

## Comment on AI Risk Management Framework: Second Draft

August 27, 2022

Part#	Section#	Section Title	Page#	Content for Comment	Comment
1	2	Audience	5	Note under Figure 1: “Risk Management should be ....., starting with the plan & design function in the application context”.	Risk analysis should be conducted throughout the AI design and development stages which include design inputs and outputs, and design verifications and validations prior to commercial releases (refer to page 7, 17 of “Content of Premarket Submissions for Device Software Functions”, a draft guidance issued by U.S. FDA on 11/4/21 for incorporating risk management in SRS (design input) and SDS (design output); and refer to 21 CFR Part 820 Section 30, Article (g) for risk analysis requirement for design validations).
1	2	Audience	6	Figure 2 “Use or impacted by” “Activities”: “...seek mitigation of impacts...”	The potential impacts should be identified and mitigated during the AI design and development stages (see the comments above).

<b>Part#</b>	<b>Section#</b>	<b>Section Title</b>	<b>Page#</b>	<b>Content for Comment</b>	<b>Comment</b>
1	2	Audience	6	Paragraph 1 under Figure 2: "TEVV allows for both mid-course remediation and post hoc risk management and mitigation".	The mid-course remediations and post hoc risk mitigations should be treated as design changes during and post designs if software changes are involved. The software changes are required to be validated (refer to 21 CFR Part 820 Section 30, Article (g) & (i) for requirements for software change validations).
1	3	Framing Risk	9	Paragraph 2 indicates that measuring risks at different stage of AI lifecycle yield different results, and risks may be latent and may increase as AI system evolve.	Risk analysis should be conducted during AI system design and development stages. Risks including any latent risks should be identified and mitigated prior to commercial releases (see the comments above).
1	3	Framing Risk	9	Paragraph 3: "AI risks measured in a laboratory or a controlled environment may differ from risks that emerge in operational setting or the real world".	AI risk analysis shall be conducted both in a lab or a controlled environment as well as in operational settings or the real world (refer to 21 CFR, Part 820, Section 30, Article (g)).

<b>Part#</b>	<b>Section#</b>	<b>Section Title</b>	<b>Page#</b>	<b>Content for Comment</b>	<b>Comment</b>
1	3	Framing Risk	9	Paragraph 2 under Sub Section "Risk Tolerance": "In the absence of risk tolerances prescribed by existing law, regulation, or norms, the AI RMF equips organizations to define reasonable risk tolerance, manage those risks, and document their risk management process".	Organizations for premarket submissions involving AI should also refer to "Content of Premarket Submissions for Device Software Functions", a draft guidance issued by U.S. FDA on 11/4/21 (see the attached file).
1	3	Framing Risk	10	Paragraph 1 under Sub Section "Risk Perspectives": "Attempting to eliminate risk entirely can be counterproductive in practice-because incidents and failures cannot be eliminated".	Significant AI risks that cause incidents and failures and are harmful to humans are required to be eliminated via design changes, design change verifications and validations for risk mitigations prior to commercial releases (refer to 21CFR, Part 820, Section 30).
1	3	AI Risk and Trustworthiness	12	Box text under "Human Factor": "Biases can be induced by AI actors across the AI lifecycle via assumptions, expectations, and decisions during the modeling tasks"	During the modeling tasks, decisions in any way changes the original AI software designs need to be evaluated for design change verifications and validations (refer to 21CFR, Part 820, Section 30, Article (i)).

<b>Part#</b>	<b>Section#</b>	<b>Section Title</b>	<b>Page#</b>	<b>Content for Comment</b>	<b>Comment</b>
1	3	AI Risk and Trustworthiness	12	Box text under "Human Factor": "Data about the frequency and rationale with which humans overrule AI system suggestions in deployed systems can be useful to collect and analyze".	The data collected should be provided to AI system developers for evaluations of the need for design change, and design change verifications and validations.
1	3	AI Risk and Trustworthiness	13	Paragraph 1 under Sub Section "Valid and Reliable": "Deployment of AI systems which are inaccurate, unreliable, or non-generalizable to data beyond their training data creates and increases AI risks and reduce trustworthiness".	AI system shall be verified and validated for its intended applications beyond training data with risks identified and mitigated (refer to 21 CFR, Part 820, Section 30, Article (g)).
1	3	AI Risk and Trustworthiness	13	Paragraph 4 under Sub Section "Valid and Reliable": "Robustness does not only require that the system perform exactly as it does under expected use, but also that it should perform in ways that minimize potential harms to people if it is operating in an unexpected environment".	AI system shall be verified or validated under actual use conditions and in actual use environment to eliminate potential harms to humans prior to commercial releases (refer to 21 CFR, Part 820, Section 30, Article (g)).

Part#	Section#	Section Title	Page #	Content for Comment	Comment
1	3	AI Risk and Trustworthiness	14	Paragraph 2 under Sub Section "Fair-and Bias Is Managed" indicates that there are three categories of AI bias: 1) systemic; 2) computational; and 3) human.	The systemic and computational AI biases shall be addressed during AI system design and development stages. Human bias, causing design changes in any ways to the original AI system, is required to be addressed per 21CFR820.30(g) regulation requirements for design changes and design verifications, and validations.
Appendix B	Appendix B	How AI Risks Differ from Traditional Software Risks	30-31	Examples of the differences: 1) "datasets used to train AI systems may become detached from their original and intended context, or may become stale or outdated relative to deployment context"; 2) AI system and complexity (billions and trillions decision points); and 3) risks associated with 3rd-Party technologies where AI systems may be trained for decision-making outside an organization's security controls or trained in one domain and then "fine-tuned" for another.	The list of examples indicate that the AI systems released for commercial uses have not been controlled via adequate AI system validations. Please refer to applicable regulations in 21CFR & guidances for submissions of premarket-applications for commercial AI systems used in healthcare industries.