviate from their original training objectives.	
GMTA	995	5.6	The terms used here are unclear.	Recommend revising as follows:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				"For AI-enabled medical devices specifically, it is essential to follow best practices for mapping data sources and controlling data suppliers and labelers as part of QMS processes. This includes appropriate data provenance, dataset versioning, and chain-of-custody controls to support traceability and integrity of training, validation, and post-market data over time. This can help control…"	
GMTA	996-1002	5.6	The suggested revision is intended to clarify that ongoing monitoring and responsive alerting systems may be particularly relevant for certain types of AI-enabled medical devices, rather than implying that such capabilities are necessary for all devices. In practice, the need for ongoing monitoring depends on the characteristics of the model and its sensitivity to real-world inputs. For example, adaptive or generative AI systems may be more susceptible to changes in input data or evolving behavior over time, which may warrant more active monitoring approaches, as compared to more traditional locked AI.	"Alerting and Reporting: For certain AI-enabled devices that may be more sensitive to real-world inputs (e.g., certain adaptive or generative AI systems), ~~T~~the use of ~~continuous~~ monitoring and responsive alerting and reporting systems may be needed ~~helps~~ to ensure reliability, stability and effectiveness of AI-enabled medical devices at deployed sites. For example, use of thresholds for critical performance metrics, automated notifications and alerts are some of the ways manufacturers can ~~continuously~~ monitor their deployed AI-enabled medical devices and promptly address risks. In addition to individual incidents, it may be ~~is~~ important to monitor for changes in data distribution and performance over time, e.g., drift as described above."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	1021-1024	5.6	Section 5.6 provides a comprehensive overview of considerations related to operations and monitoring of AI-enabled medical devices; however, several portions of the section may benefit from clearer distinction between high-level lifecycle considerations and operational expectations that may not be feasible, proportionate to risk of the device, or within the manufacturer's visibility or control. In particular, references to continuous monitoring of user interactions, workflow behaviors, prompt-response histories, adaptive learning trajectories, and emergent behaviors may unintentionally imply extensive post-market observability and governance infrastructure expectations that may not apply universally across AI-enabled medical devices or deployment models. We recommend the document emphasize that monitoring approaches should remain risk-based, proportionate to intended use and device autonomy, and dependent upon factors such as deployment architecture, manufacturer access to post-market data, site-specific integrations, and applicable jurisdictional requirements.	Recommend replacing the sentence that begins "To effectively capture…" with: "Where feasible and appropriate to the device risk, intended use, deployment architecture, and manufacturer visibility into post-market use, manufacturers may consider risk-based monitoring approaches for advanced AI models."	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	1044-1047	5.7	The distinction between “Operations and Monitoring” activities in Section 5.6 and “Real-World Performance Evaluation” activities in Section 5.7 may benefit from further clarification. The example describing clinicians overriding AI-enabled medical device outputs appears to rely on workflow and user interaction information that may not typically be available through operational monitoring infrastructure alone and may depend on site-specific workflows, local integrations, or additional post-market evaluation activities. As currently written, the example may unintentionally imply continuous manufacturer monitoring of clinician behavior or workflow interactions that may not be feasible, proportionate to risk of the device, or within the manufacturer’s visibility or control.	Recommend replacing the sentence that begins, “For example, if clinicians are frequently choosing to override aspects of the AI-enabled medical device’s outputs...” with: “For example, operational monitoring infrastructure may indicate that the model is performing within expected technical parameters, while separate real-world performance evaluation activities or user feedback may identify workflow, usability, or clinical utility considerations, such as frequent user overrides of AI-generated outputs.”	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	1060-1063	5.7	The revision introduces more flexible, risk-based language to better reflect that subgroup analyses may not be feasible or relevant for all AI-enabled medical devices. Replacing “it is important” with “manufacturers may also consider,” and adding “relevant to the intended use,” clarifies that evaluating performance across subgroups should be done where it is meaningful and practical based on the device’s intended use, available data, and deployment context. This helps avoid implying that subgroup analyses are expected in all cases while still encouraging manufacturers to consider potential differential performance where it could impact safety, effectiveness, or equity.	Recommend revising as follows: “Where feasible and relevant to the intended use, manufacturers may consider evaluating ~~it is important for~~ device performance indicators across ~~to be defined not only in aggregate but also for~~ clinically relevant subgroups, such as demographics,...”	
GMTA	1099-1102	5.8	It is not clear why this sentence would be specific to AI-enabled devices. Therefore, it is suggested removing the intro clause.	Please revise as follows: ~~For AI-enabled medical devices specifically, it~~ It is important to communicate sunsetting plans to stakeholders, both internal teams and external users and sites, with clear timelines and reasons for the decision, as well as alternative solutions or support to help them transition, as applicable.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
GMTA	1123	6	Transparency and explainability needs may differ significantly depending on the intended user, clinical workflow, and potential impact to patient health. Clarifying that transparency and labeling should follow a risk-based approach would improve consistency in implementation and help ensure that information provided to users is appropriate to the clinical context, intended reliance on AI outputs, and associated level of risk.	As with all medical devices, AI-enabled medical devices should be accompanied by information that is truthful, accurate, and sufficient to support their use by the intended user and to ensure that they are safe and effective for their intended use or purpose. Transparency, explainability, and labeling expectations may vary depending on the intended user (e.g., clinician, healthcare provider, patient, or lay user), clinical context, and the level of risk associated with the device and its outputs. A risk-based approach should be utilized when determining the type, detail, and presentation of transparency and labeling information.	
GMTA	1141-1142	6	Revisions are proposed to clarify that this sentence is referring specifically to when electronic labeling is provided in lieu of paper labeling rather than in addition to paper labeling, that electronic labeling must be permitted within the relevant jurisdiction, and to note that e-labeling can support more timely updates to labeling.	Manufacturers may also wish to consider electronic labeling in lieu of paper labeling, where permitted by the relevant authority, which can support increased availability, utility, interactivity, ~~and~~ accessibility, and timely updating of labeling throughout the medical device lifecycle.	
GMTA	1145-1147	6	It is unclear why it would be relevant to users whether a change required regulatory approval or was	Recommend revising as follows:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			implemented under an authorized predetermined change control plan (PCCP). This information does not appear necessary to support the safe and effective use of the device. Therefore, we propose removing this sentence.	"For models that are expected to undergo discrete updates, tailored transparency information can be provided at the time of each update~~,~~ ~~including whether the update required regulatory approval, for example as part of PCCP~~."	
GMTA	1185	Appendix B	Accuracy for GenAI may be evaluated as more than just predictions.	It is suggested to change the word "predictions" to "outputs" to ensure comprehensive coverage of technologies.	
	1215	Appendix C	This edit is proposed to reflect that only subgroup performance that is relevant to the specific device should be evaluated and shared in the labeling.	Please revise as follows: • measured performance and relevant subgroup performance;	
	1217	Appendix C	This edit is proposed to clarify that only a high-level summary of the testing datasets should be included in the labeling. Highly detailed descriptions may risk disclosing proprietary information and could overwhelm the reader without meaningfully contributing to the safe and effective use of the device.	Please revise as follows: • pre-market testing conditions and a high-level summary of testing datasets;	
	1219	Appendix C	Explicit disclosure regarding the use of AI within a medical device would strengthen transparency and support safe and effective use by helping	For AI-enabled medical devices, information that may be needed to support safe and effective use includes, but is not limited to:	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
			users understand the role, limitations, and intended reliance on AI-generated outputs. Clarifying whether AI outputs are advisory, diagnostic, therapeutic, autonomous, or supportive in nature, along with the expected level of human oversight, would improve user understanding, support appropriate clinical decision-making, and better align transparency expectations across AI-enabled medical devices.	• the intended use or purpose and intended user; • any professional or other qualifications that are necessary to use the device safely and effectively; • clear information on compatibility requirements for inputs (if applicable); • how the device should be used, including any prerequisite steps or necessary preparations; • the populations in which the device can be used and the conditions for which it can be used (including known limitations); • measured performance and subgroup performance; • when and for how long the device should be used; • pre-market testing conditions and datasets; • how to interpret the output of the device, known limitations, model description or characteristics.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by ******)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial

Name/Organization	Line number	Section	Comments	Proposed change	Resolution
				clear disclosure that AI is used in the device, including the role of the AI model, the type of output generated, whether the output is advisory, diagnostic, therapeutic, autonomous, or used to support another device function, and the expected level of human oversight.	

1 **MB** = Member body / **NC** = National Committee (enter the ISO 3166 two-letter country code, e.g. CN for China; comments from the ISO/CS editing unit are identified by **)
2 **Type of comment:** **ge** = general **te** = technical **ed** = editorial