GOOD MACHINE LEARNING PRACTICE FOR THE MEDICAL DEVICE TOTAL PRODUCT LIFE CYCLE

Shawn Forrest

Digital Health Specialist, CDRH Digital Health Center of Excellence

Center for Devices & Radiological Health (CDRH), US FDA

www.fda.gov/digitalhealth

Outline

- Digital Health and the CDRH DH Center of Excellence
- Artificial Intelligence/ Machine Learning (AI/ML)-Enabled Medical Devices
- Good Machine Learning Practice
- Observations from the Pre-Cert Pilot

Outline

- **Digital Health and the CDRH DH Center of Excellence**
- Artificial Intelligence/ Machine Learning (AI/ML)-Enabled Medical Devices
- Good Machine Learning Practice
- Observations from the Pre-Cert Pilot

Patients are at the Heart of What We Do

CDRH Vision

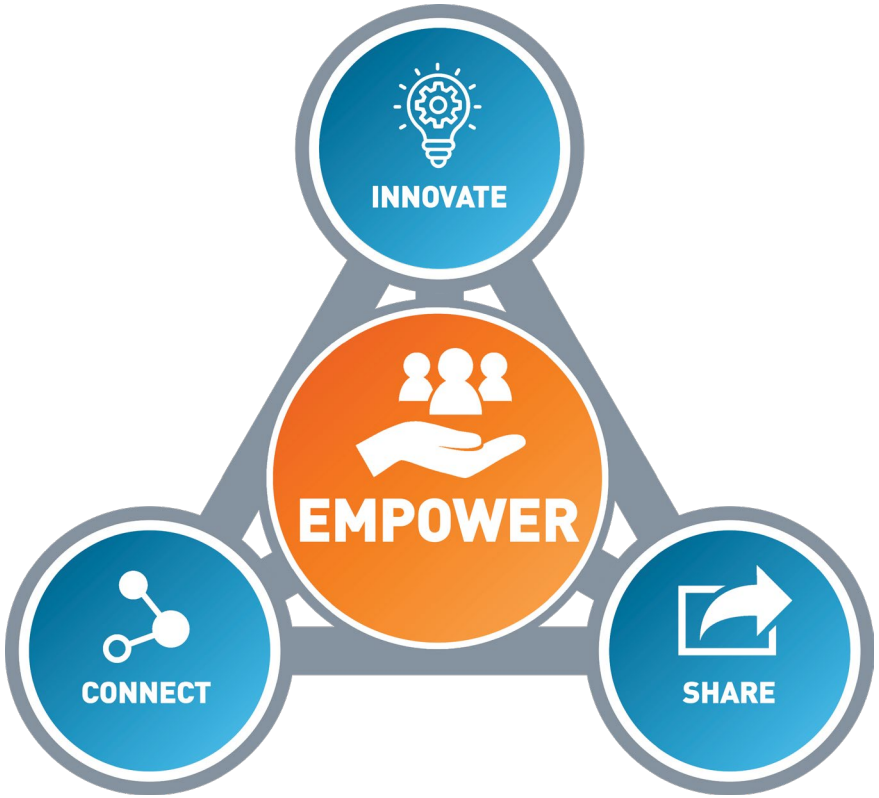
Patients in the U.S. have access to high-quality, safe, and effective medical devices of public health importance first in the world

CDRH's Digital Health Center of Excellence

Empowering All to Advance Healthcare

Our goal: Empower stakeholders to advance public health by fostering responsible and high-quality digital health innovation.

- **Connect and build** partnerships to accelerate digital health advancements.
- **Share knowledge** to increase awareness and understanding, drive synergy, and advance best practices.
- **Innovate regulatory approaches** to provide efficient and least burdensome oversight while meeting the FDA standards for safe and effective products.



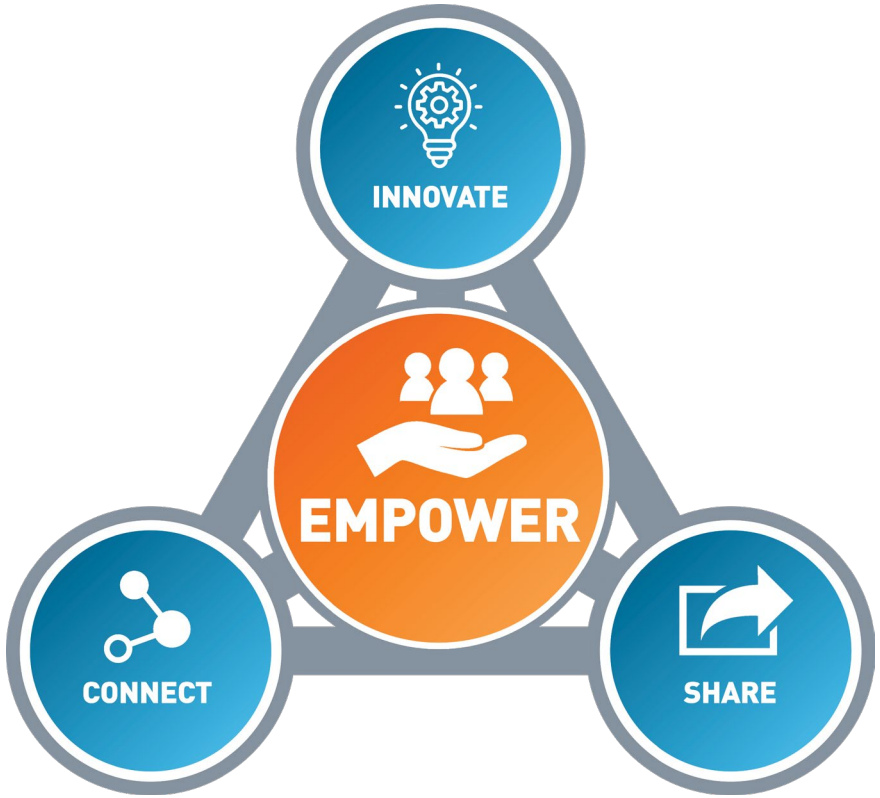
CDRH's Digital Health Center of Excellence provides world-class digital health expertise and policy direction

Our team brings extensive experience:

105+ YEARS	90+ YEARS	50+ YEARS	45+ YEARS	40+ YEARS	40+ YEARS	15+ YEARS
FDA	DIGITAL HEALTH INDUSTRY	ACADEMIA	OTHER INDUSTRY (NON-DIGITAL HEALTH)	MEDICAL	NONPROFIT	OTHER GOVERNMENT (NON-FDA)

Current as of June 2022

We are setting the stage for the advancement of digital health to help protect and promote public health



Select Accomplishments

29	14	80+	3400+	5000+
Public Documents	Guidance Documents since FY18	Experts in the Network of DH Experts	Inquiries addressed since FY18	Medical device submissions with DH considerations since FY18

Notable Authorizations

- Substance use disorder
- Opioid use disorder
- Colon Cancer
- Chronic Pain Reduction
- Capture Cardiac Ultrasounds
- Diabetes Management

New Regulatory Frameworks

- Released AI/ML Action Plan
- Piloted a software pre-certification program
- Published list of currently marketed AI/ML-enabled medical devices
- Jointly published list of 10 Good Machine Learning Practice (GMLP) Principles with the UK and Canada
- Workshop on Transparency of AI/ML-enabled medical devices

CDRH's Digital Health Center of Excellence (DHCoE) launched in September 2020.

Our work extends across the FDA and beyond the Agency, both nationally and internationally

External to FDA

- ✓ Provide clarity on regulation
- ✓ Advance international harmonization on device regulatory policy
- ✓ Facilitate and build strategic partnerships
- ✓ Communicate FDA research interests
- ✓ Advance digital health device international standards

FDA-Wide

- ✓ Provide DH expertise across the Agency
- ✓ Offer training opportunities for FDA staff
- ✓ Disseminate shared resources
- ✓ Foster collaboration across FDA in common interest areas
- ✓ Facilitate synergies in regulatory science research in DH

Medical Device Focus

- ✓ Set/lead strategic direction and launch initiatives in DH
- ✓ Establish and promote best practices
- ✓ Enable efficient, transparent, and predictable product review with consistent evaluation quality
- ✓ Build new capacity to oversee and leverage DH technologies including shared resources
- ✓ Coordinate the development of cross cutting DH policies

Internal Collaboration Touchpoints

FDA Digital Health Advisory Board

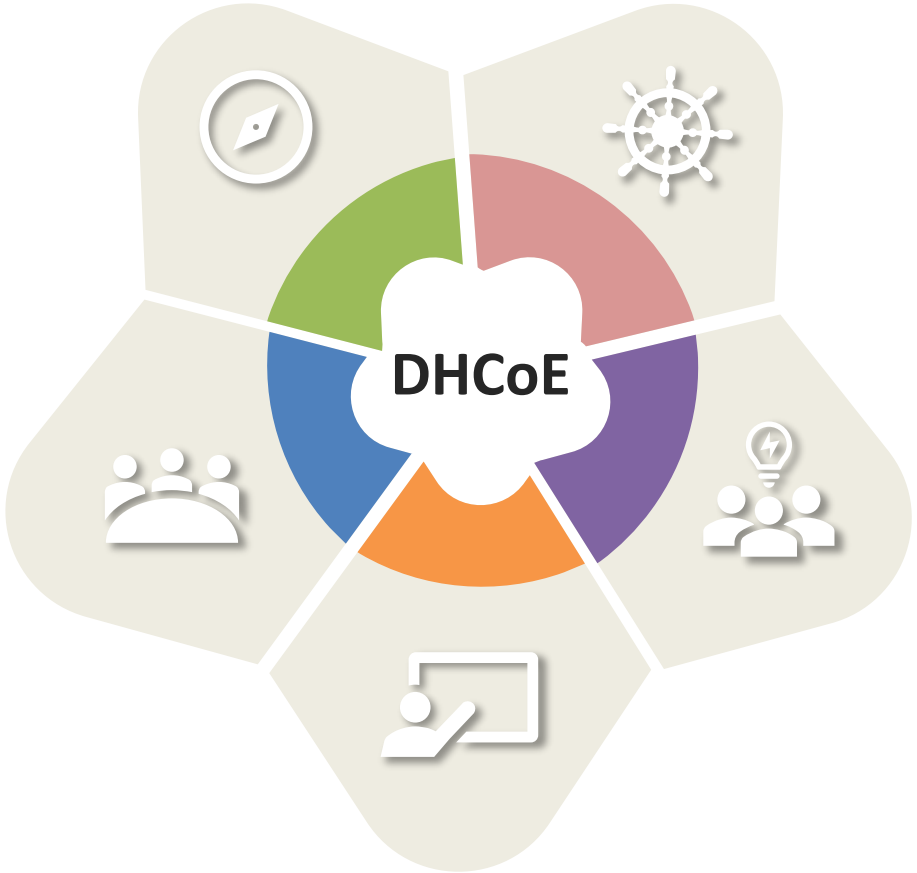
Senior leaders from FDA centers advising the DHCoE to identify and drive coordination on topic areas of work of common interest

Subcommittees

Reporting to the DHAB, subcomittees are charged to coordinate efforts on digital health topic areas affecting submissions to FDA.

Current Subcommittees:

- AI/ML
- Digital Health Technology evaluation



CDRH Digital Health Steering Committee

CDRH level steering committee established in 2015 to consistently apply policies to novel medical device submissions related to digital health technologies including software and identify policy development needs.

Program Directors Forum

Forum that brings program directors representing efforts within CDRH to stay coordinated and drive synergy within CDRH on digital health including cybersecurity, advanced manufacturing and patient science.

Training and Education

Forums bringing speakers and thought leaders to FDA.

We foster digital-health focused collaborations and interactions that advance public health

CERSIs

FDA's Centers of Excellence in Regulatory Science and Innovation

Collaborations between FDA and academic institutions through innovative research, training, and scientific exchanges.

Collaborative Communities

Continuing forums in which private- and public-sector members work together on medical device challenges. These groups can invite CDRH to participate.

Public-Private Partnerships

Initiatives that promote the development of new tools, methods and approaches to foster innovation and efficiency into FDA regulated product development.

NoDEx

FDA Network of Digital Health Experts

A pool of vetted experts who share knowledge and experience regarding digital health issues with FDA staff on an as-needed basis.

FDA Digital Health Inbox

Help navigating the FDA's current policies on digital health products and providing informal feedback.

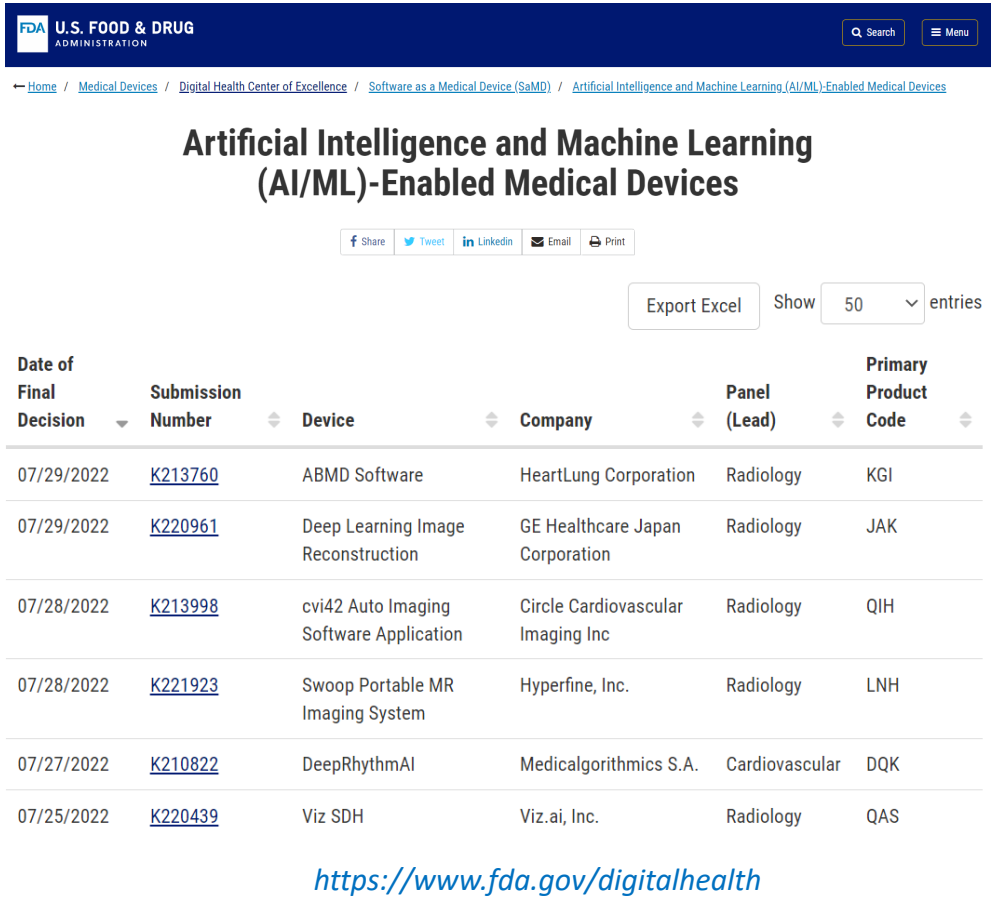
Outline

- Digital Health and the CDRH DH Center of Excellence
- **Artificial Intelligence/ Machine Learning (AI/ML)-Enabled Medical Devices**
- Good Machine Learning Practice
- Observations from the Pre-Cert Pilot

DHCoE: Areas of Focus

- Software as a Medical Device (SaMD)
- Artificial Intelligence/ Machine Learning
- Wearables
- Software in a Medical Device (SiMD)
- Wireless connectivity
- Interoperability
- Medical Device Cybersecurity
- Virtual Reality/ Augmented Reality
- Real-world Evidence and Advanced Clinical Studies
- Advanced Manufacturing
- Patient-Generated Health Data

Updated List of AI/ML-Enabled Medical Devices



U.S. FOOD & DRUG ADMINISTRATION

Search | Menu

Home / Medical Devices / Digital Health Center of Excellence / Software as a Medical Device (SaMD) / Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices

Artificial Intelligence and Machine Learning (AI/ML)-Enabled Medical Devices

Share | Tweet | LinkedIn | Email | Print

Export Excel Show 50 entries

Date of Final Decision	Submission Number	Device	Company	Panel (Lead)	Primary Product Code
07/29/2022	K213760	ABMD Software	HeartLung Corporation	Radiology	KGI
07/29/2022	K220961	Deep Learning Image Reconstruction	GE Healthcare Japan Corporation	Radiology	JAK
07/28/2022	K213998	cvi42 Auto Imaging Software Application	Circle Cardiovascular Imaging Inc	Radiology	QIH
07/28/2022	K221923	Swoop Portable MR Imaging System	Hyperfine, Inc.	Radiology	LNH
07/27/2022	K210822	DeepRhythmAI	Medicalgorithmics S.A.	Cardiovascular	DQK
07/25/2022	K220439	Viz SDH	Viz.ai, Inc.	Radiology	QAS

https://www.fda.gov/digitalhealth

Updated in October 2022

Currently Marketed AI/ML-Enabled Medical Devices

This list is meant to be:

1. A public resource on these devices and the FDA's work in this area
2. Show how AI/ML is being used across medical disciplines

A non-exhaustive list based on publicly available information

AI/ML-Enabled Medical Devices: Opportunities & Challenges

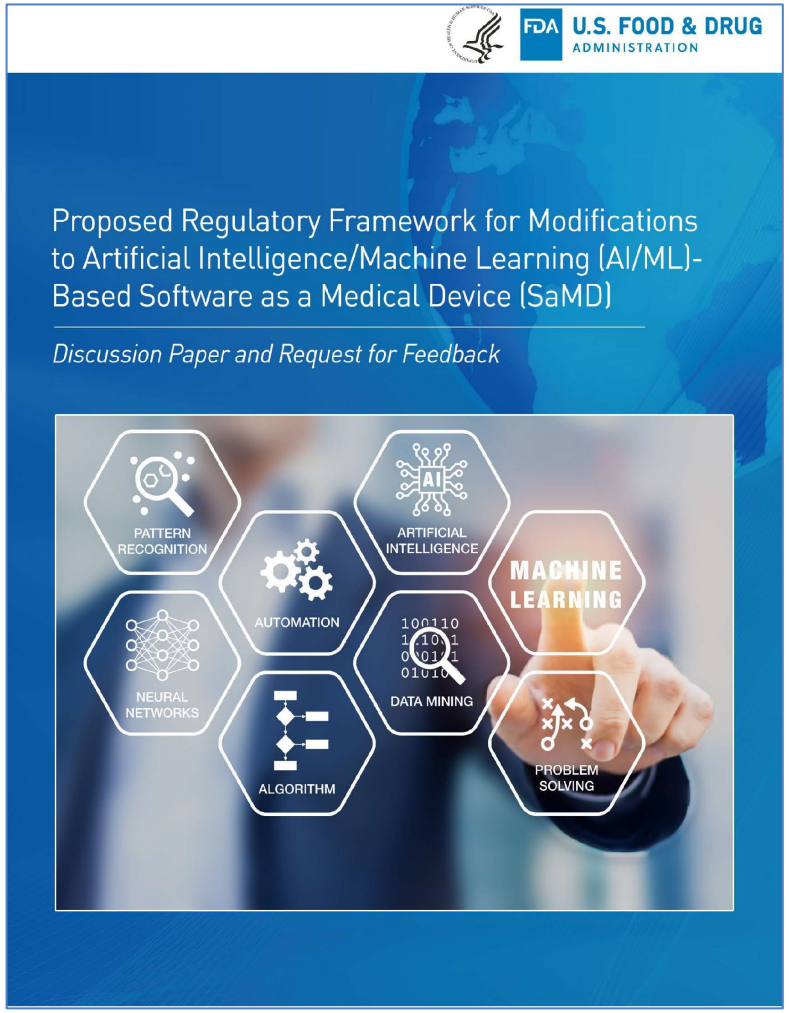
OPPORTUNITIES

- **Significant positive impact on health care**
  - Earlier disease detection
  - More accurate diagnosis
  - New insights into human physiology
  - Personalized diagnostics and therapeutics
- **Applications across all medical fields**
- **Ability to learn, adapt, and improve performance**

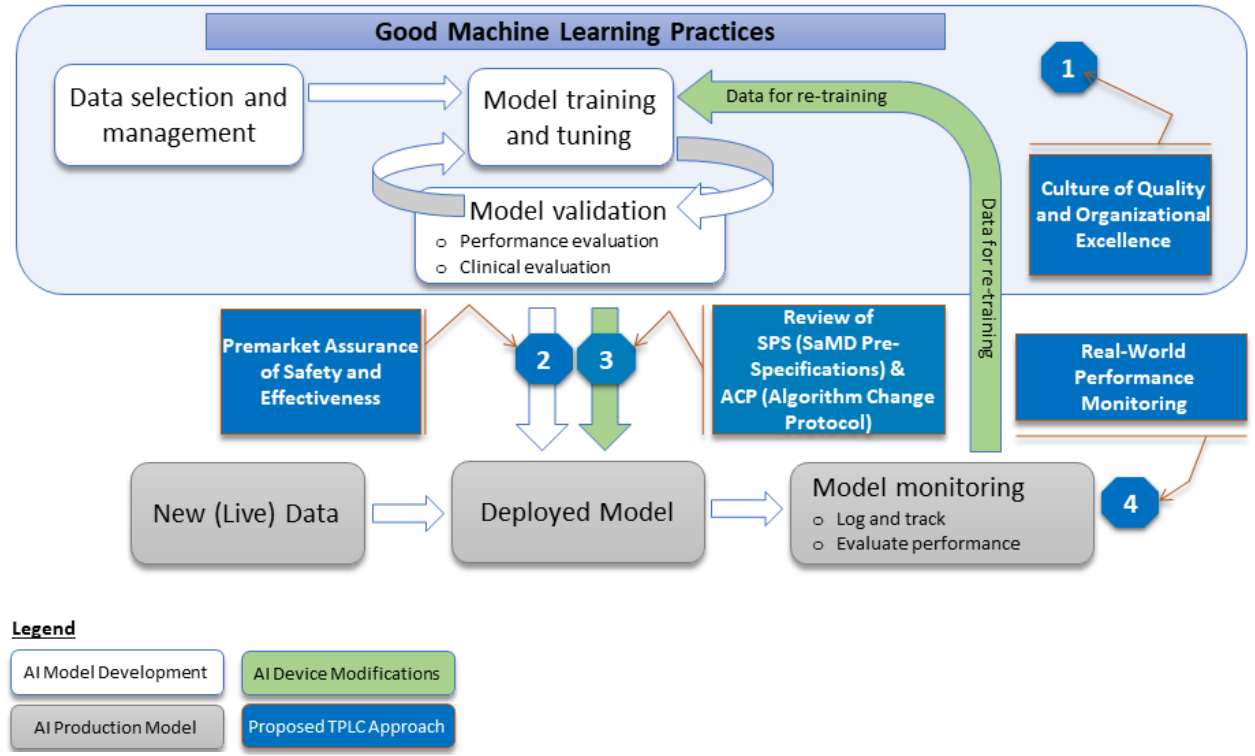
CHALLENGES

- **Fit-for-purpose data sets for development and testing, including diversity**
- **Identification and minimization of bias**
- **Opacity of some algorithms**
- **Providing oversight for an adaptive system**
- **Ensuring transparency to users**

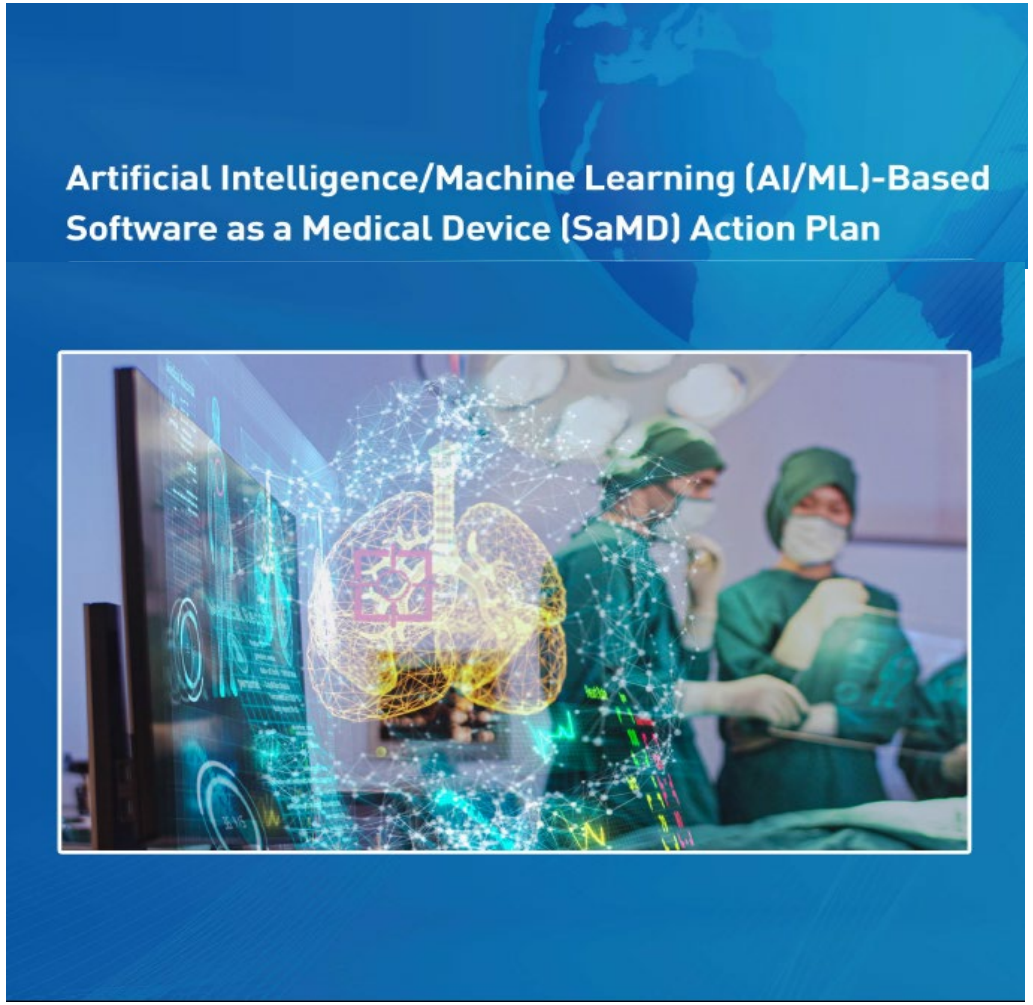
Proposed Regulatory Framework for AI/ML-Enabled Device Software



Overlay of FDA's TPLC Approach on AI/ML Workflow



Next Steps: Tailoring a Regulatory Framework



Published in 2021

Action Plan for AI/ML-Based SaMD

Outlines five next steps to advancing access:

1. Update the proposed AI/ML regulatory framework
2. Strengthen FDA's role in harmonizing GMLP
3. Foster a patient-centered approach
4. Support development of regulatory science methods
5. Advance real-world performance pilots

Stakeholder Feedback on AI/ML Approach

What we heard from stakeholders:

1. **Regulatory Framework**: Requested further development of proposed regulatory framework for AI/ML-based SaMD
2. **Good Machine Learning Practices (GMLP)**: Supported the idea of GMLP and the need for harmonization of its efforts
3. **Transparency**: Asked for further discussion with FDA on how these technologies interact with people, including transparency to users
4. **Regulatory Science**: Described need for improved methods related to algorithmic bias and robustness.
5. **Real-World Performance (RWP)**: Sought clarity on RWP monitoring for AI/ML software.

Stakeholder Feedback on AI/ML Approach

What we heard, and what we'll do

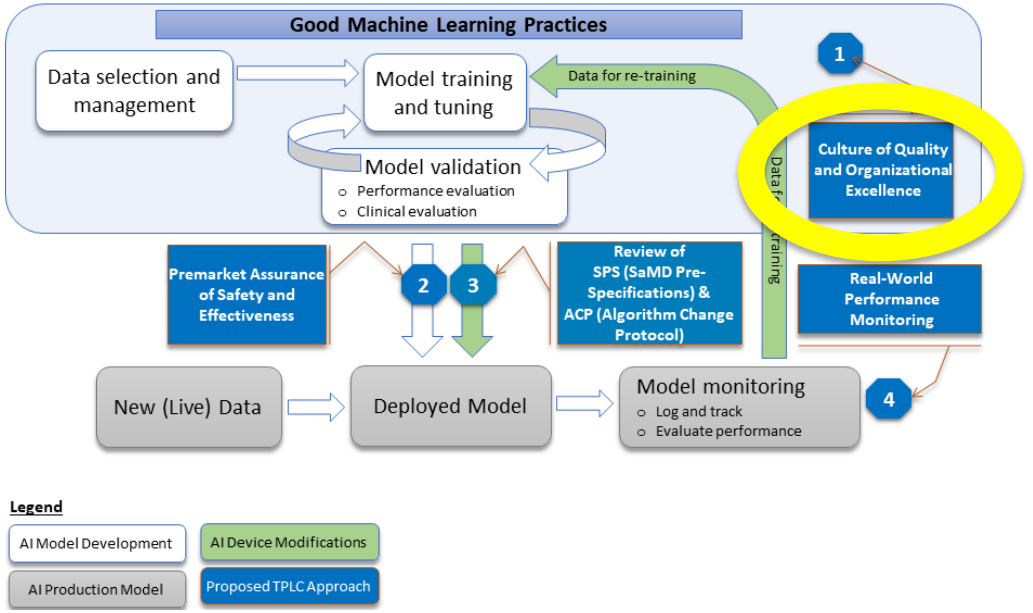
What we heard from stakeholders:	What we'll do -- The AI/ML Action Plan:
1. **Regulatory Framework**: Requested further development of proposed regulatory framework for AI/ML-based SaMD	1. **Update the proposed AI/ML framework**, including through Guidance
2. **Good Machine Learning Practices (GMLP)**: Supported the idea of GMLP and the need for harmonization of its efforts	2. **Strengthen FDA's role in harmonizing GMLP** through standards development & other initiatives
3. **Transparency**: Asked for further discussion with FDA on how these technologies interact with people, including transparency to users	3. **Foster a patient-centered approach**, starting with a workshop on transparency to users
4. **Regulatory Science**: Described need for improved methods related to algorithmic bias and robustness.	4. **Support development of regulatory science methods** related to algorithm bias and robustness
5. **Real-World Performance (RWP)**: Sought clarity on RWP monitoring for AI/ML software.	5. **Advance real-world performance pilots** in coordination with stakeholders and other programs

Outline

- Digital Health and the CDRH DH Center of Excellence
- Artificial Intelligence/ Machine Learning (AI/ML)-Enabled Medical Devices
- **Good Machine Learning Practice**
- Observations from the Pre-Cert Pilot

Good Machine Learning Practice (GMLP)

- Accepted practices in AI/ML algorithm design, development, training, and testing that facilitate the quality development and assessment of AI/ML-enabled devices
- Based on concepts from quality systems, software reliability, machine learning, and data analysis



Overlay of FDA's TPLC approach on AI/ML workflow

Adapted from Proposed Regulatory Framework for Artificial Intelligence/Machine Learning (AI/ML)-Based SaMD

Good Machine Learning Practice (GMLP)

Examples of Collaborative Efforts

- Standards Development:
  - IEEE AI Medical Device Working Group
  - ISO/IEC SubCommittee on AI 42 (ISO/ IEC JTC 1/SC 42)
  - AAMI/BSI Initiative on AI in Medical Technology
  - CTA R13 Artificial Intelligence in Healthcare
- Collaborative Communities:
  - Collaborative Community on Ophthalmic Imaging
  - Pathology Innovation Collaborative Community
  - Digital Health Measurement Collaborative Community
  - AFDO/RAPS Healthcare Products Collaborative*
- Other Collaborations:
  - IMDRF AI Medical Devices WG

Collaborative Community on Ophthalmic Imaging

Pathology Innovation Collaborative Community

Digital Health Measurement Collaborative Community

Healthcare Products Collaborative Community

*Recently transitioned from Xavier AI World Consortium Collaborative Community to the Association of Food and Drug Officials/Regulatory Affairs Professionals Society (AFDO/RAPS)

Good Machine Learning Practice Principles

U.S. FOOD & DRUG ADMINISTRATION | Health Canada Santé Canada | Medicines & Healthcare products Regulatory Agency

These guiding principles are intended to:

- Help inform the development of GMLP and encourage broad stakeholder engagement
- Promote global harmonization in efforts for the identification of best practices and the creation of standards

Good Machine Learning Practice for Medical Device Development: Guiding Principles	
Multi-Disciplinary Expertise are Leveraged Throughout the Total Product Life Cycle	**Good Software Engineering and Security Practices are Implemented**
Clinical Study Participants and Data Sets are Representative of the Intended Population	**Training Data Sets are Independent of Test Sets**
Selected Reference Datasets are Based Upon Best Available Methods	**Model Design is Tailored to the Available Data and Reflects the Intended Use of the Device**
Focus is Placed on the Performance of the Human-AI Team	**Testing Demonstrates Device Performance during Clinically Relevant Conditions**
Users are Provided Clear, Essential Information	**Deployed Models are Monitored for Performance and Re-training Risks are Managed**

https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

Good Machine Learning Practice Principles

Principle 1

Multi-Disciplinary Expertise Is Leveraged Throughout the Total Product Life Cycle:

- In-depth understanding of a model's intended integration into clinical workflow, and the desired benefits and associated patient risks, can help ensure that ML-enabled medical devices are safe and effective and address clinically meaningful needs over the lifecycle of the device.

Good Machine Learning Practice Principles

Principle 2

Good Software Engineering and Security Practices Are Implemented:

- Model design is implemented with attention to the “fundamentals”: good software engineering practices, data quality assurance, data management, and robust cybersecurity practices.
- These practices include methodical risk management and design process that can appropriately capture and communicate design, implementation, and risk management decisions and rationale, as well as ensure data authenticity and integrity.

Good Machine Learning Practice Principles

Principle 3

Clinical Study Participants and Data Sets Are Representative of the Intended Patient Population:

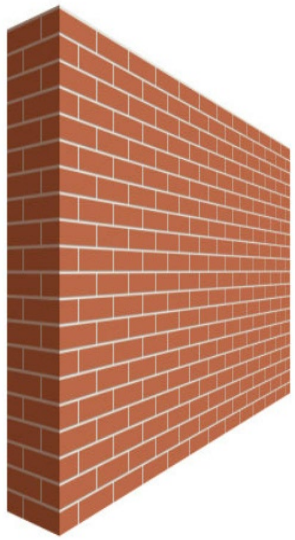
- Data collection protocols should ensure that for clinical study, training and test datasets, the following are sufficiently represented in a sample of adequate size:
  - the relevant characteristics of the intended patient population (e.g., age, gender, sex, race, and ethnicity),
  - use, and
  - measurement inputs

  so that results can be reasonably generalized to the population of interest.
- This is important to manage any bias, promote appropriate and generalizable performance across the intended patient population, assess usability, and identify circumstances where the model may underperform.

Good Machine Learning Practice Principles

Principle 4

Training Data Set



Test Data Set

Training Data Sets Are Independent of Test Sets:

- Training and test datasets are selected and maintained to be appropriately independent of one another.
- All potential sources of dependence, including patient, data acquisition, and site factors, are considered and addressed to assure independence.

Good Machine Learning Practice Principles

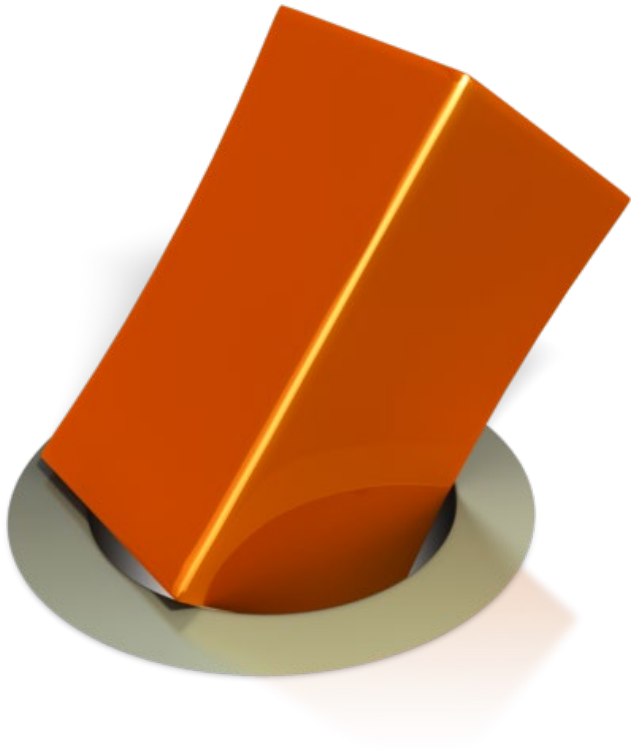
Principle 5

Selected Reference Datasets Are Based Upon Best Available Methods:

- Accepted, best available methods for developing a reference dataset (that is, a reference standard) ensure that clinically relevant and well characterized data are collected and the limitations of the reference are understood.
- If available, accepted reference datasets in model development and testing that promote and demonstrate model robustness and generalizability across the intended patient population are used.

Good Machine Learning Practice Principles

Principle 6



Model Design Is Tailored to the Available Data and Reflects the Intended Use of the Device:

- Model design is suited to the available data and supports the active mitigation of known risks, like overfitting, performance degradation, and security risks.
- The clinical benefits and risks related to the product are well understood, used to derive clinically meaningful performance goals for testing, and support that the product can safely and effectively achieve its intended use.
- Considerations include the impact of both global and local performance and uncertainty/variability in the device inputs, outputs, intended patient populations, and clinical use conditions.

Good Machine Learning Practice Principles

Principle 7

Focus Is Placed on the Performance of the Human-AI Team

- Where the model has a “human in the loop,” human factors considerations and the human interpretability of the model outputs are addressed with emphasis on the performance of the Human-AI team, rather than just the performance of the model in isolation.

Good Machine Learning Practice Principles

Principle 8

Testing Demonstrates Device Performance During Clinically Relevant Conditions:

- Statistically sound test plans are developed and executed to generate clinically relevant device performance information independently of the training data set.
- Considerations include:
  - the intended patient population,
  - important subgroups
  - clinical environment and use by the Human-AI team,
  - measurement inputs, and
  - potential confounding factors.

Good Machine Learning Practice Principles

Principle 9

Users Are Provided Clear, Essential Information:

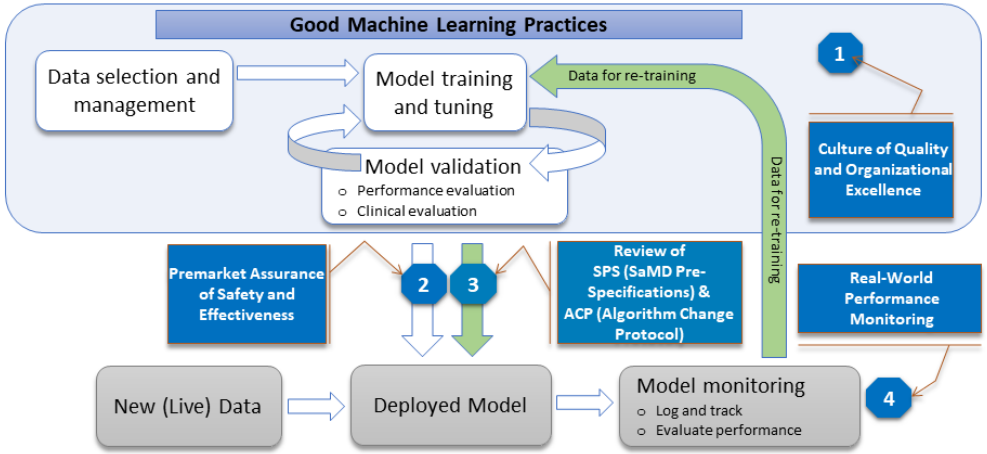
- Users are provided ready access to clear, contextually relevant information that is appropriate for the intended audience (such as health care providers or patients) including:
  - the product's intended use and indications for use,
  - performance of the model for appropriate subgroups,
  - characteristics of the data used to train and test the model,
  - acceptable inputs,
  - known limitations,
  - user interface interpretation,
  - and clinical workflow integration of the model.
- Users are also made aware of device modifications and updates from real-world performance monitoring, the basis for decision-making when available, and a means to communicate product concerns to the developer.

Good Machine Learning Practice Principles

Principle 10

Deployed Models Are Monitored for Performance and Re-training Risks Are Managed:

- Deployed models have the capability to be monitored in "real world" use with a focus on maintained or improved safety and performance.
- Additionally, when models are periodically or continually trained after deployment, there are appropriate controls in place to manage risks of overfitting, unintended bias, or degradation of the model (for example, dataset drift) that may impact the safety and performance of the model as it is used by the Human-AI team.



Good Machine Learning Practice Principles

Good Machine Learning Practice for Medical Device Development: Guiding Principles	
Multi-Disciplinary Expertise are Leveraged Throughout the Total Product Life Cycle	**Good Software Engineering and Security Practices are Implemented**
Clinical Study Participants and Data Sets are Representative of the Intended Population	**Training Data Sets are Independent of Test Sets**
Selected Reference Datasets are Based Upon Best Available Methods	**Model Design is Tailored to the Available Data and Reflects the Intended Use of the Device**
Focus is Placed on the Performance of the Human-AI Team	**Testing Demonstrates Device Performance during Clinically Relevant Conditions**
Users are Provided Clear, Essential Information	**Deployed Models are Monitored for Performance and Re-training Risks are Managed**

We envision these guiding principles may be used to:

- Adopt good practices that have been proven in other sectors;
- Tailor practices from other sectors so they are applicable to medical technology and the health care sector; and
- Create new practices specific for medical technology and the health care sector.

https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

General AI/ML-Related Standards

ISO/IEC Subcommittee on Artificial Intelligence 42 (ISO/ IEC JTC 1/SC 42)

Working Group 1	Working Group 2	Working Group 3	Working Group 4	Working Group 5
Foundational (terminology, framework)	Big Data (vocabulary, reference architecture)	Trustworthiness (risk, robustness, bias)	Use Cases & Applications	Computational approaches & characteristics of AI
22989 Concepts & Terminology 23053 Framework for AI using ML 38507 Governance implications 42001 Management system 25059 Quality Model for AI systems	20546 Big Data Overview & vocab 20547-1 Framework & application 20547-2 Use cases & derived req. 20547-3 Reference architecture 20547-4 Security & privacy 20547-5 Standards roadmap 24668 Process mgmt framework 5259-1 Data quality ML - Overview 5259-2 Data quality measures 5259-3 DQ Mgmt reqs. & guidelines 5259-4 DQ Process framework 5259-5 DQ Governance	24027 Bias in AI systems… 24028 Overview of Trustworthiness 24029-1 Robustness of Neural Nets 24029-2 Formal methods 12791 IT AI unwanted bias in C & R 24368 Ethical & societal concerns 25059 SQuaRE (quality reqs.) TR5469 Functional safety TS5471 Quality evaluation 5723 Trustworthiness Vocabulary TS6254 Explainability of ML	24030 Use cases and application 5338 AI system life cycle processes 5339 Guidelines for AI applications	24372 Computations approaches 4213 IT AI Classification perform. 5392 IT AI Ref architect knowl eng Ontologies, KE, & Represent

Joint Working Group 1 (with SC40) Governance implications of AI

Joint Working Group 2 (with SC7) Testing of AI-based systems

Red text indicates published document

General IEEE AI/ML-Related Standards

2049 series	2807 series	Other 2800 series	3000 series	7000 series
Human Augmentation	Knowledge Graphs (KG)			Addressing Ethical Concerns
P2049.1 Taxonomy & Definitions P2049.2 Privacy & Security P2049.3 Identity P2049.4 Methodologies, Processes for Ethical Considerations	P2807 Framework of KG P2807.1 Evaluation of KG P2807.2 Guide to App of Finance KG P2807.4 Guide to App of Science KG IC20-012 Roadmap for KG	P2817 Verification of Auton. System 2830 Trusted Exec. Shared ML P2840 Responsible AI Licensing P2841 Deep Learning Evaluation P2846 Auto Vehicle Behavior Model P2863 Org Governance of AI P2894 Framework Explainable AI IC20-027 Responsible Innovation of AI and Life Sciences	P3333.1.3 DL Visual Experience HF P3652.1 App of Federated ML	7000 Model Process P7001 Transparency of Autonomous Systems P7002 Data Privacy Process P7003 Algorithmic Bias P7006 Personal Data AI Agent 7007 Ontological Standard P7008 Ethically Driven Nudging P7009 Fail-Safe Design 7010 Wellbeing Metrics 7010.1 Environmental Social Gov. P7011 Trustworthiness of News P7012 Machine Readable Privacy Terms P7014 Emulated Empathy P7015 Data and AI Literacy

Red text indicates published document

Outline

- Digital Health and the CDRH DH Center of Excellence
- Artificial Intelligence/ Machine Learning (AI/ML)-Enabled Medical Devices
- Good Machine Learning Practice
- **Observations from the Pre-Cert Pilot**

Pre-Cert Pilot Program

Vision

Principles

- Patient Safety
- Product Quality
- Clinical Responsibility
- Cybersecurity Responsibility
- Proactive Culture

Software Precertification (Pre-Cert) Pilot Program

In September 2022, FDA completed an important step towards identifying regulatory approaches to software that can better promote and protect public health

The report provides an overview of Pre-Cert and summarizes key findings from the Pilot.

The report's **appendices** describe the types of objectives a new approach could better support and the types of data, or key performance indicators (KPIs), that may provide insight into those outcomes.

Excellence Appraisal Fundamental Premises (Appendix A)

1. Consistent and rigorous application of appropriate software development, monitoring, and maintenance processes results in high quality, safe, and effective medical device software;
2. Some information currently reviewed in premarket submissions describes software development processes and practices that are applicable to devices across the organization and are not solely device-specific; and
3. A regulatory framework that ensures continuous and consistent application of, and promotes continuous improvements to, organizational processes can be a more effective and efficient regulatory approach than repeated, standalone device-specific reviews of this information.

Descriptive Key Performance Indicator (KPI) Objectives (Appendix A)

1. Processes engage the right people, at the right times, to the right degree:
   "We use knowledgeable, qualified, and multi-disciplinary teams throughout the TPLC."

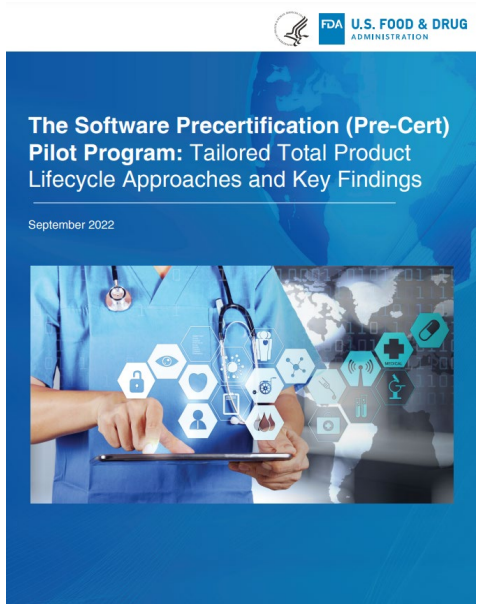
2. Development process results in well-characterized software:
   "Our software behaves as expected."

3. Deployment and monitoring process confirms well-characterized software in context of use:
   "Our software behaves as expected in the real-world."

4. Patching process ensures timely resolution of issues across the entire installed base:
   "We can fix our software when it doesn't behave as expected."

5. Update process ensures modifications meet user needs identified through real-world use:
   "We can identify and implement improvements to the expected behavior of our software."

Example KPI Formulas & Data Structures (Appendix B)



2. Data Quality

2.1. Examples of derived measure(s)/formula(s):

2.1.1. Ratio of data to errors: the number of data errors per data set per given time period divided by the count of the total number of items of the data set per given time period

2.1.2. Data transformation error rate: the number of data transformation operation fails per given time period divided by the total number of the data transformation units per given time period

2.2. Examples of observed base measures from data log events and calculations:

2.2.1. Number of data errors

2.2.2. Number of empty values

2.2.3. Average and quantiles of time data issue open

2.3. Examples of observed attributes of interest:

2.3.1. Systems and processes risk priority

2.3.2. Issue risk priority (e.g., high, medium, low)

2.3.3. Issue types (e.g., duplicate data, inaccurate data, inconsistent data, etc.)

Conclusions

GMLP is essential to ensure safe and effective AI/ML-enabled medical devices

Harmonized standards development and collaborative communities are critical for establishing accepted GMLP

FDA is exploring innovative ways to incorporate excellence appraisal into a tailored regulatory approach

Further Questions or Feedback

www.fda.gov/digitalhealth

DigitalHealth@fda.hhs.gov

SHAWN FORREST

Digital Health Specialist, Digital Health Center of Excellence (DHCoE)

Center for Devices and Radiological Health, U.S. Food and Drug Administration

Email: shawn.forrest@fda.hhs.gov