|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| A black and white logo  Description automatically generated with low confidence | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2022-2024 | | FG-AI4H-P-202 | |
| **ITU-T Focus Group on AI for Health** | |
| **Original: English** | |
| **WG(s):** | | Plenary | Helsinki, 20-22 September 2022 | |
| **DOCUMENT** | | | | |
| **Source:** | | FG-AI4H | | |
| **Title:** | | DEL02: Overview of regulatory concepts on artificial intelligence for health | | |
| **Contact:** | | Shada Alsalamah World Health Organization | | Email: [Alsalamahs@WHO.int](mailto:Alsalamahs@WHO.int) |
| **Contact:** | | Sameer Pujari World Health Organization | | Email: [Pujaris@who.int](mailto:Pujaris@who.int) |

|  |  |
| --- | --- |
| **Abstract:** | This publication contains an overview of regulatory concepts on artificial intelligence for health that is not intended as a guidance, as a regulatory framework, or policy. Rather, it is a discussion of key regulatory concepts and a resource that can be considered by all relevant stakeholders, including but not limited to, developers who are exploring and developing AI systems, regulators and policymakers who might be in the process of identifying approaches to manage and facilitate AI systems, manufacturers who design and develop AI-enabled medical devices, and health practitioners who deploy and use such medical devices and AI systems. This Deliverable contains considerations in six general topic areas: Documentation and transparency, total product lifecycle approach and risk management, intended use and analytical and clinical validation, data quality, privacy and data protection, and engagement and collaboration. Stakeholders are invited to take into account the considerations detailed in this Deliverable as they continue to develop frameworks and best practices for the use of AI in healthcare and therapeutic development.  This deliverable was approved by the FG-AI4H at its meeting P in Helsinki, 20-22 September 2022 based on P-047. |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | **International Telecommunication Union** | | |
|  | |  | | |
| **ITU-T** | **FG-AI4H Deliverable** | |
| TELECOMMUNICATION STANDARDIZATION SECTOR OF ITU | | (22 September 2022) |
|  |  | | | |
|  | DEL02  Overview of regulatory concepts on artificial intelligence for health | | | |
|  |  | | | |

Summary

This publication contains an overview of regulatory concepts on artificial intelligence for health that is not intended as a guidance, as a regulatory framework, or policy. Rather, it is a discussion of key regulatory concepts and a resource that can be considered by all relevant stakeholders, including but not limited to, developers who are exploring and developing AI systems, regulators and policymakers who might be in the process of identifying approaches to manage and facilitate AI systems, manufacturers who design and develop AI-enabled medical devices, and health practitioners who deploy and use such medical devices and AI systems. This Deliverable contains considerations in six general topic areas: Documentation and transparency, total product lifecycle approach and risk management, intended use and analytical and clinical validation, data quality, privacy and data protection, and engagement and collaboration. Stakeholders are invited to take into account the considerations detailed in this Deliverable as they continue to develop frameworks and best practices for the use of AI in healthcare and therapeutic development.

Keywords

Artificial Intelligence; Health; AI for health; Regulatory Considerations; guidance; life cycle; software as a medical device (SaMD)

Change Log

This document contains Version 1 of the Deliverable DEL02 on "*Overview of regulatory concepts on artificial intelligence for health*" approved at the ITU-T Focus Group on AI for Health (FG-AI4H) meeting held in (22 September 2022). This deliverable is an output of its Working Group on Regulatory Considerations on AI for Health (WG-RC).

Acknowledgements

Development of this guidance document was led by Sameer Pujari (Department of Digital Health and Innovation), under the overall guidance of Bernardo Mariano (Former Director, Digital Health and Innovation), Alain Labrique (Director, Digital Health and Innovation), and Soumya Swaminathan (Chief Scientist).

Shada Alsalamah (Department of Digital Health and Innovation) was the lead writer and subgroup lead of the Data Quality and Risk Management and AI Systems Development Lifecycle Approaches topic areas along with Judith Van Andel (Department of Digital Health and Innovation). The remaining topic areas were led by the following subgroup leads: M. Khair ElZarrad, (U.S. Food and Drug Administration, USA) led the Documentation & Transparency topic area; Monique Kuglitsch (Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, HHI, Germany) and Dean Ho (National University of Singapore, Singapore) jointly led the Engagement & Collaboration topic area; Naomi Lee (The Lancet, UK) led the Intended Use and Analytical & Clinical Validation topic area; and Rose Purcell (U.S. Food and Drug Administration, USA) led the Privacy and Data Protection topic area.

WHO is grateful to the following individuals who contributed to development of this publication:

Subgroup technical experts

Pat Baird (Philips, USA), Data Quality topic area subgroup technical expert.

Tim Kelsey (HIMSS, UK), Privacy and Data Protection and Documentation & Transparency topic areas subgroups technical expert.

Andrea Keyter (Key Largo Consulting Group, South Africa), Documentation & Transparency and Engagement & Collaboration topic areas subgroups technical expert.

Xiaoxuan Liu (University of Birmingham, UK), Intended Use and Analytical & Clinical Validation and Risk Management and AI Systems Development Lifecycle Approaches topic areas subgroups technical expert

Ugo Pagallo (University of Turin, Italy), Privacy and Data Protection and Engagement & Collaboration topic areas subgroups technical expert

Tayab Waseem (Continuum Health Ventures, USA), Risk Management and AI Systems Development Lifecycle Approaches topic area subgroup technical expert.

External expert group

Najeeb Al-Shorbaji, eHealth Development Association, Jordan

Paolo Alcini, European Medicines Agency, The Netherlands

Fazilah Shaik Allaudin, Ministry of Health Malaysia, Malaysia

Abdulgader Almoeen, National Center for AI (NCAI), Saudi Data & AI Authority (SDAIA), Saudi Arabia

Michael Berensmann, Federal institute for drugs and medical devices (BfArM), Germany

Hélio Bomfim de Macêdo Filho, Anvisa, Brazil

Luca Foschini, Evidation Health, USA

Ghazal Hassan, Mohammed VI University of Health Sciences, Casablanca, Morocco

Indra Joshi, NHSX, NHS, UK

Kassandra Karpathakis, Harvard TH Chan School of Public Health, Cambridge, USA and NHSX, NHS, UK

Vladimir Kutichev, Russian Federal Service for Surveillance in Healthcare, Russia

Marc Lamoureux, Health Canada, Canada

Tze-Yun Leong, National University of Singapore and AI Singapore, Singapore

Liang Hong, China's Center for Medical Device Evaluation (CMDE), China

Lin Anle, Health Sciences Authority, Singapore

Junaid Nabi, Harvard Business School, USA

Mats Ohlson, Mats Ohlson MT-Advisor, Sweden

Adrian Pacheco-Lopez, CENETEC Ministry of Health, Mexico

Maria Beatrice Panico, Medicines and Healthcare products Regulatory Agency (MHRA), UK

Andres Pichon-Riviere, Institute for Clinical Effectiveness and Health Policy - IECS, Argentina

Julie Polisena, Health Canada, Canada

Pierre Quartarolo, Danish Medicines Agency (DKMA), Denmark

Chandrashekar ranga, The Central Drugs Standard Control Organisation (CDSCO), India

Mansooreh Saniei, King's College London, UK

Tala H. Fakhouri, U.S. Food and Drug Administration, USA

Raymond Francis R. Sarmiento, University of the Philippines Manila, Philippines

Kanako Sasaki, Ministry of Health, Labour and Welfare, Japan

Robin Seidel, BfArM, Germany

Rama Sethuraman, Health Sciences Authority, Singapore

Robert Ssekitoleko, Makerere University, Uganda

Bev Townsend, University of York, UK

Georg Zimmermann, Paracelsus Medical University, Austria

Ali AlDalaan, Saudi Food and Drug Administration, Saudi Arabia

External reviewers

Thomas Wiegand, Fraunhofer Heinrich Hertz Institute, Germany

Luis Oala, Fraunhofer Heinrich Hertz Institute, Germany

Safaa Dirar Almajthoub, Digital Health Center of Excellence, Ministry of Health, Saudi Arabia

Asma Ibrahim Al Mannaei, Drugs & Medical Products Division, Department of Health- Abu Dhabi, UAE.

Brian Scarpelli, Connected Health Initiative, USA

Observers

Johner Christian, Johner-institut, Germany

Wolfgang Lauer, BfArM, Germany

Peng Liang, China's Center for Medical Device Evaluation, China

Dinsie Williams, Utoronto, Canada

WHO staff

Shada Alsalamah, Technical Officer, Digital Health and Innovation, Geneva

Samvel Azatyan – Team Lead, Regulatory Convergence and Networks, Geneva

Adriana Velazquez Berumen, Team Lead, Medical Devices and in Vitro Diagnostics, Geneva

Josee Hansen – Consultant, Geneva

Wouter 'T Hoen, HR Officer, Geneva

Agnes Sitta Kijo – Technical Officer, Facilitated Product Introduction, Geneva

Rohit Malpani, Consultant, Health Ethics and Governance Unit, Geneva

Bernardo Mariano, Chief Information Officer, Geneva

Sameer Pujari, Technical Officer, Digital Health and Innovation, Geneva

Andreas Reis, Co-Lead, Health Ethics and Governance Unit, Geneva

Anita Sands, Technical Officer, Regulation and Prequalification, Geneva

Soumya Swaminathan, Chief Scientist, Geneva

Mariam Shokralla, Consultant, Digital Health and Innovation, Geneva

Yu Zhao, Technical Officer, Digital Health and Innovation, Geneva

Disclaimer

This publication reflects the listed contributors' personal views and perspectives that may not necessarily reflect the positions and opinions of their organizations. Furthermore, these concepts are not inclusive and regulatory bodies may have additional or different approaches.

Executive Summary

The World Health Organization (WHO)'s mission to promote health, keep the world safe, and serve the vulnerable is articulated in its global strategy on digital health 2020–2025[[1]](#footnote-1). At the heart of this strategy, the WHO aims to improve health for everyone, everywhere by accelerating the development and adoption of appropriate, accessible, affordable, scalable and sustainable person-centric digital health solutions to prevent, detect and respond to epidemics and pandemics, developing infrastructure and applications. There are many international organizations and global players contributing to this area along with WHO including, but not limited to, The International Medical Device Regulators Forum (IMDRF), Global Harmonization Working Party (GHWP), U.S. Food and Drug Administration (US FDA) and Health Canada, The International Coalition of Medicines Regulatory Authorities (ICMRA), International Organization for Standardization (ISO), The Organization for Economic Co-operation and Development (OECD), The UK Medicines and Healthcare Products Regulatory Agency (MHRA), South African Health Products Regulatory Authority (SAHPRA), European Commission (EC), Singapore Health Sciences Authority (HSA), International Council for Harmonization of Technical Requirements for Pharmaceutical for Human Use (ICH), Japanese Pharmaceuticals and Medical Devices Agency (PMDA), Swissmedic, and Australian Therapeutic Goods Administration (TGA). These international and regional organizations, and national authorities collectively recognize the potential of Artificial Intelligence (AI) in enhancing health outcomes by improving clinical trials, medical diagnosis- and treatment, self-management of care, personalized care, as well as creating more evidence-based knowledge, skills, and competencies for professionals to support health care. Furthermore, with the increasing availability of healthcare data and the rapid progress of analytics techniques, AI has the potential to transform the health sector to meet a variety of stakeholders' needs in healthcare and therapeutic development.

In order to facilitate the safe and appropriate use of AI technologies for the development of AI systems in healthcare, the International Telecommunication Union (ITU) and the WHO have established a Focus Group on AI for Health (FG-AI4H). To support its work, FG-AI4H created several working groups, including a Working Group on Regulatory Considerations (WG-RC) on AI for Health. The WG-RC consists of members representing multiple stakeholders including regulatory authorities, policy makers, academia, and industry who explored regulatory and health technology assessment concepts and emerging "good practices" for the development and use of AI in healthcare and therapeutic development. The work of the WG-RC represents a multi-disciplinary, international effort to increase dialogue and examine key concepts for the use of AI in healthcare.

This publication, which is based on the work of the WG-RC, aims to deliver an *Overview of Regulatory Concepts on Artificial Intelligence for Health* that covers the following six general topic areas: documentation and transparency, total product lifecycle approach and risk management, intended use and analytical and clinical validation, data quality, privacy and data protection, and engagement and collaboration. This overview is *not* intended as a guidance, as a regulatory framework, or policy. Rather, it is a discussion of key regulatory concepts and a resource that can be considered by all relevant stakeholders, including but not limited to, developers who are exploring and developing AI systems, regulators and policymakers who might be in the process of identifying approaches to manage and facilitate AI systems, manufacturers who design and develop AI-enabled medical devices, and health practitioners who deploy and use such medical devices and AI systems. Consequently, the WG-RC recommends that stakeholders take into account the following considerations as they continue to develop frameworks and best practices for the use of AI in healthcare and therapeutic development:

1. *Documentation and Transparency:* Pre-specifying and documenting the intended medical purpose and development process, such as the selection and use of datasets, reference standards, parameters, metrics, deviations from original plans and updates during the phases of development should be considered in a manner that allows for the tracing of the development steps as appropriate. A risk-based approach should be considered also for the level of documentation and record keeping utilized for the development and validation of AI systems.
2. *Risk Management and AI Systems Development Lifecycle Approaches:* A total product lifecycle approach should be considered throughout all phases in the life of an AI system: pre-market development management, post-market management and change management. In addition, it is key to consider arisk management approach that addresses risks associated with AI systems, such as cybersecurity threats and vulnerabilities, underfitting, algorithmic bias, etc.
3. *Intended Use, and Analytical and Clinical Validation*: Initially, providing transparent documentation of the intended use of the AI system should be considered. Details of the training dataset composition underpinning an AI system, including size, setting and population, input and output data and demographic composition should be transparently documented and provided to users. In addition, it is key to consider demonstrating performance beyond the training data through external analytical validation in an independent test dataset. This test dataset should be representative of the population and setting in which the AI system is intended to be deployed and appropriately independent of the AI training algorithm during training. Transparent documentation of the external dataset and performance metrics should be provided. Furthermore, considering a graded set of requirements for clinical validation based on risk is important. Randomized clinical trials are the gold standard for evaluation of comparative clinical performance and could be appropriate for the highest risk tools or where the highest standard of evidence is required. In other situations, consider prospective validation in a real-world deployment and implementation trial which includes a relevant comparator using accepted groups. Finally, consider a period of more intense post-deployment monitoring through post-market surveillance and market surveillance for AI systems.
4. *Data Quality:* Developers should consider whether available data is of sufficient quality to support the development of the AI system that can achieve the intended purpose. Furthermore, developers should consider deploying rigorous pre-release evaluations for AI systems to ensure that they will not amplify any of the issues discussed in Section 4 of this document, such as biases and errors. Moreover, careful design or prompt troubleshooting can help identify data quality issues early on and prevent or ameliorate possible resulting harm. Finally, stakeholders should consider mitigating data quality issues that arise in healthcare data and the associated risks, as well as continue to work to create data ecosystems to facilitate the sharing of good-quality data sources.
5. *Privacy and Data Protection:* Privacy and data protection should be considered during the design and deployment of AI systems. Early in the development process, developers should consider gaining a good understanding of applicable data protection regulations and privacy laws and ensure the development process meets or exceeds such legal requirements. Finally, it is important to consider implementing a compliance program that addresses risks and ensures that the privacy and cybersecurity practices take into account potential harm, as well as the enforcement environment.
6. *Engagement and Collaboration:* It is important to consider the development of accessible and informative platforms that facilitate engagement and collaboration, where applicable and appropriate, among key stakeholders during the AI innovation and deployment roadmap. It is fundamental to consider streamlining the oversight process for AI regulation through such engagement and collaboration to accelerate practice-changing advances in AI.

Finally, the WG-RC has provided a forum for regulators and subject matter experts to discuss regulatory concepts for the use of AI technologies and development of AI systems for health and medical purposes. The WG-RC recognizes that the AI landscape is rapidly evolving and that the concepts in this deliverable may require expansion as technology and its uses develop. It recommends that stakeholders, including regulators and developers, continue to engage and that the community at large works towards shared understanding and mutual learning. In addition, established national and international groups, such as the IMDRF and ICMRA, should continue to work on topics of AI for potential regulatory convergence and harmonization.

**CONTENTS**

| Page |
| --- |
| [1 Introduction 9](#_Toc120899773)  [2 Purpose 10](#_Toc120899774)  [3 Acronyms, definitions, and fundamental concepts 10](#_Toc120899775)  [3.1 Definitions and fundamental concepts 10](#_Toc120899776)  [3.2 Acronyms 13](#_Toc120899777)  [4 Key AI applications in healthcare and therapeutic development 14](#_Toc120899778)  [5 Topic areas of regulatory concepts 15](#_Toc120899779)  [5.1 Documentation & transparency 16](#_Toc120899780)  [5.2 Risk management and AI systems development lifecycle approach 19](#_Toc120899781)  [5.3 Intended use and analytical & clinical validation 26](#_Toc120899782)  [5.4 Data quality 32](#_Toc120899783)  [5.5 Privacy and data protection 37](#_Toc120899784)  [5.6 Engagement & collaboration 41](#_Toc120899785)  [6 Recommendations for the way forward 48](#_Toc120899786)  [7 Final considerations 50](#_Toc120899787) |

**List of Tables**

| Page |
| --- |
| [Table 1 – Six Key Topic Areas of Regulatory Concepts 16](#_Toc120899788)  [Table 2 – AI systems risk classification 23](#_Toc120899789)  [Table 3 – General data quality considerations. 37](#_Toc120899790)  [Table 4 – Examples of regulators' approaches to engagement and collaboration with stakeholders about the use of AI in healthcare and therapeutic development 42](#_Toc120899791)  [Table 5 – List of key recommendations for regulatory concepts for AI in Health based on each of the six topic areas. 48](#_Toc120899792) |

**List of Figures**

| Page |
| --- |
| [Figure 1 – A general spectrum of AI research and development in healthcare delivery and therapeutic development. 15](#_Toc120899793)  [Figure 2 – Examples of key development decision points in AI systems' development. 18](#_Toc120899794)  [Figure 3 – The process of developing and deployment of the AI system 19](#_Toc120899795)  [Figure 4 – AI system Total Product Lifecycle approach on AI workflow,, 21](#_Toc120899796)  [Figure 5 – IMDRF schematic representation of the security risk management process. 22](#_Toc120899797)  [Figure 6 – General AI medical device risk management approach. 22](#_Toc120899798)  [Figure 7 – The UK NHS Buyer's Guide to AI in Health and Care 25](#_Toc120899799)  [Figure 8 – Post-Approval Change Management Protocol (PACMP) for medical devices. 26](#_Toc120899800)  [Figure 9 – Domains of health technology regulation, assessment and management for drugs and devices. 26](#_Toc120899801)  [Figure 10 – Overview of framework for clinical evaluation of AI models in health developed by the WG-Clinical Evaluation. 29](#_Toc120899802)  [Figure 11 – The 10Vs of data. 33](#_Toc120899803)  [Figure 12 – Examples of quality check principles. 35](#_Toc120899804)  [Figure 13 – NIST Privacy Framework – Cybersecurity and Privacy Risk Relationship 40](#_Toc120899805) |

ITU-T FG-AI4H Deliverable DEL02

Overview of regulatory concepts on artificial intelligence for health

# Introduction

The World Health Organization (WHO)'s mission to promote health, keep the world safe and serve the vulnerable is articulated in its global strategy on digital health 2020–2025[[2]](#footnote-2). At the heart of this strategy, the WHO aims to improve health for everyone, everywhere by accelerating the development and adoption of appropriate, accessible, affordable, scalable and sustainable person-centric digital health solutions to prevent, detect and respond to epidemics and pandemics. This should enable countries to use health data to promote health and wellbeing to achieve the health-related Sustainable Development Goals[[3]](#footnote-3) and the triple billion targets of WHO's Thirteenth General Programme of Work, 2019–2023[[4]](#footnote-4).

In addition to WHO's efforts, there is a wave of interest by many other international and regional organizations and national authorities, where each and every one is contributing to a global ecosystem. This is mainly to collectively bridge fundamental gaps in this area. Some of the key players and their efforts include, but not limited to, the International Medical Device Regulators Forum (IMDRF)[[5]](#footnote-5), Global Harmonization Working Party (GHWP), International Coalition of Medicines Regulatory Authorities (ICMRA)[[6]](#footnote-6), International Organization for Standardization (ISO)[[7]](#footnote-7), the Organization for Economic Co-operation and Development (OECD)[[8]](#footnote-8), International Council for Harmonization of Technical Requirements for Pharmaceutical for Human Use (ICH). Moreover, there are national efforts sharing the same goal[[9]](#footnote-9).

The digital transformation of healthcare and therapeutic development, including exploring Artificial Intelligence (AI) uses, has the potential to enhance health outcomes by improving medical diagnosis, digital therapeutics, clinical trials, self-care, and evidence-based knowledge. For the purpose of this document AI is defined as "a branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviors such as learning, making decisions and making predictions. The subset of AI known as Machine Learning (ML) allows computer algorithms to learn through data, without being explicitly programmed, to perform a task"[[10]](#footnote-10). With the increasing availability of healthcare data and the rapid progress of analytics techniques, AI has the potential to transform the health sector, one of the most important sectors for societies and economies worldwide.

# Purpose

The International Telecommunication Union (ITU) is the UN specialized agency for information and communications technology. While, on the other hand, The WHO is the UN specialized agency for health. Both organizations partnered to establish an open group of experts to develop a generalizable benchmarking framework for health solutions based on AI, the ITU/WHO Focus Group on AI for Health (FG-AI4H). In order to facilitate the safe and appropriate use of AI technologies[[11]](#footnote-11) for the development of AI systems[[12]](#footnote-12) in healthcare and support its work, the FG-AI4H created a Working Group on Regulatory Considerations (WG-RC) on AI for Health. The WG-RC consists of multiple stakeholders including representatives from regulatory bodies, policy makers, academia and industry who explored regulatory and health technology assessment concepts and emerging "good practices" for the development and use of AI in healthcare and therapeutic development.

This publication is a general, high-level and nonexclusive overview of key regulatory concepts' topic areas developed by the WG-RC to support the overarching FG-AI4H framework. Recognizing that a single publication cannot address the specifics of the various AI systems that can be used for therapeutic development or healthcare applications generally, the WG-RC's overview will highlight some of the key regulatory principles and concepts, such as risk/benefit assessments and considerations for the evaluation and monitoring of the performance of AI systems developed using AI technologies.Throughout the process of developing this publication, the WG-RC took into consideration different stakeholder perspectives, as well as different global and regional settings. The WG-RC's overview is not intended as guidance, as a regulatory framework or policy. Rather, it is meant as a listing of key regulatory concepts and a resource that can be considered by all relevant stakeholders, including but not limited to, developers who are exploring and using AI technologies and developing AI systems, regulators who might be in the process of identifying approaches to manage and facilitate AI systems, manufacturers who design and develop AI systems that are embedded into medical devices, and health practitioners who deploy and use such medical devices and AI systems.

# Acronyms, definitions, and fundamental concepts

## Definitions and fundamental concepts

For the purpose of this document, some key terms and concepts are defined and/ or explained below.

This section applies to terms and concepts as they are used for the purpose of this document as part of the WG-RC; for more general terms across the FG, please refer to FG-AI4H Deliverable 0.1 (2022-09) with terms and definitions deliverable[[13]](#footnote-13).

3.1.1 Artificial intelligence:

AI is a branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviours such as learning, making decisions and making predictions. The subset of AI known as ML allows computer algorithms to learn through data, without being explicitly programmed, to perform a task[[14]](#footnote-14).

3.1.2 Trustworthiness:

Trustworthy AI in the context of this document refers to AI systems and technologies that meet stakeholder's expectation in terms of bias, explainability, provenance and other desirable characteristics. Therefore, stakeholders involved in the development, deployment, or operation of such AI-based systems should be held accountable for their proper functioning.

3.1.3 Transparency:

The term transparency, in the context of this document, refers to issues such as sharing and making available to the appropriate entities, the relevant plans, decisions, and associated reasoning and the data/datasets utilized in the conception, development and ongoing deployment and monitoring of AI systems. Transparency is multifaceted and may include public dissemination by publications in peer-reviewed journals, publishing and documenting pre-specifications for development processes including clinical trials, etc. Considerations should be given to factors, such as data privacy and intellectual property, among other things.

3.1.4 Documentation:

For the purpose of this document, the term documentation refers to processes and methods used to document, retain, and prespecify critical development ideas, including the initial conception, validation and deployment, and post-deployment plans, as well as relevant key decisions and choices and supporting rationale (e.g., selection of data/datasets) used in the development of AI systems for health and therapeutic development throughout the total life cycle (e.g., from conception to post-deployment). Methods and approaches for risk and error management, reporting, and detection of bias are all key areas for documentation. Documentation can also help facilitate the understanding of algorithm decision-making process (explainability). Documentation should allow for the tracing and audits of the development process and the steps taken in the development and validation of the AI system if needed and appropriate. This includes ensuring that changes and deviations from prespecified approaches and protocols are tracked, recorded, and justified. Although effective documentation is only one element that supports transparency, it is a key regulatory principle.

3.1.5 Privacy

Privacy is a broad and multidimensional concept.It is a universally accepted fundamental human right[[15]](#footnote-15). In nearly every nation, numerous statutes, constitutional rights, and judicial decisions seek to protect privacy. The concept of privacy includes the control over personal information, often referred to as data or information privacy. Data privacy is focused on the use and governance of personal data, including implementing policies to ensure that consumers' personal information is being collected, shared, and used in appropriate ways[[16]](#footnote-16). Privacy risks include reidentification, as well as the release of unwanted inferences about a data subject (e.g., whether they have a certain disease)[[17]](#footnote-17).

3.1.6 Data integrity

Data Integrity can be defined as "the completeness, consistency, and accuracy of data."[[18]](#footnote-18).

3.1.7 Data protection

Data protection is a more technical issue under the broader umbrella of privacy which includes more domains beyond the protection of a person's personal data. However, for the context of this document, data protection includes the requirements and methods used to store and organize data in a physically secured manner to prevent unauthorized access and use. Data protection, although also a legal issue, is focused on securing data against malicious attacks and preventing the potential exploitation of stolen data for profit. While security is necessary for protecting data, it may not be sufficient for addressing privacy[[19]](#footnote-19).

3.1.8 Health data

Health data is personal data relating to the physical or mental health of a person, and includes the provision of healthcare services, and information regarding a person's health status[[20]](#footnote-20). Health data is often considered a special category of personal data, or 'sensitive' personal data, because of the nature and influence such data has on human lives and its impact on their fundamental rights and freedoms.

3.1.9 Sources of health data

Sources of health data include data acquired from digital health and medical technologies, such as wearable devices, digital health (or electronic health) applications, and medical devices and sensors; electronic health records and administrative hospital data; data from aggregated clinical trials; bioimaging and genomic data from the sequencing of human biological materials; health-related geospatial and contact tracing data; insurance claims; and data from social media, smart phones, and other electronic devices. Sources of health data is illustrated in WHO Evolving Health Data Ecosystem diagram[[21]](#footnote-21). The health data, or special personal data, derived from these sources, including heart rate, blood glucose, genetic predispositions, fitness levels, age, weight and so on, may be subject to data protection and privacy laws. Although these laws may vary from country to country as later discussed in this section, they will inform how the data is processed and for what purpose.

3.1.10 Software as a medical device (SaMD)

SaMD is defined by the IMDRF as "software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device"[[22]](#footnote-22).

3.1.11 AI system

The IMDRF[[23]](#footnote-23) defines an AI system as a software that is developed with one or more of the techniques and approaches listed below\* and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.

\*AI techniques and approaches:

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;

(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c) Statistical approaches, Bayesian estimation, search and optimization methods.

3.1.12 AI technology

For the context of this publication, the term AI technology is referred to any AI technology (e.g. machine learning, deep learning, natural language processing, computer vision, etc.) used to develop an AI system.

## Acronyms

For the purpose of this document, the following acronyms apply.

|  |  |
| --- | --- |
| AI | Artificial Intelligence |
| CDSS | Clinical Decision Support System |
| CONSORT-AI | Consolidated Standards of Reporting Trials for AI |
| CQC | Care Quality Commission |
| CRM-N | Clinical Research Materials Notification |
| DAISAM | Data and artificial intelligence assessment methods |
| EC | European Commission |
| EU | European Union |
| FG-AI4H | Focus Group on Artificial Intelligence for Health |
| GDPR | General Data Protection Regulation |
| GHWP | Global Harmonization Working Party |
| HIPAA | Health Insurance Portability and Accountability Act |
| HSA | Health Sciences Authority |
| ICH | International Council for Harmonization of Technical Requirements for Pharmaceutical for Human Use |
| ICMRA | International Coalition of Medicines Regulatory Authorities |
| iDAIR | The International Digital Health & AI Research Collaborative |
| IMDRF | International Medical Device Regulators Forum |
| IoT | Internet of Things |
| ISO | International Organization for Standardization |
| ITU | International Telecommunication Union |
| MAS | Multi-Agent Systems |
| MHLW | Ministry of Health, Labour and Welfare |
| MHRA | Products Regulatory Agency |
| ML | Machine Learning |
| NHS | National Health Service |
| NHS | National Health Service |
| NICE | National Institute for Health and Care Excellence |
| NIST | National Institute of Standards and Technology |
| NLP | Natural Language Processing |
| OECD | The Organisation for Economic Co-operation and Development |
| PACMP | Post-Approval Change Management Protocol |
| PDPC | Singapore Personal Data Protection Act |
| PMDA | Japanese Pharmaceuticals and Medical Devices Agency |
| PRISM | Pharmaceutical Regulatory Information System |
| QMS | Quality Management System |
| SAHPRA | South African Health Products Regulatory Authority |
| SaMD | Software as a Medical Device |
| SANAS | South African National Accreditation System |
| SARs | Special Access Routes |
| SPIRIT-AI | Standard Protocol Items: Recommendations for Interventional Trials for AI |
| TGA | Therapeutic Goods Administration |
| TPLC | Total Product Lifecycle |
| UNCTAD | United Nations Conference on Trade and Development |
| US FDA | United States Food and Drug Administration |
| WG-RC | Working Group on Regulatory Considerations on Artificial Intelligence for Health |
| WHO | World Health Organization |

# Key AI applications in healthcare and therapeutic development

AI is increasingly being explored to advance healthcare on multiple fronts. The blending of technology and medicine in research and development is facilitating a wealth of innovation that continues to improve[[24]](#footnote-24). Many health-related AI systems already exist or are continuously being developed to meet a variety of stakeholders' needs in healthcare and therapeutic development. These solutions have wide-ranging uses across the spectrum of healthcare delivery and therapeutic development. For example, AI systems in healthcare are being used to support patients throughout the diagnosis and treatment of a disease, such as solutions to support adherence to therapeutics and enhance communication capabilities with care providers.

Healthcare is becoming more patient-centric with personalized approaches to decision-making[[25]](#footnote-25). This would allow data to be used to improve patient and population wellness, patient education and engagement, prevention and prediction of diseases and care risks, medication adherence, disease management, disease reversal/remission, and individualization and personalization of treatment and care. Toward these ends, AI is being increasingly incorporated and utilized in the clinical setting. For example, AI-enabled medical devices are being utilized to support clinical decision making, such as AI systems that facilitate clinical evaluations and care triaging. AI systems are also being used in medical product development and evaluation, including during drug discovery to identify potential therapeutic candidates and in clinical research for patient enrichment. Figure 1 below illustrates areas of AI research and development across the spectrum of healthcare delivery and therapeutic development. This figure is not an exhaustive listing of all AI applications but instead provides examples that are meant to show the broad range of current and potential uses of AI systems.

Diagram

Description automatically generated

Figure 1 – A general spectrum of AI research and development in healthcare delivery and therapeutic development.

The above spectrum assists in determining what regulatory concepts may be applicable and how to implement them. This document articulates a selection of key regulatory concepts and discusses topic areas relevant to many stakeholders in the current AI for health ecosystem.

# Topic areas of regulatory concepts

As mentioned previously, AI systems may be utilized across all aspects of healthcare and therapeutic development. Regardless of the AI system application category, regulators are keen not only to ensure that the AI systems are safe and effective for intended use, but also that such promising tools reach those who need them as fast as possible. Dialogue between all stakeholders participating in the AI for health ecosystem, especially developers and regulators, is highly advised as the community matures. Therefore, this publication aims to establish a common understanding around the use of the AI systems in health that can be relevant to stakeholders.

Towards achieving this aim, an extensive literature review, which included current guidelines, allowed the identification of a list of topic areas of regulatory concepts for the use of AI in healthcare and therapeutic development. At its first meeting, the WG-RC discussed the proposed topic areas and agreed to focus its deliverable on the six key areas listed in Table 1 and discussed in the remaining sections of this publication. The working group was divided into six subgroups composed of subject matter experts to draft a section on each topic area.

Table 1 – Six Key Topic Areas of Regulatory Concepts

|  |  |
| --- | --- |
| **Topic Area No.** | **Topic Area** |
| Topic Area # 1 | Documentation & Transparency |
| Topic Area # 2 | Risk Management and AI Systems Development Lifecycle Approaches |
| Topic Area # 3 | Intended Use and Analytical & Clinical Validation |
| Topic Area # 4 | Data Quality |
| Topic Area # 5 | Privacy and Data Protection |
| Topic Area # 6 | Engagement & Collaboration |

The WG-RC would like to highlight that this list is not a fully inclusive list of key concepts and hopes that this list will serve as a starting point for future deliberations and subsequent updates.

## Documentation & transparency

Documentation and transparency are critical concepts that are essential to facilitate scientific and regulatory assessments of AI systems. They also help ensure trust not only in the AI system itself, but also between developers and end-users. Accurate and comprehensive documentation is key to allowing a transparent evaluation of AI systems for health. This includes undertaking a total product lifecycle approach to prespecifying and documenting processes, methods, resources and decisions made in the initial conception, training, deployment, validation (data curation or model tuning), and post-deployment of health-related AI systems that may require regulatory oversight. The following discussion focuses on some elements related to documentation and transparency, but is not fully inclusive of all of the factors that are relevant to this important area.

Effective documentation and transparency are key elements that inform regulatory decision making. They also help establish trust and guard against biases and data dredging. The same regulatory expectations and standards that ensure the safety and effectiveness of regulated products apply for AI systems used in regulated areas. It is important for regulators to be able to trace back the development process and to have appropriate documentation of essential steps and decision points. For example, specifying the problem that developers are attempting to address, the context in which the AI system is proposed to function, the selection, cleaning and processing of data/datasets used in the development process are all aspects requiring careful documentation.

Documentation should allow for the tracking, recording, and retaining of records of essential steps and decisions, including justifications and reasoning for deviating from prespecified plans. Effective documentation may also help show that the developers are taking into consideration the full complexity of the context within which the AI system is expected to operate, and how the AI system is addressing the needs of end-users. In absence of transparent documentation, it becomes difficult to ascertain whether the proposed approaches can be generalized from the data presented in the submission material to real-world deployments in new settings, which may markedly reduce performance[[26]](#footnote-26). Figure 2 below shows some examples of essential steps and decision points that developers are encouraged to consider for documentation purposes.

Different entities, disciplines and multidisciplinary expertise are likely to be involved in the development of AI systems for health and therapeutic development. There is a need to develop a shared understanding of procedures required for transparent documentation and to show that decisions were scientifically sound. Systems used to track and document the development processes and key decision points should record access and be protected against data manipulation and adversarial attacks.

Documentation and transparency should not be viewed as a burden, but as an opportunity to show the strength of a science-based development that considers the full context in which the AI system is expected to be utilized, including the characteristics of end-users. Tools and processes for documentation should be proportional to the risks involved. Conversation with regulatory organizations prior to or at early stages of development is encouraged and may provide vital help in informing documentation needs.

Beyond the regulatory perspective, it is important to note that effective documentation and other steps that help ensure transparency are important ways to establish trust and a shared understanding of AI systems in general. Other steps to facilitate transparency include, but are not limited to, publishing in peer-reviewed journals; sharing data and datasets; making code available to foster mutual learning and facilitate additional studies. Academic institutions, medical journals, and regulatory organizations among other stakeholders are working on advancing transparency for the use of AI in diagnostic and therapeutic development.

Collaborations, such as Consolidated Standards of Reporting Trials for AI (CONSORT-AI)[[27]](#footnote-27), and Standard Protocol Items: Recommendations for Interventional Trials for AI (SPIRIT-AI)[[28]](#footnote-28) have provided useful guidance about how to design trials where AI systems are used as well as how to publish results from such trials. Transparency is not only an important consideration for building trust, but it can also be an essential tool to educate end-users. This can be achieved by adapting the communication strategies to the needs of the end-users and other stakeholders if appropriate. Also, As outlined in Figure 2, the development process of an AI system is multifaceted. A methodical approach to documentation throughout the full development cycle, including deployment and post-deployment, should be considered.

Timeline

Description automatically generated

Figure 2 – Examples of key development decision points in AI systems' development.

The following are some elements that might be useful to consider from a documentation and record retention perspective.

1. Documentation across the total product lifecycle – ensuring a quality continuum

Developers should design, implement and document approaches and methods to ensure a quality continuum across the development phases. Effective documentation outlining all phases of development would further enhance confidence in the AI system and would show how expected and unexpected challenges and deviations are identified and managed. Validation processes and benchmarking should be carefully documented including the decisions for selecting specific data sets, reference standard, parameters and metrics to justify such processes. For example, careful consideration should be given to documenting how and why specific data or data sets are selected to train, externally validate and retrain the model (e.g., post-deployment retraining).

1. Pre-specification and documenting the purpose, clinical context, and development

The intended purpose/function of the AI systems should be clearly documented. For example, what is the problem that the AI system is aiming to resolve? This should take into consideration the full clinical and health contexts in which the tools are expected to function. For example, clinical care environments can be vastly complex and involve several individuals with different roles and expectations. Documenting how the AI system should function in such active environments needs to be considered. As shown in Figure 3, there are multiple processes, testing/validation steps and protocols that should be pre-specified and documented. Pre-specification is one of the most important elements that supports trust and confidence in the development process. This will show evidence of a coherent development process and will be the basis for providing justification and rationale for any future deviations and modifications.

1. Deployment and Post-deployment

AI systems may be designed using data and datasets from specific populations. As with any therapeutic, once deployed, the AI systems will be utilized by a larger population and potentially with variable end-users. Careful deployment plans and justification for targeting different end-users should be considered. Deviations from prespecified plans, updates and/or modifications of the AI system, post-deployment performance, data capture and approaches to continued assessment of the AI system should also be documented. Such approaches will be increasingly relevant once learning AI systems that may change after deployment become more common.

1. Risk Based approach and proportionality

Generally, regulatory frameworks recommend a risk-based approach where processes to identify and mitigate errors, biases and other risks should be put in place in ways that are proportional to their importance. Developers of AI systems should keep in mind that regulatory organizations have avenues for dialogue and discussions that can be utilized to shed light on regulatory requirements.

## Risk management and AI systems development lifecycle approach

There are many categories that AI systems fall into, for example, devices that rely on AI and used as a medical device (commonly known as SaMD short for Software as a Medical Device). Such category of AI systems are defined by the IMDRF as "*software intended to be used for one or more medical purposes that perform these purposes without being part of a hardware medical device*"[[29]](#footnote-29). However, the regulatory concepts for such category of AI systems are similar to typical software that are regulated as medical devices, in addition to specific concepts including but not limited to, continuous learning capabilities, level of human intervention, training of models, and retraining[[30]](#footnote-30). Furthermore, a holistic risk management approach that includes addressing risks associated with an AI system's cybersecurity threats and vulnerabilities should be considered throughout the total product lifecycle. This topic area aims to present a holistic risk-based approach for AI systems in general and those used as medical device in particular throughout its lifecycle, including during pre- and post-market deployment.

* ***AI systems during the development and deployment process***

The block diagram in Figure 3 below illustrates the process of the development and deployment of an AI systems. Developers and implementers of AI systems should establish measures to ensure their responsible development.

Diagram

Description automatically generated

Figure 3 – The process of developing and deployment of the AI system[[31]](#footnote-31)

As illustrated in Figure 3 above, all activities related to the design, development, training, validation, retraining, and deployment of AI systems should be performed and managed under an ISO 13485 based quality management system[[32]](#footnote-32). Next to clinical endpoints, AI-specific monitoring dimensions include confidence[[33]](#footnote-33), bias and robustness[[34]](#footnote-34), among others.

* ***AI Systems Development Lifecycle***

An AI system development lifecycle approach can facilitate continuous AI learning and product improvement while providing effective safeguards. This can be achieved if such an approach involves appropriate development practices for AI systems throughout their development and deployment. Furthermore, this approach could potentially increase the trustworthiness and safety of the AI systems. An example of one such approach is a Total Product Lifecycle (TPLC) approach that could include the following four key components:

* demonstration of a culture of quality and organizational excellence of the company producing the AI systems;
* premarket assurance of safety and effectiveness;
* review of AI systems pre-specifications and algorithm change protocol; and
* real-world performance monitoring.

Diagram

Description automatically generated

Figure 4 – AI system Total Product Lifecycle approach on AI workflow[[35]](#footnote-35),[[36]](#footnote-36),

* ***Holistic risk management***

Holistic risk evaluation and management should be considered, and it should take into consideration the full context within which the AI system may be utilized. This could include not only the software or AI system being developed, but also other software that may be used within the same environment or context. Other risks, such as those associated with cybersecurity threats and vulnerabilities should also be considered and managed. Risks associated with cybersecurity threats and vulnerabilities should be considered throughout all phases in the life of a medical device. Therefore, AI systems manufacturers should employ a risk-based approach to ensure the design and development of AI systems used as medical devices include appropriate cybersecurity protections. Doing so necessitates that manufacturers take a holistic approach to device cybersecurity by assessing risks and mitigations throughout the AI system development life cycle. In order to achieve this, the IMDRF has published a security risk management process, illustrated in Figure 5.

Diagram

Description automatically generated

Figure 5 – IMDRF schematic representation of the security risk management process[[37]](#footnote-37).

However, to facilitate AI systems risk management, a general holistic management approach is introduced in this subsection with three broad management categories: pre-market development management, post-market management, and change management. These categories are illustrated in Figure 6 and discussed below.

Figure 6 – General AI medical device risk management approach.

1. **Pre-market development management**

The controls and measures put in place to ensure a developed AI system functions as expected while minimizing any risk should be proportional to the risks that could be imposed if the AI system were to malfunction. For example, failure of an AI system that is designed to encourage adherence to a healthy diet is different than one that is designed to diagnose or treat certain diseases and pathologies. Therefore, developers should consider a risk-based approach throughout all involved processes to prioritize safety. Developers need to consider the intended use of the AI system and the clinical context to evaluate the level of risk. For example, the IMDRF risk framework for SaMD identifies two major factors that may also contribute to the impact or risk of AI system. The first factor is the significance of the information provided by the AI system to the healthcare decision. The significance is determined by the intended use of the information – to treat or diagnose, to drive clinical management, or to inform clinical management. The second factor is the state of the healthcare situation or condition, which is determined by the intended user, disease or condition, and the intended population for the AI systems – critical, serious, or non-serious healthcare situations or conditions. Taken together, these factors describing the intended use can be used to place the AI system- into one of four categories from lowest (I) to highest risk (IV) to reflect the risk associated with the clinical situation and device use.

Table 2 – AI systems risk classification[[38]](#footnote-38)

| State of healthcare situation or condition | Significance of information provided by AI system to healthcare decision | | |
| --- | --- | --- | --- |
| Treat or diagnose | Drive clinical management | Inform clinical management |
| Critical | IV | III | II |
| Serious | III | II | I |
| Non-serious | II | I | I |

The intended use and risk classificationshould be considered when testing different models and balancing trade-offs, such as transparency and accuracy. In cases where training data sets are limited, simpler models, such as regression or decision tree models, often provide equivalent or better results than more complex models and provide the added benefit of more transparency and interpretability. On the other hand, in cases with larger, complex data sets, complex models, such as deep learning networks, may not lend themselves to being explainable but may provide greater accuracy than simpler models. However, in cases in which there is a greater risk of harm, stakeholders should consider discussing the risks and benefits of choosing a more complex model and whether there are ways to mitigate the lack of interpretability and transparency and build trust in the model through additional validation measures.

Furthermore, depending on the level of risk, some AI system may be approved as being available for full deployment whereas others may be initially authorized for deployment in more 'AI-ready' institutions. 'AI-ready' institutions are those which are certified based on having stringent levels of surveillance in place with responsive back-up systems to handle any failure of the algorithm to minimize risk of patient harm.

Overall, it is important to achieve transparency between all AI-system stakeholders including the developers/ manufacturers, regulatory authorities, and the implementors (i.e. users in the healthcare settings such as medical practitioners). Appropriate documentation of risk management and proper auditing procedures are examples of ways that help assure transparency. Generally, auditing specific key components of the AI medical device should be considered (i.e. certain software, hardware, training data, failure cases). For instance, it is important to version control training data, as more data is added with each update. If an algorithm suddenly deteriorated in performance after an update, we may wish to inspect everything that contributed to the update. In most cases, the thing that will have changed is the addition of new training data by the developer (rather than changes to the software itself, such as modification to the neural networks). Moreover***,*** given how unpredictable changes in performance can be for AI, it is recommended to have active reporting and investigation of failure cases (in the CONSORT-AI guidelines) although it is not prescriptive, given the wide range of available reporting and investigation avenues from common-sensical clinical auditing (i.e. human inspection) to technical solutions based on inference.

Although not specific to AI, there is a thickening web of country-, nation-, and jurisdictional-specific legislations and laws that manufacturers and developers may need to consider for the development and deployment of regulated AI medical devices in healthcare. Such legislation includes, but is not limited to, the Personal Data Protection Act, Human Biomedical Research Act, Private Hospitals and Medical Clinics Act, Health Insurance Portability and Accountability Act, and General Data Protection Regulation (GDPR). Therefore, compliance with relevant laws (local, and cross-jurisdictional laws and data protection acts) needs to be demonstrated by manufacturers and developers of medical devices whether they embed an AI component or not.

1. ***Post-market management***

Post-market monitoring and surveillance of AI medical devices allows timely identification of software and hardware related problems, which may not be observed during device development, validation, and clinical evaluation since these are performed in controlled settings. New risks may surface when the software is implemented in a broader real-world context and is used by a diverse spectrum of users with different expertise. Companies involved in distributing AI medical devices (manufacturers, importers, wholesalers, authorized representatives, and registrants) are required to comply with their post-market duties and obligations which include reporting to relevant regulatory authorities in any of the following circumstances:

* Any serious public health threat
* Death, serious deterioration in state of health of patient, user or other person occurred,
* Death, serious deterioration in state of health of patient, user or other person might have occurred
* Any field safety corrective action (such as return of a type of device to the manufacturer or its representative (also known as recall, in some jurisdictions); device modification; device exchange; device destruction; advice given by the manufacturer regarding the use of the device[[39]](#footnote-39),[[40]](#footnote-40).

Furthermore, manufacturers should have an initiative to actively survey and detect possible threats as part of their post-market plan. There should be a plan outlined by the manufacturers on how they can actively monitor and respond to evolving and newly identified threats. Key considerations for the post-market plan include[[41]](#footnote-41): post-market vigilance, vulnerability disclosure, patching and updates, recovery, and information sharing. Finally, as part of the post-market duties and obligations, companies involved in distributing medical devices (manufacturers, importers, wholesalers, and registrants) are required to report adverse events associated with the use of software medical devices.

Generally speaking, there is a need for both post-market clinical performance follow-up (with the frequency of follow up determined as appropriate) and periodic safety checks. The intensity of post-market surveillance by the manufacturer may be risk-proportionate (according to consequences of failure (harm) and likelihood of early detection of failure). Finally, post-market surveillance requires a minimum level of evaluation for each site to ensure that potential algorithm vulnerabilities due to variation in local environments can be detected.

For example, the UK National Health Service (NHS) AI Lab published a guidance to accelerate a safe and effective adoption of AI in health. This guide lists ten questions, falling under four categories, to help buyers of AI products make informed decisions, problem identification, product assessment, implementation considerations, and procurement and delivery (as illustrated in Figure 7 below).

Diagram

Description automatically generated

Figure 7 – The UK NHS Buyer's Guide to AI in Health and Care[[42]](#footnote-42)

1. ***Change management***

Considering the character of AI systems, it is significant to settle the regulatory system for enabling continuous improvements through the development lifecycle. One of the models is the change management approach implemented by the Ministry of Health, Labour and Welfare (MHLW) of Japan. This approach is adapted in the Pharmaceuticals and Medical Devices Act as Post-Approval Change Management Protocol (PACMP) for medical devices[[43]](#footnote-43). The basic concept was published by the International Conference of Harmonization in application to pharmaceutical product lifecycle management, but this concept is also translatable to medical devices, such as those that utilize AI with the intent of continuous improvement through the product lifecycle[[44]](#footnote-44). The PACMP is illustrated below in Figure 8. The PACMP is a kind of Total Product Lifecycle Approach of medical devices. The PACMP focuses on the change management process in the total product lifecycle of medical devices which can be continuously improved.

Text, timeline

Description automatically generated

Figure 8 – Post-Approval Change Management Protocol (PACMP) for medical devices.

## Intended use and analytical & clinical validation

In principle, regulatory mechanisms are in place to answer the question: Do the available data (included in regulatory submission) support the conclusion that an investigational or experimental therapeutic is sufficiently safe and effective to justify entry into the market and public access? In addition to the principles in 5.1 and 5.2, assessing if the use of the solution is safe (will not harm the user) and if the claims made about the performance are robust (the device is effective) should also be considered (see Figure 9 below). To evaluate these claims for AI tools requires a clear use case description, demonstration of analytical and clinical validation, and assessment of the potential for bias or discrimination in the tool.



Figure 9 – Domains of health technology regulation, assessment and management for drugs and devices.

* ***Use case description, analytical, and clinical validation***

Demonstrating safety and consistently delivering expected performance is a critical part of product, development, deployment, clinical evaluation and regulation. Independent evaluation of a tool is important, not only for clinicians and patients and other end users of a product, but also as a quality mark for developers and procurers. For AI tools, setting out best practice in analytical and clinical validation is challenging. Not only is the regulatory landscape changing, but the technical capabilities underpinning many tools is developing rapidly, and there is a growing body of research on digital and AI interventions in health.

The performance of AI systems can be rapidly changed, not only as a result of code change, but also with the provision of different or additional training or tuning data. Evaluation that considers steps from development, to analytical and clinical validation, and post market surveillance is therefore considered best practice for AI systems.

This topic area covers the concepts of use case descriptions (including intended use statements) and analytical and clinical validation. These concepts, follow the framework proposed by the WHO/ITU FG-AI4H Working Group on Clinical Evaluation (see Figure 10). A full description of this framework, can be found in the accompanying deliverable for the Working Group on Clinical Evaluation. The following section describes the key concepts and best practices, and builds on important work from international regulatory and national bodies, such as IMDRF. It is not intended to replace the work done by these regulatory and national bodies. By outlining key concepts, this deliverable describes where challenges remain in this rapidly changing field. For example, particular consideration is given to under-resourced settings which may have less national regulatory capacity. This document also explores the role of benchmarking in the evaluation of AI systems in health, and relates these evaluation principles to this topic area, and to the WHO/ITU FG-AI4H work, in which benchmarking exploration is a key component[[45]](#footnote-45).

* ***Intended Use***

AI systems are complex, dependent not only on the constituent code, but also on the training data, clinical setting, and user interaction. They are often situated in a complex clinical pathway or are being introduced in new clinical pathways altogether (for example, into new telemedical pathways or part of the addition of new triage tools). Therefore, for AI tools, safety and performance can be highly context dependent. The description of the use case has a substantial role both to inform end users where the tool can safely and appropriately be utilized, and for regulated tools (the statement of intended use), to allow regulators to assess if the evidence of the analytical and clinical validation steps taken are appropriate and sufficient for the intended use case.

When developing a health-related AI system, it is important for stakeholders to consider and describe the relevant use case. This consideration should cover the setting (geography, type of care facility), the population (ethnicity, race, gender, age, disease type, disease severity, co-morbidities) the intended user (healthcare provider or patient facing), and the clinical situation for which it is intended. Many interventions, tests, and guidelines are prone to bias, this is a particularly important consideration for AI systems which are highly sensitive to the characteristics of the data they were trained upon and are prone to failure in unseen data types (such as a new disease feature or population type or context that was previously unencountered). Developers should also provide a clear clinical and scientific explanation of the tools' intended performance, including the populations and contexts in which it has been validated for use. Standardized reporting templates common to all stakeholders can help to more effectively communicate the intended use[[46]](#footnote-46),[[47]](#footnote-47),[[48]](#footnote-48). For some intended use cases there may be clear reasons why analytical performance of the tool would differ in different settings[[49]](#footnote-49) (for example a symptom checker may perform differently in areas with different disease epidemiology to the data on which it was trained). If this is the case, systematic known differences in performance should be included in the intended use statement. For other intended use cases, there may be emerging evidence that the tool under consideration, or other very similar tools, have been shown to have similar analytical performance in a wider setting than those in which they were initially developed and validated[[50]](#footnote-50) (for example retinal tools have been shown to have a similar performance in different populations[[51]](#footnote-51)). Understanding of the generalizability of similar tools may be considered when providing a statement of the intended use or description of the use case[[52]](#footnote-52).

As part of the risk management process, regulators may wish to request evidence that developers have considered if there are situations in which a tool should not be used (for example if there is insufficient training data for a particular patient group, or absence of validation in a particular setting), or if there are potential risks from use outside of intended settings.

* ***Analytical Validation (also referred to as technical validation)***

For the purposes of this document, analytical validation refers to the process of validating the AI tool using data, but without performing interventional or clinical studies. This may also be referred to as technical validation. Appropriate analytical validation demonstrates that a model is robust and performs to an acceptable level in the intended setting. It also enables the understanding of potential bias and generalizability (and any steps taken to understand these).

Developers should provide a description of training datasets used in model training, tuning, and internal validation. (as for the intended use case description, this should cover the size, setting, population demographics, intended user and clinical situation (with input and output data), and can use standardized reporting templates). Transparency and documentation around dataset selection and characteristics are critical to ensure that tools are used appropriately. Developers and regulators may expect that the AI tool has been externally validated in a dataset that is independent from that in which it was trained in order to demonstrate external validity and generalizability of the model beyond the dataset in which it was trained. The external dataset is expected to be representative of the setting and population that are described in the intended use (gender, race, ethnicity) to demonstrate robust performance in the intended setting. The validation dataset should be of adequate quality, with appropriate robustness of labels. As part of risk management process, it is important to identify any high-risk cases or cases that may be[[53]](#footnote-53).

Diagram, text

Description automatically generated

Figure 10 – Overview of framework for clinical evaluation of AI models in health developed by the WG-Clinical Evaluation.

Although bias, errors, and missing data are not unique to AI development, they are still serious concerns, which may arise for many reasons including unequal and non-representative training or validation datasets or structural bias in the systems where training data is generated (e.g., healthcare settings). Reporting the gender, race and ethnicity of individuals in the training and validation data cohorts, if feasible, can help in addressing the potential for bias and its impact. For example, a better understanding of bias may help identify populations for which a tool may not function as expected. Post marketing surveillance could also provide insights into the impact of potential bias.

Obtaining datasets that are sufficiently representative, and of sufficient quality can be difficult. Those local, regional, and national bodies interested in procuring AI systems could hold their own hidden dataset to enable this external validation set (for example, a scheme currently underway by the UK body NHSX, which has nationally representative datasets for some common use cases). Access to representative datasets for validation is a particular issue in many low- and middle-income countries. Where datasets are available in low resource settings, there may also be limitations introduced by the quality of the data. The ability to produce robust datasets with high quality ground truth labels is likely to be affected by limitations elsewhere in the health setting where there are barriers impeding access to diagnosis and treatment. These major challenges have the potential to not only propagate inequality of access, but also to compromise safety and performance of AI-based tools, and is a potential area of future work (for example the newly launched International Digital Health &   
AI Research Collaborative (iDAIR) mentions the use of collaborative, distributed, and responsible use of data as one of their main aims[[54]](#footnote-54)).

While most regulatory agencies have national or regional remits, there is currently a reliance within some countries with limited regulatory capacity on decisions made by other major regulators. Availability of independent, hidden, representative datasets also offers particular advantages to countries that do not have their own regulatory process, or where regulatory decisions may be informed by data provided to other bodies. However, the performance of AI-based tools is highly dependent on the context of use. Local or national bodies could perform analytical validation as a second local validation step to ensure that the performance metrics obtained are consistent with that demonstrated for other regulatory approvals. This could be best prioritized through a needs-based approach, for example, the identification of key areas in which AI-based tools are promising and could provide local value, and the potential prospective creation of datasets to support validation.

In order to understand the performance of a tool, evaluation against an accepted standard should be made. The most appropriate standard for comparison may differ by intended use but commonly used standards are human performance in a similar task or other models (for example derived from logistic regression) with strong evidence-based or mandated standards of accuracy, sensitivity and specificity (for example for screening tools). Depending on the intended use case, the requirement for comparative performance may be more or less stringent (for example when used as a triage or screening tool, a different level of comparative performance may be acceptable compared to a tool used for diagnosis).

Some limited comparative benchmarking of AI systems has been performed in a single setting but may become more common as the number of available tools increases[[55]](#footnote-55). In the future, if an AI system has proven clinical efficacy and safety in a particular setting, it may be possible and appropriate to benchmark other newer tools against the AI system to understand potential similarity of performance. Benchmarking software is being developed as part of the work of the Open Code Initiative[[56]](#footnote-56). Platforms like this may also be useful as a way to perform repeated algorithmic validation of models that have been updated. However, this is currently not the case for any use cases, and benchmarking thus far only has been used to understand comparative analytical performance. In addition, repeatedly using the same data for benchmarking multiple updated models (and thus, even if inadvertently, training to the test) risks introducing bias, and this should be taken into account when benchmarking is considered.

A designated FG-AI4H working group on Data and AI Solution Assessment Methods[[57]](#footnote-57) provides guidance on the methods, processes and software development for the analytical validation of health-related AI systems[[58]](#footnote-58).

1. ***Clinical validation***

Analytical validation performed retrospectively on an existing dataset gives measures of performance (accuracy, negative predictive value, positive predictive value), but does not allow evaluation of other factors that may affect performance of the tool (user interaction, workflow integration and unintended consequences of tool within a complex clinical pathway).

Both national and international bodies have proposed a graded set of requirements based on risk for digital health tools (significance of the information provided by the tool and the state of the health condition)[[59]](#footnote-59),[[60]](#footnote-60). The IMDRF document on clinical evaluation of SaMD (illustrated earlier in Table 2[[61]](#footnote-61)) proposes that devices in category I are the lowest risk tools that have analytical validity evidence, and that a novel tool in this category would require manufacturers to collect real world performance data and generate a demonstration of scientific validity. For higher risk SaMD, clinical evaluation evidence is expected based on analytical validity evidence. The appropriate level of evidence of adequate clinical performance for a novel AI tool before deployment is not universally agreed on and is the subject of a separate working group within the FG-AI4H (Working Group on Clinical Evaluation).

Randomised clinical trial data are the gold standard evaluation of comparative clinical performance, and may be appropriate for the highest risk devices, where an AI tool has no demonstrated performance in that setting or for large national procurement bodies that seek evaluation of performance before national expenditure. A trial that is expected to guide clinical practice should have a clinically meaningful primary endpoint (morbidity, mortality), but in certain situations, event rate or time lag between the trial and the endpoint may result in a more feasible surrogate endpoint. Reporting guidelines backed by the widely accepted EQUATOR network are now available for protocols and clinical trials using AI systems[[62]](#footnote-62). However, currently there remains a small number of actively recruiting or completed randomised trials in this field[[63]](#footnote-63).

Randomised clinical trials have potential limitations that may make other options preferable (trials can be slow, expensive, and may evaluate performance in specific groups under trial conditions). Where randomised evidence may not be necessary (for example evidence required may be proportional to the risk or cost of a tool), prospective validation, in a real-world deployment and implementation trial, with a relevant comparison group showing improvement in meaningful outcomes using validated tools or widely accepted and verified endpoints, and with systematic safety reporting may be appropriate. Clinical performance should be considered in the context of the capability of the health workers, available internet bandwidth and health informatics infrastructure, and real-time data pipelines, and developers should provide a description of the steps taken to perform clinical validation in a similar context to that which is available at the intended use setting.

Further consideration of the most appropriate level or type of clinical evaluation for a digital health intervention will be provided by the working group on clinical evaluation.

The following special considerations are specific to:

1. ***Post market monitoring***

Post-market monitoring in some regulatory contexts rely heavily on adverse event reporting. Recent WHO guidance recommends proactive post-market surveillance must be carried out by the manufacturer. As part of a total product life cycle approach to regulation in this context, further prospective post-market clinical follow-up should be completed after deployment. Regulators must be notified of reportable incidents (adverse events), and findings from more continuous monitoring using real world data may help developers and regulators better understand and assure the safety and performance of these devices in real world use. For prospective monitoring of real-world data, significant investment will be required in prospectively curating and labelling validation data. A defined period of close monitoring may be appropriate for AI based tools for those with high risk given their tendency to overfit on erroneous data features and produce unpredictable errors on unseen data features, and the lack of data from use in real world settings with long term results. Regulators may recommend that manufacturers develop specific market surveillance measures that are appropriate for AI systems.

1. ***Changes to the AI tool***

An update of an AI tool, by a code change, change of the user interface, or provision of further training data may alter the analytical or clinical performance of an AI system. The group are not aware of currently approved medical AI systems that are 'continuously learning' but anticipate that these may be developed. Such AI systems would require a risk-benefit evaluation in keeping this the concepts in this document and that of Clinical Evaluation of AI systems for health. Taking 'checkpoints', by evaluating the tool as it is currently performing at regular intervals enables regular evaluation and could signal changes in performance. Depending on the risk of the AI systems and the extent of the changes, the appropriate validation must be agreed by the developer and the regulator. Analytical validation against previously unseen datasets- or benchmarking against approved datasets representative of the intended setting or population could be useful in this scenario.

1. ***Low- and middle-income countries***

There is considerable variation in the experience of international regulatory bodies with deployed AI technologies and developed AI systems. Some countries also lack a dedicated national regulatory body. The WG-RC meetings have provided a forum for the sharing of expertise and discussion of common problems, including regulatory bodies and other interested stakeholders, some with aligned remits. Furthermore, there are important regulatory concepts related to the intended use and analytical and clinical validation of AI systems in health. First, in low- and middle-income countries, one of the potential uses of AI technologies is in bringing specialized AI-based systems or knowledge to areas which do not have the relevant medical specialist (for example interpreting retinal scans, histopathology slides or radiology images). In high income countries, AI systems are more often positioned as an adjunct to medical professionals. Leveraging the evaluation performed to support regulation in a high income setting to inform how such AI systems are used in low- or middle-income settings therefore may not be appropriate. Thus, the full context of healthcare infrastructure and resources should be considered. Second, some regulatory bodies rely on decisions from other bodies to support their regulatory work. Given that the performance of AI systems may be highly context dependent, additional steps may be required. There is a concern that developers may not ensure adaptation or evaluation for resource limited settings if the market is less attractive. Regulatory agencies in high income countries could support this adaptation, which could also increase generalisability and robustness of AI systems, but requires adaptive studies to ensure wider use in low and middle income countries or by using incentives to encourage additional development, testing and validation. The availability of a range of representative datasets would support local or national analytical validation. Finally, health AI systems can be highly sensitive to shifts in data distribution and features. They may therefore be sensitive to differences in disease prevalence when moving from high to low-income counties, with the possibility of lower performance without appropriate evaluation or tuning with local data.

## Data quality

* ***Data in current health ecosystems***

The health sector has been very receptive to the benefits of AI thanks to the explosion of data and accessibility to computational power. Data is the most important ingredient for training AI/ML algorithms, and can be classified based on format, structure, volume, and many other factors. It can take any form, including: character, text, words, numbers, pictures, sound, or video. Also, these data can be either structured, semi-structured, or unstructured[[64]](#footnote-64). Structured data is normally stored in databases that are structured in a manner that follows a specific model or scheme such as data stored in electronic medical records, mobile devices, and Internet of Things (IoT) devices. Regardless of the format, structure, volume of the data, a more general classification can be based on the following 10Vs[[65]](#footnote-65) (as illustrated in Figure 1 below): Volume, Veracity, Validity, Vocabulary, Velocity, Vagueness, Variability, Venue, Variety, and Value.

* ***Good quality data in health AI systems***

All AI tasks and solutions use some form of data regardless of its characteristics to facilitate machines to learn, adapt, and improve of their learning. However, data quality greatly influences the success of such solutions' safety and effectiveness. Good quality data is an ambiguous term that is open to misinterpretation. Therefore, gaining a good understanding of the datasets used, for example, from the 10Vs perspective (mentioned above) is crucial to assess data quality in AI systems during development and even after. The section below highlights key challenges and considerations for all stakeholders, including developers and regulators when handling data in AI systems in order to achieve good data quality.

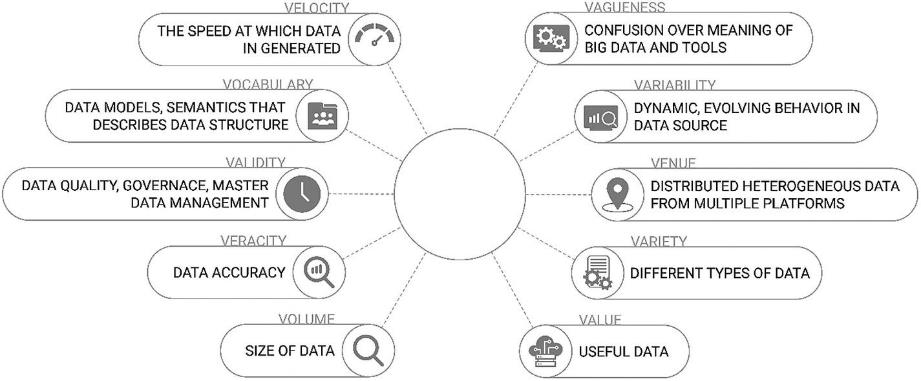


Figure 11 – The 10Vs of data[[66]](#footnote-66).

* ***Key quality data challenges and considerations for health AI systems***

The availability of good quality datasets that are clinically relevant is one of the key challenges that developers face. However, data of varying quality can still be used depending on purpose, and thus, developers should determine if available data is of sufficient quality to support the development of systems that can achieve their intended goal. The lack of good quality datasets to be used in the development of AI systems may hinder their effectiveness and potential benefits. Data that is not of sufficient quality for the intended purpose can also lead to many issues, such as bias and errors. In this section, some data quality issues that often arise when developing AI systems and need to be considered by all stakeholders are discussed and summarized in Table 1. These issues and considerations can relate directly to dataset management, ML model, infrastructure used to manage the data, or general governance aspects:

1. **Dataset management.** When managing datasets for ML models, a clear data management plan should be pre-specified and well-documented. Data management approaches should be risk based and fit for purpose. This may include, but is not limited to, data selection volume (including volume of data used and volume of available data), splitting, cleansing (including any AI algorithms that were used to clean the data), data usability (including how well the dataset is structured in a machine-readable format), labelling, dependencies, augmentation, and streaming. If data augmentation is relevant, it is important to develop a clear data augmentation strategy. The developers should also consider putting in place good data accountability practices for those handling the data to ensure quality and integrity of data is maintained throughout the lineage of data. This is also essential for knowledge management and transfer in a highly evolving field. Finally, in addition to the handling of the data, the capacity to plan for and conduct data analyses is also an important consideration.
2. **Data inconsistency.** High heterogeneity in the syntax of the data may require harmonization in order to address issues related to multiple data sources with varying standards, formats, schemas, structures, and ambiguous semantics into a coherent dataset for the purpose of its comprehensive analysis, which is especially challenging when using healthcare data. For example, much of the data collected from various information silos is inconsistent, incompatible, or not executable in machine-readable formats. For multiple data sources, there may be variations in how the data are captured (e.g., definitions of individual variables).
3. **Dataset selection and curation.** Knowing the source of data and an initial assessment of the data quality can help to determine the potential for selection and information bias. Selection bias results when data used to produce the model are not fully representative of the actual data or environment that the model may receive or function in. In addition to selection bias, measurement bias is another relevant aspect that results when the data collection device causes the data to be systematically skewed in a particular direction. Therefore, developers should be aware of data quality limitations when attempting to curate and utilize these large-scale datasets. Moreover, developers and regulators need to know where the data originally came from and how it was collected and curated. This is especially important when the datasets are from an open-source database where the original source and specifications of the dataset may not be available. When the origin of data is difficult to establish, it would be prudent for developers to assess the risks of using such data and manage them accordingly. Finally, even if datasets are collected from reliable sources, mitigation of bias and assessment and mitigation of other risks to data robustness remain essential for a heterogeneous dataset.
4. **Data usability.** Knowing if the data used for development of the algorithm was intended for that training is essential, as developers need to convey their full understanding of the dataset and why it was suitable for their purpose. For example, data from a third-party source may be representative data intended for training purposes (e.g. case studies in tertiary education) and may not be suitable for training of an AI model intended to diagnose a disease or condition.
5. **Data integrity.** Data Integrity can be defined as "the completeness, consistency, and accuracy of data"[[67]](#footnote-67). Lack of data integrity is also important issue. This can be best understood by how well extraction and transformation have been performed on the dataset. To maintain data integrity, data verification checks may be developed. Data verification checks are a key component of data quality assurance when utilizing real world data. Data verification checks should also be the first step of data preparation for any ML workflow.

Timeline

Description automatically generated with medium confidence

Figure 12 – Examples of quality check principles[[68]](#footnote-68).

1. **Model training.** AI-algorithms are usually trained on a separate dataset, named the training dataset and validated on a separate dataset, to reliably measure the performance of the algorithm. The training datasets should be well represented (for instance by considering prevalence of disease/condition) to avoid 'class imbalance'. Medical record data is inherently biased, and thus, there is a need to incorporate non-medical data such as the social determinants of health[[69]](#footnote-69). Furthermore, underrepresentation of important diagnostic features may limit performance of the model and cause bias. This can be avoided by ensuring inclusion and exclusion criteria at the patient level and input data level do not create a selection bias. Furthermore, ensuring the datasets are reflective of the setting in which the model will be applied - a lack of diverse data (age, race, geographic areas) could limit the generalizability and accuracy of developed AI system. This is demonstrated by a recent study by Stanford University[[70]](#footnote-70) that shows 71% of patient data from just three U.S. states train most AI diagnostic tools used in the USA.

* **Data labelling.** It is important to ensure consistent, reliable, and accurate labeling of datasets for testing following good practices. In cases where subjective reference standards are used, quality will be influenced by many factors, such as the independence and qualifications of the graders, the number of graders per label, whether the reference standard is validated to correlate to patient outcomes, and whether the reference standard follows published metrics.
* **Documentation and transparency.** Often the algorithm and data supporting it are not available, or not well documented for all AI system stakeholders. Therefore, this makes it difficult to assess the quality of the underlying data. Transparency and careful documentation are important not only on the methodology used in the collection of data, but also for the actual selection and modifications of datasets used for training, validation, and testing. Therefore, good documentation is fundamental to achieve transparency that would enable verification and traceability. Transparency of methods should ensure data quality. Beyond CONSORT-AI and SPIRIT-AI reporting guidelines already mentioned, specific checklists for reporting representativeness, completeness, and other data quality characteristics have been devised by the machine learning community[[71]](#footnote-71),[[72]](#footnote-72).

Eventually, developers should consider deploying rigorous pre-release trials for AI systems to ensure that they will not amplify any of the issues discussed above like biases and errors due to any issues with the training data, algorithms, or other elements of system design. Furthermore, careful design or prompt troubleshooting can help identify data quality issues early on. This could potentially prevent or ameliorate possible resulting harm. Finally, to mitigate data quality issues that arise in healthcare data and the associated risks, stakeholders should continue to work to create data ecosystems to facilitate the sharing of good-quality data sources.

Table 3 is a classified list that summarizes the key data quality considerations for AI systems' safety and effectiveness[[73]](#footnote-73).

Table 3 – General data quality considerations.

| Category | Data quality consideration item |
| --- | --- |
| Dataset | Splitting |
| Selection volume and size |
| Selection bias |
| Individual variables definitions in each dataset |
| Raw data vs "cleaned" data |
| Data wrangling and cleansing |
| Parameters and hyperparameters |
| Usability |
| Characterization |
| Labelling |
| Dependencies |
| Augmentation |
| Manipulation |
| Streaming |
| Interfaces |
| Integrity |
| Unique requirements |
| Data source |
| Data infrastructure | Storage size |
| Storage format |
| Transformation medium |
| AI/ ML model | Data Training |
| Tuning Data |
| Verification set |
| Validation set |
| Testing |
| Development set |
| Static AI vs dynamic AI |
| Open AI vs closed AI |
| Governance management | Liability |
| Data access |
| Risk Management |
| Data Protection |
| Privacy |
| Adoption education for clinical practice |
| Good practices |
| Standards (of care, governance, interoperability, etc.) |
| Scope of practice and AI model use |
| Technical checklist |
| Documentation |
| Transparency |

## Privacy and data protection

The WHO Global Strategy on Digital Health 2020–2025 classifies health data as sensitive personal data, or personally identifiable information, that requires a high safety and security standard. Therefore, it emphasizes the need for a strong legal and regulatory framework to protect the privacy, confidentiality, integrity, availability, and processing of personal health data. A responsive legal and regulatory framework can also address issues of cybersecurity, trust building, accountability and governance, ethics, equity, capacity building, and literacy. This will help ensure that good quality data are collected and subsequently shared to support planning, commissioning, and transformation of services.

To develop and maintain adequate data security strategies, it is important for AI system developers, deployers, and manufacturers to understand the thickening web of privacy and data protections laws. This section discusses high-level considerations for privacy and data protection. For other ethical considerations, you may want to refer to the deliverable of the Working Group on Ethical Considerations on AI for Health[[74]](#footnote-74).

* ***Current Landscape***

As the demand for health-related data increases, protecting privacy is creating a unique challenge for all stakeholders wishing to benefit from the many opportunities of AI systems and technologies. One of the main reasons for this is that the high dimensionality of big data could make it difficult to apply anonymization and de-identification methods. Additionally, securing large-scale data sets against unauthorized access at each stage of the development process – collection, storage and management, transport, analysis, sharing, and destruction– is an important consideration.

Approximately145 countries and regions have data protection regulations and privacy laws regulating the collection, use, disclosure, and security of personal information[[75]](#footnote-75). There are many different definitions and interpretations of "data protection" and "privacy." In some cases, data protection and privacy are used interchangeably. However, although these concepts are similar and often overlap, their meanings are different, and developers should be aware of the legal and ethical implications that result from these differences.

Laws and regulations that cover "the management of personal information" are typically grouped under "privacy policy" in the United States and under "protection policy" in the EU and elsewhere. These laws are often complex and may have conflicting obligations. When developing an AI system for therapeutic development or healthcare applications, early in the development process developers should consider gaining an understanding of applicable data protection regulations and privacy laws, including special regulatory provisions related to sensitive data, such as genetic data. Developers should also consider national, as well as regional laws. For example, in the United States, although the Health Insurance Portability and Accountability Act (HIPAA) sets a baseline for protecting health data, states are empowered to enact stricter privacy laws (e.g., California Consumer Privacy Act of 2018).

It is important to understand the jurisdictional scope of the various laws. For instance, because the scope of the GDPR is broad and its impact is significant, companies may want to at least consider the possibility and evaluate the extent to which they are subject to it. Most privacy laws, including PDPC, only apply to personal data processed within the country, whereas the GDPR[[76]](#footnote-76) may apply to the personal data of European Union (EU) data subjects, regardless of the location where data is processed.[[77]](#footnote-77) As a result, companies subject themselves to compliance obligations under the GDPR if they are located in the EU (including if any component of that organization is located in the EU); offer goods and services to individuals located in the EU; or, monitor the behavior of individuals located in the EU.

It is also important for developers to understand the varied legal contexts and requirements for privacy-related concepts, such as "identifiable," "anonymous," and "consent." For example, Chapter One of the UK's draft anonymization, pseudonymization, and privacy enhancing technologies guidance warns that referring to datasets as "anonymized" when they still may contain personal data in a pseudonymized form poses the risk of violating the UK data protection law in the mistaken belief that the processing does not involve personal data[[78]](#footnote-78). Consent requirements also vary depending on the jurisdiction. For example, various jurisdictions may require "explicit consent," with heightened information requirements for the processing of health-related data[[79]](#footnote-79). Therefore, developers may want to consider the varied legal contexts when documenting how they address privacy-related concepts, including measures taken to meet consent requirements and the how they define anonymous or identifiable information.

In addition, certain jurisdictions have data protection regulatory frameworks that introduce reciprocity-based rules and place restrictions on the movement or transfer of data across borders. This might have a significant impact on the way in which data is processed and shared between countries. These provisions serve to curtail transnational data flows into and out of areas that are considered not to provide an "adequate" level of data protection.

Adequacy assessments may be required to determine if a recipient-country has thresholds of data protection laws and protections "essentially equivalent" or "substantially similar" to the jurisdiction from which the data was transferred. The GDPR, as a significant driver of emerging global data protection regimes, provides that the free transfer of personal data to third countries, non-European Union member states, can primarily occur where such third country is considered by the EU Commission as having an "adequate" level of protection[[80]](#footnote-80). To date, the EU Commission has only recognized 13 countries as providing adequate protection[[81]](#footnote-81).

Therefore, developers should be aware of the nuances of the different jurisdictions regulations and laws and consider documenting their data protection practices accordingly. In general, companies should consider keeping current on new laws and requirements, leveraging governance, risk analysis, policies, trainings, and other strategies in a comprehensive and coherent way.

* ***Documentation and Transparency***

Documentation and transparency are critical to facilitate trust regarding privacy and data protection. Detailed privacy policy disclosures provide regulators with a benchmark by which to examine a company's data handling. These disclosures should identify significant uses of personal information for algorithmic decisions. Depending on the jurisdiction, the disclosures may require the inclusion of other relevant information, such as the types of health data collected and processed; the sources of the health data collected and processed; the identity of the persons or organizations which determined the purpose or means of processing personal data; the identity of the person or organization which processed the data; the legal bases for processing the data; how the data was collected, including whether adequate notice was provided to the data subject and how consent requirements were met; and, technical and organizational information concerning the storage of data, including security measures.

Developers should take privacy into account as they design and deploy AI systems. This includes designing, implementing, and documenting approaches and methods to ensure a quality continuum across the development phases to protect data privacy[[82]](#footnote-82). Privacy protections should not just be limited to addressing cybersecurity risks, especially since some privacy risks (e.g., harms to one's dignity, which may cause embarrassment or stigma, or more tangible harms, such as discrimination, economic loss, or physical harm)[[83]](#footnote-83) can also arise by means unrelated to cybersecurity incidents. Therefore, when developing solutions to address risks, developers should have a general understanding of the different origins of cybersecurity and privacy risks and develop their risk management practices accordingly (see Figure 13 below[[84]](#footnote-84)).

Diagram, venn diagram

Description automatically generated

Figure 13 – NIST Privacy Framework – Cybersecurity and Privacy Risk Relationship[[85]](#footnote-85)

A compliance program should consider risks and develop privacy compliance priorities that take into account any specific potential harm, as well as the enforcement environment. Developers may want to consider including in their documentation a description of the operations involved in the processing of personal data, a risk assessment, and the measures implemented to mitigate risks that take into account the interests of data subjects.

Certain regulations outline prescriptive security requirements to address cybersecurity and privacy risks, such as the GDPR's data protection by design and default[[86]](#footnote-86) and India's proposed data privacy by design policy[[87]](#footnote-87), while others include the duty to implement and maintain reasonable security practices and procedures appropriate to the risk.[[88]](#footnote-88) Privacy frameworks often include privacy impact assessments, which are a widely used privacy management tool to proactively evaluate and mitigate privacy risks. Some jurisdictions, including the EU[[89]](#footnote-89), require companies to conduct these assessments[[90]](#footnote-90). Although the U.S. law does not require privacy impact assessments, the U.S. Department of Commerce's National Institute for Standards and Technology's (NIST) Privacy Framework recommends that developers conduct them. According to NIST, "identifying if data processing could create problems for individuals, even when an organization may be fully compliant with applicable laws or regulations, can help with ethical decision-making in system, product, and service design or deployment" [[91]](#footnote-91).This in turn can increase trust in the system.

Developers may also want to consider annotating their AI and having audit trails that explain what kinds of choices are made during the development process. Annotated notes provide "after the fact" transparency to outside parties and can help to explain the manner in which privacy was embedded, if applicable[[92]](#footnote-92). Such explanations and documentation should be at different levels of detail, targeted at different audiences – regulators, managers, developers, operators, and users. The nature of the information and explanations required may be different, but all of the assumptions, constraints, data sources, expected input and output, and major risks and limitations at each level should be clearly documented. In addition, an audit trail not only shows that controls have been applied, it could also potentially show how damage was mitigated in the case of a data breach.

Many jurisdictions enforce certain cybersecurity requirements or publish guidance on cybersecurity for developers of medical devices to consider. Although an in-depth discussion of cybersecurity requirements is outside the scope of this subsection, it is important to understand the key role that cybersecurity plays in the protection of personal health information. Cybersecurity focuses on specific technical implementations needed to protect systems and networks against cyberattacks, which could compromise both the security of health-related systems and data as well as an individual's privacy, which could result in harm. To provide transparency about cybersecurity practices, developers may want to consider documenting practices and approaches for data security, including policies that help protect the confidentiality, integrity, and availability of personal data throughout its lifecycle, such as appropriate encryption, access controls, logging methods, risk monitoring, and methods of secure destruction. Developers may also consider documenting systems and approaches used to protect against data manipulation and adversarial attacks[[93]](#footnote-93).

## Engagement & collaboration

Engagement and collaboration, where applicable and appropriate, among developers, manufacturers, healthcare practitioners, patients, patient advocates, policymakers, regulatory bodies, and other stakeholders can improve the safety and quality of an AI system. This section focuses on the engagement and collaboration approaches of regulatory bodies with stakeholders in the area of AI in healthcare and therapeutic development. Although many regulatory bodies have adopted engagement and collaboration approaches in this area, this section discusses the approaches of five regulatory authorities: the MHRA, South African Health Products Regulatory Authority (SAHPRA), European Commission (EC), Singapore Health Sciences Authority (HSA), and US FDA. Table 4 lists examples of with whom, why, and how these regulators foster engagement and collaboration. The examples are not meant to be comprehensive, but instead are intended to highlight general approaches. Table 4 is followed by an analysis that discusses the similarities and differences in the approaches.

The second subsection examines two examples of engagement and communication between regulators and AI developers resulting in positive clinical outcomes (CURATE.AI and IDentif.AI). The last subsections, consider the practical implications for engagement and collaboration in resource-limited settings, and recommend ways that regulatory bodies in countries without past experience in engagement and collaboration can initiate this process. This is supplemented by several narratives: how to apply engagement tools (based on experience) and how to position the regulator as a partner in the context of accessible dialogue, guidance and recommendations during the development process.

Table 4 – Examples of regulators' approaches to engagement and collaboration with stakeholders about the use of AI in healthcare and therapeutic development

|  |  |  |  |
| --- | --- | --- | --- |
| Regulator | With Whom? | Why? | How? |
| 1. MHRA  (United Kingdom) | *Examples of stakeholders with whom MHRA engages and collaborates:*   Patients/Patient Advocates   Academia   Healthcare Professionals  (e.g., providers in the National Health Service (NHS) and private healthcare providers)   Industry  (e.g., medical device and in vitro diagnostic industry, health tech industry)   Domestic Government Partners (e.g., Department of Health and Social Care (DHSC), NHS England and Improvement, National Institute for Health and Care Excellence (NICE), and Care Quality Commission (CQC)) | *Examples of reasons why the MRA engages and collaborates with stakeholders:*   Alert users to problems with medical devices and medicines   Answer enquiries about role in regulation or raise awareness of safety issues   Seek feedback on development of regulatory policy, managing adverse incidents and risks   Interface with UK government and NHS including stakeholders aligned to digital and AI-related activities | *Examples of ways in which the MHRA engages and collaborates with stakeholders:*  ● Central alerting system to the NHS and healthcare providers or through professional groups  ● Media, public, and other stakeholder inquiries via MHRA Customer Service Centre, dedicated email inboxes, and Press Office  ● Connecting with Expert Advisory Groups, networks, and stakeholder groups on specific issues  ● Consultation on engagement with patients and public[[94]](#footnote-94)   Working-level meetings with national stakeholders, bilateral meetings with other parts of NHS, government and international counterparts |
| 2. SAHPRA (South Africa) | *Examples of stakeholders with whom SAHPRA engages and collaborates:*   Patients/Patient Advocates   Academia   Healthcare Professionals   Industry  (e.g., manufacturers/ distributors, trade associations)   National Government Partners  (e.g., National Department of Health, National Department of Trade & Industry, South African National Accreditation Service) | *Examples of reasons why the SAHPRA engages and collaborates with stakeholders:*   Facilitate the approval of innovative AI systems   South African National Accreditation System (SANAS) to ensure Conformity Assessment Body network is established in country to certify quality management system (QMS) | *Examples of ways in which the SAHPRA engages and collaborates with stakeholders:*  ● The framework for engagement and collaboration has not yet been formalized   Recommended that stakeholder engagement adopt the five-step engagement model developed by the Therapeutic Goods Administration (TGA) Australia[[95]](#footnote-95) |
| 3. EC (European Union) | *Examples of stakeholders with whom the EC engages and collaborates:*   Patients/Patient Advocates   Academia   Healthcare Professionals | *Examples of reasons why the EC engages and collaborates with stakeholders:*   To "support the Commission in the development of actions for the digital transformation of health and care in the EU." | *Examples of ways in which the EC engages and collaborates with stakeholders:*   By providing "advice and expertise to the Commission, particularly on topics set out in the communication[[96]](#footnote-96) on enabling the digital transformation of health and care in the Digital Single Market, that was adopted in April 2018." In particular, such topics regard health data interoperability and record exchange formats, digital health services, data protection and privacy, AI, and "other cross cutting aspects linked to the digital transformation of health and care, such as financing and investment proposals and enabling technologies." |
| 4. HSA (Singapore) | *Examples of stakeholders with whom the HSA engages and collaborates:*   Academia (e.g., research institutions)   Healthcare Professionals   Industry  (e.g., software and AI developers, trade associations)   National Government Bodies | *Examples of reasons why the HSA engages and collaborates with stakeholders:*  ● Early engagement and support to innovators to facilitate regulatory compliance thus facilitating timely access to safe innovations for patients   Actively consult on new policies and guidelines related to AI and software medical devices to receive and incorporate stakeholders' inputs and perspectives Regulatory Guidelines for Software Medical Devices – A Life Cycle Approach[[97]](#footnote-97)   To work with other agencies responsible for implementation and deployment of AI and software medical devices in healthcare system to facilitate greater adoption of innovative technologies in the healthcare system |  Rapid, streamlined engagement portals are available for several facets of product regulation[[98]](#footnote-98)   Specific processes that can be straightforwardly addressed include Medical Device Information Communication System (for application submissions for licenses, permits, registrations, etc.)   Online self-help tools to determine the product classification and risk classification for medical devices and simple forms to seek advice and confirmation from the HSA[[1]](https://euc-word-edit.officeapps.live.com/we/wordeditorframe.aspx?ui=en-us&rs=en-us&wopisrc=https%3A%2F%2Fworldhealthorg.sharepoint.com%2Fsites%2FAI4HWGRC%2F_vti_bin%2Fwopi.ashx%2Ffiles%2F906e1d4253d14ebd88b6f1058a0586ce&wdenableroaming=1&mscc=1&hid=-341&uiembed=1&uih=teams&hhdr=1&dchat=1&sc=%7B%22pmo%22%3A%22https%3A%2F%2Fteams.microsoft.com%22%2C%22pmshare%22%3Atrue%2C%22surl%22%3A%22%22%2C%22curl%22%3A%22%22%2C%22vurl%22%3A%22%22%2C%22eurl%22%3A%22https%3A%2F%2Fteams.microsoft.com%2Ffiles%2Fapps%2Fcom.microsoft.teams.files%2Ffiles%2F965658113%2Fopen%3Fagent%3Dpostmessage%26objectUrl%3Dhttps%253A%252F%252Fworldhealthorg.sharepoint.com%252Fsites%252FAI4HWGRC%252FShared%2520Documents%252FGeneral%252FWG-RC%2520DEL02%2520Versions%252FWG-RC_DEL02-v2-With-Tracked-Changes.docx%26fileId%3D906e1d42-53d1-4ebd-88b6-f1058a0586ce%26fileType%3Ddocx%26ctx%3Daggregate%26scenarioId%3D341%26locale%3Den-us%26theme%3Ddefault%26version%3D21021008600%26setting%3Dring.id%3Ageneral%26setting%3DcreatedTime%3A1618040598657%22%7D&wdorigin=TEAMS-ELECTRON.aggregatefiles.aggregate&wdhostclicktime=1618040598334&jsapi=1&jsapiver=v1&newsession=1&corrid=ad11001e-6e86-4468-b462-baacfb513824&usid=ad11001e-6e86-4468-b462-baacfb513824&sftc=1&sams=1&accloop=1&sdr=6&scnd=1&hbcv=1&htv=1&hodflp=1&instantedit=1&wopicomplete=1&wdredirectionreason=Unified_SingleFlush&rct=Medium&ctp=LeastProtected#_ftn1)   Medical Device Development Consultation: Online appointment booking system that allows innovators and developers to seek scientific and regulatory advice during medical device development phase to facilitate regulatory compliance   Online stakeholder consultation process for all new and revised policies and guidelines   Regular focus group discussions and engagements with industry associations and companies |
| 5. FDA (United States) | *Examples of stakeholders with whom the FDA engages and collaborates:*   Patients/Caregivers/Patient Advocates   Academia (e.g., research institutions)   Healthcare Professionals   Industry (e.g., developers, device manufacturers, drug companies, trade associations)   National Government Partners (e.g. NIH, ONC, FCC)   Foreign Government Partners   International Organizations  (e.g., IMDRF, ICH)   Consumers/General Public | *Examples of reasons why the FDA engages and collaborates with stakeholders*   Facilitate patient access to technologies that can benefit them in a timely manner   Support novel, innovative medical product development through early interactions with stakeholders   Provide timely feedback on FDA policies to reduce uncertainty   Communicate to the public about AI/ML devices   Receive feedback on policies, guidance, and discussion papers | *Examples of ways in which the FDA engages and collaborates with stakeholders:*   Hold different types of pre-submission meetings to provide early feedback to sponsors   Participate and lead international harmonization efforts (e.g., IMDRF, ICH)   Engage as members of Public-Private Partnerships and Collaborative Communities   Collaborate in pre-competitive space on regulatory science research to advance scientific community understanding   Receive formal comments on policies and guidance through the Federal Register   Hold workshops and other engagement events to get feedback from patients, industry, and other stakeholders |

* ***Discussion about strategies of profiled regulatory bodies***

Within the series of tables, we find the approaches of four national bodies and one regional (in the case of EC) regulatory body to foster engagement and collaboration. In the first category ("with whom"), there are considerable similarities among these bodies. The shared targets for engagement and collaboration include health professionals (indicated by US FDA, SAHPRA, MHRA, EC, and HSA), academia (US FDA, SAHPRA, MHRA, EC, and HSA), industry (US FDA, SAHPRA, MHRA, EC, and HSA), patients or patient advocates (US FDA, SAHPRA, MHRA, and EC), domestic government bodies (US FDA, SAHPRA, and MHRA), media (national and trade press; US FDA and MHRA), health providers (US FDA and MHRA), and consumers (US FDA and MHRA). Interestingly, the strategy paper by the U.S. Department of Commerce's National Institute of Standards and Technology (NIST)—provided under "Additional sources of relevant literature" at the end of this section—also refers to academia and domestic government bodies as targets for engagement and collaboration.

In the second category ("why"), SAHPRA mentions the importance of communicating the benefits and intended use of devices, presumably to protect and promote public health (listed by US FDA and implied by MHRA). US FDA also stresses the importance of bilateral communication with stakeholders so that regulators are aware about developments in industry (or academia) and so that these stakeholders, in turn, are aware about developments in regulation. Similarly, MHRA indicates the importance of acquiring feedback about medical devices from stakeholders. This supports the objectives given by both SAHPRA and EC: to facilitate approval of innovative solutions and support the digital transformation of health and care. The Health Sciences Authority (HSA) acknowledges the importance of early engagement with the innovators and developers to provide greater clarity in regulatory requirements and improve transparency in regulatory processes.

For the third category ("how"), US FDA lists steps that are taken to foster engagement (e.g., hosting workshops, producing digital and print material, and offering training modules or other types of education). MHRA also notes the importance of holding meetings with stakeholders (including domestic government institutes and international counterparts). HSA has introduced Pre-market Consultation Scheme to support innovation and device development by providing scientific and regulatory advice to enable regulatory compliance by software and AI developers, who unlike traditional medical device players are not familiar with regulatory requirements[[99]](#footnote-99),[[100]](#footnote-100).

* ***Two successful instances of engagement***

To understand the value of engagement and collaboration among regulatory bodies and stakeholders, we provide two real-world examples (hereafter, Cases 1 and 2). Clear avenues for engagement between regulators and AI developers play a major role in ensuring that rigorous evaluation and accelerated delivery of impactful modalities can be seamlessly realized. One aspect is the area of interventional AI/digital medicine, which involves the application of software/devices (e.g., AI-based drug development and/or dosing platforms) and/or the application of resulting drug compounds and/or combinations recommended by these platforms[[101]](#footnote-101),[[102]](#footnote-102),[[103]](#footnote-103). In this context, integrating regulator accessibility with emerging innovation, sometimes under urgent circumstances, will ultimately result in life-saving outcomes. Importantly, these outcomes will not solely be confined to post-approval treatment, but to substantial patient benefit during the investigational stages of validation as well.

In Case 1, the developmental roadmap and validation of CURATE.AI and foundational technology of IDentif.AI was discussed with the Medical Devices Branch[[104]](#footnote-104) of the Health Sciences Authority (HSA) in Singapore. This interactive session included an in-depth review of the key findings of the technology platforms, the process of implementing both platforms, emerging statistical analysis strategies to effectively assess personalized medicine treatment outcomes, and regulatory routes. A broader discussion pertaining to how clinical trial designs themselves may evolve due to the emergence of AI was also conducted[[105]](#footnote-105),[[106]](#footnote-106),[[107]](#footnote-107). A clear pathway for subsequent inquiries was established, as multiple and frequent guidance requests were expected due to the nature of the trial designs that were envisioned. These included *N*-of-1 study designs for a broad range of indications designed for each patient. Specifically, these designs were personalized based on (e.g.,) the individualized dosage calibrations of the drug regimen (clinician-selected regimen), serial efficacy and toxicity measurements, efficacy-guided treatment protocol, and safety parameters.Subsequent submissions have included engagement with regulators for risk classifications associated with the device for each trial and subsequent discussion for submission of special access routes (SARs) for the potential rapid implementation of [[108]](#footnote-108) 1-of-1 trials and for treatment purposes if needed.Rapid and informative responses and active engagement from HSA regulatory team members resulted in efficient turnaround times for trial initiation, which ultimately resulted in a positive outcome for a refractory oncology patient. A sustained track record of engagement with the regulatory community has played a key role in helping a clear process flow to be developed for downstream guidance requests.

Case 2 was developed in response to the COVID-19 pandemic. Specifically, a patient-derived live virus strain was harnessed for IDentif.AI-driven combination therapy optimization to serve as a clinical decision support system (CDSS). Contrary to traditional AI-based approaches, this strategy did not use pre-existing patient datasets. Instead, prospective experimentation was used alongside an AI-derived small data analytics strategy to pinpoint prospective data-backed rankings of combinations for potential further clinical consideration and potentially for the elimination of certain combinations from further clinical consideration. The foundational technology for IDentif.AI was previously discussed in detail with the HSA Medical Devices Branch, and additional IDentif.AI SARS-CoV-2 study information was provided in the context of clinical decision support, developing optimized combinations pinpointed by IDentif.AI, and potential trials being designed with clinical partners. With regards to regulator engagement, the Medical Devices Branch of the HSA was contacted to provide device risk classification guidance for the submission of a Clinical Research Materials Notification (CRM-N) for study purposes. Obtaining a CRM-N is a required segment of the submission of a clinical validation program because it stipulates a prerequisite of an initial assessment of device risk from the HSA[[109]](#footnote-109). Of note, the submission portal and portal interaction were particularly straightforward to navigate and integrated with a uniform access portal, which was streamlined for efficient oversight and monitoring with regulatory bodies. This further demonstrates the straightforward process of interaction with the HSA. This case served as an example of the critical importance of straightforward regulator accessibility and the profoundly positive impact that this can have on the advancement of promising technologies towards further clinical assessment and validation.

* ***Recommended approaches for countries without past experience***

For countries with limited experience in engagement and collaboration (and/or limited resources), several considerations should be made. For instance, what levels of engagement and collaboration are desired and what steps can and should be taken to achieve this? Also, what challenges are presented by this technology (e.g., AI explainability)?

In many cases, it is desirable to adopt regulatory models that are adaptable, flexible, modular, and scalable, to accommodate the uncertainties of innovation through appropriate forms of oversight and coordination. These features not only fit the specific challenges of emerging technologies, but also the regulatory approach of countries without past experience in this field, or with scarce economic resources. Priorities, on the one hand, should be scalable, so that growing amounts of work can be suitably addressed by adding resources to the regulatory model. On the other hand, priorities should be determined in accordance with the modular adaptability of the steps and levels of engagement. In ecology, adaptability regards the ability to cope with unexpected disturbances in the environment. In engineering, modularity refers to the interrelation of the separate parts of a software package or also to the partitioning of the design to make it manageable. In multi-agent systems (MAS), it refers to the efficient usage of computational resources. We can profit from this notion to create adaptable policies that can be combined into regulatory systems for legal governance. The aim should be to address the uncertainties of innovation, aligning with society's preferences on emerging innovation, while allowing regulators to capture expanding understanding of technological challenges with increasing normative granularity.[[110]](#footnote-110)

* ***Narrative on using engagement tools in practice based on practical experience of using them***

For all countries—from those with limited experience in engagement and collaboration (and/or limited resources) to those on the other end of these spectra—project and programme management tools can help organizations (including regulators) structure and execute their engagement with stakeholders and users. No matter which tool is chosen, the key to valuable engagement is investing the time, energy, and thought into how best to engage stakeholders and following through on that engagement for the duration of a project or program. Often, engagement fails if the investment is seen as a short- rather than long-term relationship.

The Australian Government's recommended five-step model for engagement[[111]](#footnote-111) is a good starting point for considering how a regulator could engage with developers of AI health products and services. In this model, engagement starts with thinking through the purpose of engagement (based on what it is hoped to achieve) and identifying relevant stakeholders. When planning out the different levels of engagement with stakeholders, it is recommended to map out existing relationships and to define the type of engagement and relationship that is needed with the stakeholder (and what type of relationship they would be open to having). For example, a digital health developer building an app to support parents with children above a healthy weight may find the primary health body is an influential stakeholder because it sets policies around managing children's weight. However, it is not a body with whom the developer of the app needs to engage regularly, so the developer may only "inform" the health body of the project. Contrarily, a developer will want to work with parents of children above a healthy weight to co-design the app and ensure it fits their needs. It would, therefore, be important for the developer to "collaborate" with a representative group of parents and establish two-way/multi-way communication and shared learning and decision-making over the course of the project.

Another similar approach for making sure that stakeholders are provided with the right information at the right time and using the optimal communication channels is outlined by the leading product development software company Atlassian[[112]](#footnote-112). Within the stakeholder communication "play," importance is placed on who the stakeholders are, the desired method of communication, and the frequency of communication. For example, for an internal government project developing a digital health product, there will be internal stakeholders (like funders of the project and policy leads) and external stakeholders (like leading academics). The communications plan should outline how each stakeholder group will be addressed (e-mail, face-to-face conversations, video call, and/or social media) and how often will be the contact with the stakeholder group (daily, fortnightly, and/or yearly) based on what the relationship with the stakeholder brings to the overall goals (i.e., information sharing, co-design, and/or quality assurance). This plan can then be mapped out in a simple table (example headings: method, audience/stakeholder, content to share, why, and frequency) for the whole development team to follow.

* ***Narrative positioning the regulator as a partner in the development process***

As demonstrated in the tables and discussed in the subsequent text, multiple regulatory bodies emphasize the importance of open (bilateral) communication with stakeholders so that regulators are aware about developments in AI-based technology developments and so that these stakeholders, in turn, are aware about changes in regulation. This is because AI-based technology is constantly changing and regulation needs to be able to keep pace with its iterative nature. The development, deployment, post-market surveillance, and iteration of AI products and services in healthcare should, therefore, be an ongoing conversation between developers and regulators.

It is recommended that regulators look at AI-based technology in healthcare from a mindset of accessible engagement that potentially, when applicable, enables working alongside the developer to ensure compliance with regulatory requirements throughout the development and implementation process. An engagement mindset approach to regulation is about building trusting, collaborative relationships between developers and the regulatory body(s) along with a two-way dialogue that enables developers to learn from regulators and vice/versa.

Furthermore, depending on a country's regulatory arrangements one or more regulators may be responsible for AI-based health products and services. This means a developer often has to work with (and meet the standards of) more than one regulatory body. To ensure that this is a smooth and positive experience for AI developers, it is again recommended that regulators take a service approach. By this, it is meant that a single, clearly marked pathway should be established to be followed by an AI developer when ensuring the compliance of a product or service. Regulators need to collaborate with each other on clear messaging to developers, consistent levels of engagement with developers at the right point, and sharing learnings about different engagements with developers.

If a country wants to take an accessible engagement mindset approach to regulation of AI products and services one step further, co-regulation, could be explored. As outlined by Clark[[113]](#footnote-113), with a coregulation approach regulators outlined a regulatory framework based on needed compliance to the legislative act(s) and the detail of how this is applied in practice is jointly developed by regulators and a representative sample of developers[[114]](#footnote-114). Similar to the above point about regulation from a service mindset, a co-regulatory approach, when appropriate and with any potential conflicts of interest properly managed, is about generating buy-in from developers through engaging them in the design and implementation of the regulatory process. It is also about designing a regulatory process that reflects and acknowledges the needs of developers as well, not just those of the regulatory body and associated bodies. However, ultimately regulators must remain fully independent of developers to make decisions that put the public's safety first, as well as ensuring that public and private healthcare resources are only used for technologies that meet independently developed standards of quality, safety, and efficacy.

# Recommendations for the way forward

Based on its work, the WG-RC recommends that stakeholders examine the key 18 concepts discussed in Section 5 above and summarized in Table 5 below as they continue to develop frameworks and best practices for the use of AI in healthcare and therapeutic development.

Table 5 – List of key recommendations for regulatory concepts for AI in Health based on each of the six topic areas.

| Topic Area | Recommendation list |
| --- | --- |
| 1. Documentation and Transparency | 1.1 Consider pre-specifying and documenting the intended medical purpose and development process, such as the selection and use of datasets, reference standards, parameters, metrics, deviations from original plans, and updates/ changes, during the phases of development should be considered in a manner that allows for the tracing of the development steps as appropriate. |
| 1.2 Consider a risk-based approach also for the level of documentation and record keeping utilized for the development and validation of AI systems. |
| 2. Risk Management and AI Systems Development Lifecycle Approach | 2.1 Consider a total product lifecycle approach throughout all phases in the life of a medical device: pre-market development management, post-market management/ surveillance, and change management. |
| 2.2 Consider arisk management approach that addresses risks associated with AI systems, such as cybersecurity threats and vulnerabilities, underfitting, and algorithmic bias, etc. |
| 3. Intended Use, and Analytical and Clinical Validation | 3.1 Consider providing transparent documentation of the intended use of the AI system. Details of the training dataset composition underpinning an AI system, including size, setting and population, input and output data and demographic composition should be transparently documented and provided to users. |
| 3.2 Consider demonstrating performance beyond the training dataset through external, analytical validation in an independent test dataset. This test dataset should be representative of the population and setting in which the AI system is intended to be deployed and transparent documentation of the external dataset and performance metrics should be provided. This test dataset should be appropriately independent of the AI training algorithm during training. |
| 3.3 Consider a graded set of requirements for clinical validation based on risk. Randomized clinical trials are the gold standard for evaluation of comparative clinical performance and could be appropriate for the highest risk tools or where the highest standard of evidence is required. In other situations, consider prospective validation in a real-world deployment and implementation trial which includes a relevant comparator using accepted relevant groups. |
| 3.4 Consider a period of more intense post-deployment monitoring through post-market management and market surveillance for high-risk AI systems. |
| 4. Data Quality | 4.1 Consider whether available data is of sufficient quality to support the development of the AI system that can achieve the intended purpose. |
| 4.2 Consider deploying rigorous pre-release evaluations for AI systems to ensure that they will not amplify any of relevant issues, such as biases and errors. |
| 4.3 Consider careful design or prompt troubleshooting to help identify data quality issues early on, which could potentially prevent or ameliorate possible resulting harm. |
| 4.4 Consider mitigating data quality issues that arise in healthcare data and the associated risks. |
| 4.5 Consider working with other stakeholders to create data ecosystems that can facilitate the sharing of good-quality data sources. |
| 5. Privacy and Data Protection | 5.1 Consider privacy and data protection during the design and deployment of AI systems. |
| 5.2 Consider gaining a good understanding of applicable data protection regulations and privacy laws early in the development process and ensure the development process meets or exceeds such legal requirements. |
| 5.3 Consider implementing a compliance program that addresses risks and develop privacy and cybersecurity practices and priorities that take into account potential harm, as well as the enforcement environment. |
| 6. Engagement and Collaboration | 6.1 Consider the development of accessible and informative platforms that facilitate engagement and collaboration, where applicable and appropriate, among key stakeholders of the AI innovation and deployment roadmap. |
| 6.2 Consider streamlining the oversight process for AI regulation through engagement and collaboration to potentially accelerate practice-changing advances in AI. |

# Final considerations

WHO recognizes the potential of Artificial Intelligence (AI) in enhancing health outcomes by improving clinical trials, medical diagnosis, treatment, self-management of care and person-centred care, as well as creating more evidence-based knowledge, skills and competence for professionals to support healthcare. Furthermore, with the increasing availability of healthcare data and the rapid progress of analytics techniques, AI has the potential to transform the health sector to meet a variety of stakeholders' needs in healthcare and therapeutic development. For this reason, the WHO and ITU are collaborating through the Focus Group on AI for Health (FG-AI4H) to facilitate the safe and appropriate development and use of AI systemsin healthcare FG-AI4H's Working Group on Regulatory Considerations (WG-RC) on AI for Health consists of members representing multiple stakeholders including regulatory bodies, policy makers, academia, and industry who explored regulatory and health technology assessment considerations and emerging "good practices" for the development and use of AI in healthcare and therapeutic development. This publication, which is based on the work of the WG-RC, is an overview of regulatory concepts on AI for Health that covers the following six general topic areas: documentation & transparency, risk management and AI systems development lifecycle approach, intended use and analytical & clinical validation, privacy and data protection, and engagement & collaboration. This overview is not intended as a guidance, regulation, or policy. Rather, it is meant as a listing of key regulatory concepts and a resource that can be considered by all relevant stakeholders in medical devices ecosystems, including but not limited to, developers who are exploring and developing AI systems, regulators who might be in the process of identifying approaches to manage and facilitate AI systems, manufacturers who design and develop AI-embedded medical devices, health practitioners who deploy and use such medical devices and AI systems, and those working to remit. The WG-RC recommends that stakeholders examine these key considerations and other potential once as they continue to develop frameworks and best practices for the use of AI in healthcare and therapeutic development in relationship to the 6 topic areas.

The WG-RC recognizes that AI has been instrumental in rapidly advancing research in healthcare and therapeutic development. However, it also recognizes the evolving complexity of the AI landscape and the need for international collaboration to facilitate the safe and appropriate development and use of AI systems. Accordingly, international collaboration on AI regulations and standards is important for three reasons. First, sharing knowledge and best practices of evolving regulatory concepts could increase the speed of developing this regulatory landscape and reduce the gap between advancing technology and regulation. Second, international collaboration improves consistency in regulations, which is important as many tools will likely eventually cross borders. Consistencies around regulatory concepts considered for AI systems and technologies could improve standards and enable more rapid deployment. Finally, international collaboration supports countries with less regulatory capacity by ensuring that these countries can also use tools with high standards, reducing the potential for disparity in the introduction of these tools. Eventually, the WG-RC understands that the AI landscape is rapidly evolving and that the concepts in this deliverable may need to be expanded upon as the technology and its uses develop. It recommends that stakeholders, including regulators and developers and manufacturers, continue to engage and that the community at large works towards shared understanding and mutual learning. In addition, established national and international groups, such as the IMDRF, GHWP, AMDF, and ICMRA, should continue to work on topics of AI for potential regulatory convergence and harmonization.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. Global Strategy on Digital Health 2020– 2025. Geneva: World Health Organization; (2020) (<https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf?sfvrsn=f112ede5_58>) [↑](#footnote-ref-1)
2. World Health Organization. (2020). Global Strategy on Digital Health 2020– 2025. Geneva. [Online] ([https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf?‌sfvrsn=f112ede5\_58](https://www.who.int/docs/default-source/documents/gs4dhdaa2a9f352b0445bafbc79ca799dce4d.pdf?sfvrsn=f112ede5_58)) [↑](#footnote-ref-2)
3. United Nations. (2020). The 17 Goals - Sustainable Development. [Online] (<https://sdgs.un.org/goals>) [↑](#footnote-ref-3)
4. Thirteenth General Programme of Work 2019−2023. Geneva: World Health Organization; (2020) (<https://www.who.int/about/what-we-do/thirteenth-general-programme-of-work-2019---2023>) [↑](#footnote-ref-4)
5. U.S. FDA. (2019). Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback. [Online] (<https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf>) [↑](#footnote-ref-5)
6. International Coalition of Medicines Regulatory Authorities (ICMRA). (2021). Informal Innovation Network Horizon Scanning Assessment Report – Artificial Intelligence. [Online] (<https://www.icmra.info/drupal/sites/default/files/2021-08/horizon_scanning_report_artificial_intelligence.pdf>) [↑](#footnote-ref-6)
7. ISO/IEC TR 24028:2020, Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence. [Online] (<https://www.iso.org/standard/77608.html>) [↑](#footnote-ref-7)
8. The Organisation for Economic Co-operation and Development (OECD) (2019). Principles on AI. OECD. [Online] (<https://www.oecd.org/going-digital/ai/principles/>). [↑](#footnote-ref-8)
9. A non-exclusive list of national efforts: U.S. Food and Drug Administration (US FDA) and Health Canada, The UK Medicines and Healthcare Products Regulatory Agency (MHRA), South African Health Products Regulatory Authority (SAHPRA), European Commission (EC), Singapore Health Sciences Authority (HSA), Japanese Pharmaceuticals and Medical Devices Agency (PMDA), Swissmedic, and Australian Therapeutic Goods Administration (TGA). [↑](#footnote-ref-9)
10. ML-Enabled Medical Devices- A subset of AI-enabled Medical Devices: Key Terms and Definitions, International Medical Device Regulators Forum. Proposed document posted for public consultation, September 16, 2021. [↑](#footnote-ref-10)
11. For the context of this publication, the term AI technology is referred to any AI technology (e.g. machine learning, deep learning, natural language processing, computer vision, etc.) used to develop an AI system. [↑](#footnote-ref-11)
12. AI system: AI-based systems are able to perform tasks such as visual perception, speech recognition, decision-making, and translation between languages by using machine learning (including deep learning) or non-ML expert systems (based on rules like decision trees). For example, a Machine Learning-enabled Medical Device is a medical device that uses machine learning, in part or in whole, to achieve its intended medical purpose. A Machine Learning-enabled Medical Device can be considered an AI-based system. [↑](#footnote-ref-12)
13. ITU/WHO Focus Group on AI for health Deliverable 0.1, "Common unified terms in artificial intelligence for health" (September 2022), <https://itu.int/go/fgai4h>. [↑](#footnote-ref-13)
14. Machine Learning-enabled Medical Devices- A subset of Artificial Intelligence-enabled Medical Devices: Key Terms and Definitions, International Medical Device Regulators Forum. Proposed document posted for public consultation, September 16, 2021. [↑](#footnote-ref-14)
15. According to the United Nations Universal Declaration of Human Rights of 1948, “No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honor and reputation.” [↑](#footnote-ref-15)
16. What is Privacy? IAPP. (2020) (<https://iapp.org/about/what-is-privacy/>) [↑](#footnote-ref-16)
17. Kearns, Michael, and Aaron Roth. *The Ethical Algorithm: The Science of Socially Aware Algorithm Design*. Oxford University Press, 2020, 33. [↑](#footnote-ref-17)
18. # U.S. Food and Drug Authority (2021). Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. Draft Guidance for Industry. <<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory>>

    [↑](#footnote-ref-18)
19. What is Privacy? IAPP. (2020) (<https://iapp.org/about/what-is-privacy/>) [↑](#footnote-ref-19)
20. EU General Data Protection Regulation 2016/679. [↑](#footnote-ref-20)
21. E. Vayena, J. Dzenowagis, M. Langfeld (2016). Evolving Health Data Ecosystem. (<https://www.who.int/ehealth/resources/ecosystem.pdf?ua=1>) [↑](#footnote-ref-21)
22. Software as a Medical Device (SaMD): Key Definitions. IMDRF (2013) (<http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf>). [↑](#footnote-ref-22)
23. IMDRF AIMD WG/N67(PD1):2021. Machine Learning-enabled Medical Devices- A subset of Artificial Intelligence-enabled Medical Devices: Key Terms and Definitions. Proposed document posted for public consultation. [Online] (<https://www.imdrf.org/sites/default/files/2021-10/Machine%20Learning-enabled%20Medical%20Devices%20-%20A%20subset%20of%20Artificial%20Intelligence-enabled%20Medical%20Devices%20-%20Key%20Terms%20and%20Definitions.pdf>) [↑](#footnote-ref-23)
24. Arjun Panesar. (2019). *Machine Learning and AI for* Healthcare. Berkeley, CA: Apress. [Online] (<https://doi.org/https://doi.org/10.1007/978-1-4842-3799-1>). [↑](#footnote-ref-24)
25. [↑](#footnote-ref-25)
26. Wu, E., Wu, K., Daneshjou, R. *et al.* How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nat Med* **27,** 582–584 (2021). [Online] (<https://doi.org/10.1038/s41591-021-01312-x>) [↑](#footnote-ref-26)
27. Liu, X., Cruz Rivera, S., Moher, D. *et al.* Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med* **26,** 1364–1374 (2020). [Online] (<https://doi.org/10.1038/s41591-020-1034-x>) [↑](#footnote-ref-27)
28. Rivera Samantha Cruz, Liu Xiaoxuan, Chan An-Wen, Denniston Alastair K, Calvert Melanie J. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension *BMJ* 2020; 370 :m3210 (<https://www.bmj.com/content/bmj/370/bmj.m3210.full.pdf>)

    [↑](#footnote-ref-28)
29. Software as a Medical Device (SaMD): Key Definitions. IMDRF (2013) (<http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf>). [↑](#footnote-ref-29)
30. Ibid [↑](#footnote-ref-30)
31. Ibid [↑](#footnote-ref-31)
32. Ibid [↑](#footnote-ref-32)
33. Detecting Failure Modes in Image Reconstructions with Interval Neural Network Uncertainty

    Oala, Luis, Heiß, Cosmas, Macdonald, Jan, März, Maximilian, Samek, Wojciech, and Kutyniok, Gitta

    *In ICML 2020 Workshop on Uncertainty & Robustness in Deep Learning* 2020 [↑](#footnote-ref-33)
34. Data and artificial intelligence assessment methods (DAISAM) reference

    Oala, Luis, Balachandran, Pradeep, Cabitza, Federico, Calderon Ramirez, Saul, Chiavegatto Filho, Alexandre, Eitel, Fabian, Extermann, Jérôme, Fehr, Jana, Ghozzi, Stephane, Gilli, Luca, Jaramillo-Gutierrez, Giovanna, Kester, Quist-Aphetsi, Kurapati, Shalini, Konigorski, Stefan, Krois, Joachim, Lippert, Christoph, Martin, Jörg, Merola, Alberto, Murchison, Andrew, Niehaus, Sebastian, Ritter, Kerstin, Samek, Wojciech, Sanguinetti, Bruno, Schwerk, Anne, and Srinivasan, Vignesh. *In Proceedings of the ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H) - Meeting I* 2020 [↑](#footnote-ref-34)
35. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)-Based Software as a Medical Device (SaMD) - Discussion Paper and Request for Feedback. U.S. FDA Artificial Intelligence and Machine Learning. (2019) (<https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf>) [↑](#footnote-ref-35)
36. FDA response to public comments on AI medical Total Product Lifecycle approach on AI workflow of Jan 2021: <https://www.fda.gov/media/145022/download>

    [↑](#footnote-ref-36)
37. International Medical Device Regulators Forum. (2019). Principles and Practices for Medical Device Cybersecurity. [Online] (<http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-ppmdc.pdf>) [↑](#footnote-ref-37)
38. International Medical Device Regulators Forum (2014). Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations. [Online] (<https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf>) [↑](#footnote-ref-38)
39. World Health Organization. (‎2020)‎. Guidance for post-market surveillance and market surveillance of medical devices, including in vitro diagnostics. World Health Organization. <https://apps.who.int/iris/handle/10665/337551> [↑](#footnote-ref-39)
40. Singapore; The Health Sciences Authority (2020). Regulatory Guidelines for Software Medical Devices – A Life Cycle Approach. [Online] (<https://www.hsa.gov.sg/docs/default-source/hprg-mdb/gudiance-documents-for-medical-devices/regulatory-guidelines-for-software-medical-devices---a-life-cycle-approach.pdf>) [↑](#footnote-ref-40)
41. Ibid [↑](#footnote-ref-41)
42. NHSX; UK (2020). A Buyer’s Guide to AI in Health and Care. [Online] (<https://www.nhsx.nhs.uk/ai-lab/explore-all-resources/adopt-ai/a-buyers-guide-to-ai-in-health-and-care/>) [↑](#footnote-ref-42)
43. “Handling with applications for confirmation of PACMP for medical devices” MHLW PSEHBSD Notification No.0831-14, August 31, 2020. [Online] (<https://www.mhlw.go.jp/content/11120000/000665757.pdf>) [↑](#footnote-ref-43)
44. International Conference of Harmonization harmonised guideline: Technical and regulatory considerations for pharmaceutical product lifecycle management Q12. [Online] (<https://database.ich.org/sites/default/files/Q12_Guideline_Step4_2019_1119.pdf> ) [↑](#footnote-ref-44)
45. Schörverth, Elora, et.al. FG-AI4H Open Code Initiative - Evaluation and Reporting Package. *In Proceedings of the ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H) - Meeting K* 2021 [↑](#footnote-ref-45)
46. Sendak, M. P., Gao, M., Brajer, N., & Balu, S. (2020). Presenting machine learning model information to clinical end users with model facts labels. *NPJ digital medicine*, *3*(1), 1-4. [↑](#footnote-ref-46)
47. Data and artificial intelligence assessment methods (DAISAM) Audit Reporting Template. Verks, Boris, and Oala, Luis. *In Proceedings of the ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H) - Meeting J* 2020 [↑](#footnote-ref-47)
48. *Luis Oala, et. al.* ML4H Auditing: From Paper to Practice.*Proceedings of the Machine Learning for Health NeurIPS Workshop*, PMLR 136:280-317, 2020. [↑](#footnote-ref-48)
49. Post-Hoc Domain Adaptation via Guided Data Homogenization. Willis, Kurt, and Oala, Luis. *In ICLR 2021 Workshop on Robust and Reliable Machine Learning in the Real World Workshop (RobustML)* 2021 [↑](#footnote-ref-49)
50. More Than Meets The Eye: Semi-supervised Learning Under Non-IID Data. Calderon-Ramirez, Saul, and Oala, Luis. *In ICLR 2021 Workshop on Robust and Reliable Machine Learning in the Real World Workshop (RobustML)* 2021 [↑](#footnote-ref-50)
51. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. Bellemo et al. The Lancet Digital Health. Vol 1, Issue 1, 35-44. 2019. [↑](#footnote-ref-51)
52. Macdonald J., März M., Oala L., Samek W. (2021) Interval Neural Networks as Instability Detectors for Image Reconstructions. In: Palm C., Deserno T.M., Handels H., Maier A., Maier-Hein K., Tolxdorff T. (eds) Bildverarbeitung für die Medizin 2021. Informatik aktuell. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-33198-6\_79 [↑](#footnote-ref-52)
53. Oala, L. et. al. (2020, November). ML4H Auditing: From Paper to Practice. In *Machine Learning for Health* (pp. 280-317). PMLR. [↑](#footnote-ref-53)
54. I-DAIR. (2020). Advancing Digital Health and Artificial Intelligence Research through Collaboration. [Online] (<http://i-dair.org/>) [↑](#footnote-ref-54)
55. Salim, M., Wåhlin, E., Dembrower, K., Azavedo, E., Foukakis, T., Liu, Y., Smith, K., Eklund, M., and Strand, F. (2020). External Evaluation of 3 Commercial Artificial Intelligence Algorithms for Independent Assessment of Screening Mammograms. JAMA Oncology [↑](#footnote-ref-55)
56. ITU. (2022). FG-AI4H Open Code Initiative (OCI). [Online] (<https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/opencode.aspx>) [↑](#footnote-ref-56)
57. aiaudit.org. (2022) AI Audit.org. [Online] (<https://aiaudit.org/>) [↑](#footnote-ref-57)
58. Oala, Luis et.al. ML4H Auditing: From Paper to Practice *In Proceedings of the Machine Learning for Health NeurIPS Workshop* 2020 [↑](#footnote-ref-58)
59. International Medical Device Regulators Forum. (2016). Software as a Medical Device (SaMD): Clinical Evaluation. [Online] (<http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-samd-ce.pdf>) [↑](#footnote-ref-59)
60. National Institute For Health And Care Excellence (NICE). [Online] (<https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf>) [↑](#footnote-ref-60)
61. International Medical Device Regulators Forum (2014). Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations. [Online] (<https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf>) [↑](#footnote-ref-61)
62. Liu, X., Cruz Rivera, S., Moher, D. *et al.* Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med* **26,** 1364–1374 (2020). [Online] (<https://doi.org/10.1038/s41591-020-1034-x>) [↑](#footnote-ref-62)
63. Topol, E. (2020). Welcoming new guidelines for AI clinical research. Nat. Med. [↑](#footnote-ref-63)
64. Panesar A. (2019) Data. In: Machine Learning and AI for Healthcare. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-3799-1\_2 [↑](#footnote-ref-64)
65. Ibid [↑](#footnote-ref-65)
66. Ibid [↑](#footnote-ref-66)
67. # U.S. Food and Drug Administration (2021). Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products. Draft Guidance for Industry. [Online] (<https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory>)

    [↑](#footnote-ref-67)
68. Duke-Margolis Center for Health Policy. (2019). Determining Real-World Data’s Fitness for Use and the Role of Reliability. [Online] (<https://healthpolicy.duke.edu/sites/default/files/2019-11/rwd_reliability.pdf>) [↑](#footnote-ref-68)
69. Obermeyer, Powers, Vogeli, & Mullainathan. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447-453. 10.1126/science.aax2342 [↑](#footnote-ref-69)
70. Shana Lynch (2020). The Geographic Bias in Medical AI Tools. [Online] (<https://hai.stanford.edu/news/geographic-bias-medical-ai-tools>) [↑](#footnote-ref-70)
71. <https://datanutrition.org/> [↑](#footnote-ref-71)
72. Gebru, T. *et al.* (2021). Datasheets for Datasets. [Online] (<https://arxiv.org/abs/1803.09010>) [↑](#footnote-ref-72)
73. This list will be updated and harmonized with IMDRF work. [↑](#footnote-ref-73)
74. For a broader discussion of privacy and other ethical considerations for the use of AI, refer to the deliverable of the FG-AI4H’s Working Group on Ethical Considerations on AI for Health and international, regional, and national recommendations, such as the EU’s Ethics Guidelines for Trustworthy AI (https://www.who.int/publications/i/item/9789240029200). [↑](#footnote-ref-74)
75. Greenleaf, Graham, Global Tables of Data Privacy Laws and Bills (7th Ed, January 2021) (February 11, 2021). (2021) 169 Privacy Laws & Business International Report. 6-19, Available at SSRN: <https://ssrn.com/abstract=3836261> or [http://dx.doi.org/10.2139/ssrn.3836261](https://dx.doi.org/10.2139/ssrn.3836261); Global data privacy laws 2021: Despite COVID delays, 145 laws show GDPR dominance’ (2021) 169 PLBIR 1, 3-5; Global data privacy laws 2021: Uncertain paths for international standards’ (2021) 169 PLBIR 23-27. [↑](#footnote-ref-75)
76. See also India’s proposed Personal Data Protection Act [↑](#footnote-ref-76)
77. Like the GDPR, the CCPA applies to natural persons who are California residents who are “domiciled in the state or who is outside the state for a temporary or transitory purpose.” Cal. Code Regs. tit. 18, §17014. [↑](#footnote-ref-77)
78. UK Information Commissioner’s Office, Introduction to Anonymisation: Draft anonymisation, pseudonymisation, and privacy enhancing technologies guidance. https://ico.org.uk/media/about-the-ico/consultations/2619862/anonymisation-intro-and-first-chapter.pdf [↑](#footnote-ref-78)
79. GDPR, Article 9. [↑](#footnote-ref-79)
80. Data flows have increasingly become an important part of global interconnection and AI development. Although the Schrems II case pertains to the EU-US position on data transfers, the wider implications inform global data transfers and the way it which they are to be compatible with GDPR requirements, including the validity of standard contractual clauses which depend on whether effective mechanisms are in place to ensure compliance with the level of protection required under the GDPR. *Data Protection Commissioner v. Facebook Ireland Limited, Maximillian Schrems* (Case C-311/18, “*Schrems II*”). [↑](#footnote-ref-80)
81. Adequacy decisions : How the EU determines if a non-EU country has an adequate level of data protection. European Commission (<https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en>). [↑](#footnote-ref-81)
82. For example, a pillar of the data quality continuum in some jurisdictions, e.g. EU law, is the accountability principle. According to Art. 5 of the GDPR, data controllers shall abide by the five sets of principles enshrined in Art. 5(1), e.g., data minimization. Data controllers shall determine both technical and organizational measures to attain such ends (Art. 5(2)), throughout the entire cycle of data processing. Although not mentioned, the accountability principle is also at work in Art. 24(1), 25(1), and 32 of the regulation in regard to the responsibility of the controller, principle of data protection by design (and by default), and security measures. [↑](#footnote-ref-82)
83. The National Institute of Standards and Technology (NIST) Privacy Framework: A Tool for Improving Privacy Through Enterprise Risk Management. National Institute of Standards and Technology (NIST), U.S. Department of Commerce, Washington DC (2020). (https://www.nist.gov/system/files/documents/2020/01/16/NIST%20Privacy%20Framework\_V1.0.pdf) [↑](#footnote-ref-83)
84. Ibid. [↑](#footnote-ref-84)
85. Ibid. [↑](#footnote-ref-85)
86. GDPR Articles 25 and 32. [↑](#footnote-ref-86)
87. India’s Personal Data Protection Act Ch. VI, 22(1)(e), 24(1). [↑](#footnote-ref-87)
88. CCPA § 1798.150(a)(1), South Africa’s Protection of Personal Information Act of 2013, Israeli Privacy Protection Regulations (Data Security), 5777-2017 (implementing the Protection of Privacy Law, 5741-1981 of 1981), United Arab Emirates’ Federal Law No. 2 of 2019, Kingdom of Saudi Arabia’s E-Commerce Law of 2019 and its Implementing Rules. [↑](#footnote-ref-88)
89. See Article 35, “A data protection impact assessment shall be conducted if processing is likely to result in high risk to the rights and freedoms of the natural persons.” [↑](#footnote-ref-89)
90. While risk assessments are quite common in information security standards and requirements, they are rarely seen in privacy rules in the U.S. The GDPR, however, requires that an organization processing personal data has to conduct a specific Data Privacy Impact Assessment or DPIA before beginning the processing. [↑](#footnote-ref-90)
91. The NIST framework is a general guidance for data security and is not specific to health care data. NIST Privacy Framework: A Tool for Improving Privacy Through Enterprise Risk Management. NIST (2020). (<https://www.nist.gov/system/files/documents/2020/01/16/NIST%20Privacy%20Framework_V1.0.pdf>) [↑](#footnote-ref-91)
92. West, Darrell M., and John R. Allen. *Turning Point: Policymaking in the Era of Artificial Intelligence*. Brookings Institution Press, 2020, 183. [↑](#footnote-ref-92)
93. The NIST Cybersecurity Framework is an internationally recognized document that explores these concepts in more detail. Framework for Improving Critical Infrastructure Cybersecurity. NIST (2018) (<https://www.nist.gov/cyberframework>). [↑](#footnote-ref-93)
94. # How should we engage and involve patients and the public in our work, MHRA (2020) (<https://www.gov.uk/government/consultations/how-should-we-engage-and-involve-patients-and-the-public-in-our-work>)

    [↑](#footnote-ref-94)
95. Australia Government Department of Health, Stakeholder Engagement Framework, AccountAbility. AA1000SE Stakeholder Engagement Practitioner’s Perspectives (2005) [↑](#footnote-ref-95)
96. Communication on enabling the digital transformation of health and care in the Digital Single Market; empowering citizens and building a healthier society, European Commission (2018) (<https://ec.europa.eu/digital-single-market/en/news/communication-enabling-digital-transformation-health-and-care-digital-single-market-empowering>) [↑](#footnote-ref-96)
97. Health Sciences Authority (HSA). (2020). Regulatory Guidelines for Software Medical Devices – A Life Cycle Approach. [Online] (https://www.hsa.gov.sg/docs/default-source/hprg-mdb/gudiance-documents-for-medical-devices/regulatory-guidelines-for-software-medical-devices---a-life-cycle-approach.pdf"https://www.hsa.gov.sg/docs/default-source/hprg-mdb/gudiance-documents-for-medical-de) [↑](#footnote-ref-97)
98. E-services, HSA (2021) (<https://www.hsa.gov.sg/e-services>) [↑](#footnote-ref-98)
99. Australia Government Department of Health, Stakeholder Engagement Framework, AccountAbility. AA1000SE Stakeholder Engagement Practitioner’s Perspectives. 2005 [↑](#footnote-ref-99)
100. International Association for Public Participation (IAP2) Spectrum (2007) [↑](#footnote-ref-100)
101. D. Ho, Artificial Intelligence in Cancer Therapy (2020) (<https://science.sciencemag.org/content/367/6481/982>) [↑](#footnote-ref-101)
102. D. Ho, Addressing COVID-19 Drug Development with Artificial Intelligence (2020) (<https://onlinelibrary.wiley.com/doi/full/10.1002/aisy.202000070>). [↑](#footnote-ref-102)
103. A. Blasiak, J.J. Lim, S.G.K. Seah, T. Kee, A. Remus, D.H. Chye, P.S. Wong, L. Hooi, A.T.L. Truong, N. Le, C.E.Z. Chan, R. Desai, X. Ding, B.J. Hanson, E.K.-H. Chow, D. Ho, IDentif.AI: Rapidly optimizing combination therapy design against severe Acute Respiratory Syndrome Coronavirus 2 (SARS‐Cov‐2) with digital drug development, (2020) (<https://aiche.onlinelibrary.wiley.com/doi/10.1002/btm2.10196>) [↑](#footnote-ref-103)
104. Health Sciences Authority: Regulatory Guidelines for Software Medical Devices – A Lifecycle Approach; <https://www.hsa.gov.sg/docs/default-source/announcements/regulatory-updates/regulatory-guidelines-for-software-medical-devices--a-lifecycle-approach.pdf>. [↑](#footnote-ref-104)
105. D. Ho, S.R. Quake, E.R.B. McCabe, W.J. Chng, E.K. Chow, X. Ding, B.D. Gelb, G.S. Ginsburg, J. Hassenstab, C.M. Ho, W.C. Mobley, G.P. Nolan, S.T. Rosen, P. Tan, Y. Yen, and A. Zarrinpar, Enabling Technologies for Personalized and Precision Medicine (2020) (<https://www.cell.com/trends/biotechnology/fulltext/S0167-7799(19)30316-6>) [↑](#footnote-ref-105)
106. P. Shah, F. Kendall, S. Khozin, R. Goosen, J. Hu, J. Laramie, M. Ringel, N. Schork, Artificial intelligence and machine learning in clinical development: a translational perspective Artificial intelligence and machine learning in clinical development: a translational perspective, (2019) ([www.nature.com/articles/s41746-019-0148-3](file:///C:\Users\campos\AppData\Roaming\Microsoft\Word\www.nature.com\articles\s41746-019-0148-3)) [↑](#footnote-ref-106)
107. S. Harrer, P. Shah, B. Antony, J. Hu (2019) Artificial Intelligence for Clinical Trial Design; ([https://www.sciencedirect.com/science/article/pii/S0165614719301300#:~:text=AI%20techniques%20have%20advanced%20to,to%20assist%20human%20decision%2Dmakers.&text=We%20explain%20how%20recent%20advances,towards%20increasing%20trial%20success%20rates](https://www.sciencedirect.com/science/article/pii/S0165614719301300)) [↑](#footnote-ref-107)
108. Special Access Route (Medical Devices), HSA (2019) (<https://www.hsa.gov.sg/medical-devices/registration/special-access-routes/qualified-practitioner-request>) [↑](#footnote-ref-108)
109. Health Product Classification Form Health Product Classification Form, (HSA) (<https://crm.hsa.gov.sg/Webform/HPRG>) [↑](#footnote-ref-109)
110. Cf. Pagallo, U., Casanovas, P. and R. Madelin (2019) The Middle-out Approach: Assessing Models of Legal Governance in Data Protection, Artificial Intelligence, and the Web of Data, *The Theory and Practice of Legislation*, 7(1): 1-25. [↑](#footnote-ref-110)
111. Australia Government Department of Health, Stakeholder Engagement Framework, AccountAbility. AA1000SE Stakeholder Engagement Practitioner’s Perspectives,(2005) <https://www.health.gov.au/sites/default/files/stakeholder-engagement-framework_0.pdf> [↑](#footnote-ref-111)
112. Stakeholder Communications, Atlassian (<https://www.atlassian.com/team-playbook/plays/stakeholder-communications-plan>) [↑](#footnote-ref-112)
113. R. Clarke (2019). Regulatory alternatives for AI. Computer Law & Security Review. (<https://doi.org/10.1016/j.clsr.2019.04.008>) [↑](#footnote-ref-113)
114. Ibid. [↑](#footnote-ref-114)