ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Focus Group Deliverable

(09/2022)

Focus Group on Artificial Intelligence for Health
(FG-AI4H)

FG-AI4H DEL02

Regulatory considerations on artificial intelligence for health

ITU-T FG-AI4H Deliverable

DEL02 – Regulatory considerations 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).

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Change Log

This document contains Version 1 of the Deliverable DEL02 on overview of "Regulatory considerations on artificial intelligence for health" approved at the ITU-T Focus Group on AI for Health (FG-AI4H) meeting held on 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 document was led by Sameer Pujari (Department of Digital Health and Innovation) and Shada AlSalamah (Department of Digital Health and Innovation) under the overall guidance of Alain Labrique (Director, Department of Digital Health and Innovation), Jeremy Farrar (Chief Scientist, Science Division), Soumya Swaminathan (Former Chief Scientist, Science Division), Bernardo Mariano (Former Director, Department of Digital Health and Innovation), Adriana Velazquez Berumen (Access to Medicines and Health Products), and Anita Sands (Regulation and Prequalification Department).

Technical coordination of the topic areas was provided by the following subgroup leads (in alphabetical order): Shada Alsalamah (Department of Digital Health and Innovation) led the subgroup on the Data Quality and Risk Management and AI Systems Development Lifecycle Approaches; M Khair ElZarrad, (U.S. Food and Drug Administration, United States of America) led the Documentation and Transparency topic area; Monique Kuglitsch (Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Germany) and Dean Ho (National University of Singapore, Singapore) jointly led the Engagement and Collaboration topic area; Naomi Lee (The Lancet, United Kingdom) led the Intended Use and Analytical and Clinical Validation topic area; and Rose Purcell (Former U.S. Food and Drug Administration, United States of America) led the Privacy and Data Protection topic area.

FG-AI4H are grateful to the following persons who contributed to development and review of this publication (in alphabetical order): Najeeb Al-Shorbaji (eHealth Development Association, Jordan), Batoul Albaz (Health Sector, King Abdulaziz City for Science and Technology, Saudi Arabia), Paolo Alcini (European Medicines Agency, Netherlands), Ali AlDalaan (Saudi Food and Drug Administration, Saudi Arabia), Samvel Azatyan (WHO), Judith Van Andel (WHO), Fazilah Shaik Allaudin (Ministry of Health, Malaysia), 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, United Arab Emirates), Abdulgader Almoeen (National Center for AI, Saudi Data & AI Authority, Saudi Arabia), Sultan Alzahrani (Digital Health Institute, Health Sector, King Abdulaziz City for Science and Technology, Saudi Arabia), Lin Anle (Health Sciences Authority, Singapore), Housseynou Ba (WHO), Pat Baird (FG-AI4H, United States of America), Michael Berensmann (Federal Institute for Drugs and Medical Devices, Germany), Simão de Campos Neto (ITU), Marcelo D'Agostino (WHO), Jose Eduardo Díaz Mendoza (WHO), Mengjuan Duan (WHO), Clayton Hamilton (WHO), Josee Hansen (WHO), Wouter 'T Hoen (WHO), Mattias Karlsson Dinnetz (IP for Innovators Department, Technology Transfer Section, World Intellectual Property Organization), Tala H. Fakhouri (U.S. Food and Drug Administration, United States of America), Hélio Bomfim de Macêdo Filho (Brazilian Health Regulatory Agency- Anvisa, Brazil), Luca Foschini (FG-AI4H, United States of America), Mohammed VI Hassan Ghazal (University of Health Sciences, Morocco), Liang Hong (China's Center for Medical Device Evaluation, China), Indra Joshi (NHSX, National Health Service, United Kingdom), Kassandra Karpathakis (Harvard TH Chan School of Public Health, United States of America; NHSX, National Health Service, United Kingdom), Tim Kelsey (Healthcare Information and Management Systems Society, United Kingdom), Andrea Keyter (FG-AI4H, South Africa), 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), Xiaoxuan Liu (University of Birmingham, United Kingdom), Rohit Malpan (WHO), Ahmed Mandil (WHO), Junaid Nabi (Harvard Business School, United States of America), Mariam Nouh (Future Economies Sector, King Abdulaziz City for Science and Technology, Saudi Arabia), Mohamed Nour (WHO), David Novillo Ortiz (WHO), Luis Oala (Fraunhofer Heinrich Hertz Institute, Germany), Mats Ohlson (FG-AI4H, Sweden), Adrian Pacheco-Lopez (National Center for Health Technology Excellence, Ministry of Health, Mexico), Ugo Pagallo (University of Turin, Italy), Maria Beatrice Panico (Medicines and Healthcare products Regulatory Agency, United Kingdom), Andres Pichon-Riviere (Institute for Clinical Effectiveness and Health Policy, Argentina), Julie Polisena (Health Canada,

Canada), Pierre Quartarolo (Danish Medicines Agency, Denmark), Chandrashekar Ranga (The Central Drugs Standard Control Organisation, India), Andreas Reis (WHO), Mansooreh Saniei (King's College London, United Kingdom), Raymond Francis R. Sarmiento (University of the Philippines Manila, Philippines), Kanako Sasaki (Ministry of Health, Labour and Welfare, Japan), Brian Scarpelli (FG-AI4H, United States of America), Denise Schalet (WHO), Rama Sethuraman (Health Sciences Authority, Singapore), Agnes Sitta Kijo (WHO), Robert Ssekitoleko (Makerere University, Uganda), Mariam Shokralla (WHO), Bev Townsend (University of York, United Kingdom), Tayab Waseem (FG-AI4H, United States of America), Thomas Wiegand (Fraunhofer Heinrich Hertz Institute, Germany), Yu Zhao (WHO) and Georg Zimmermann (Paracelsus Medical University, Austria).

Abbreviations and acronyms

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
DHSC	Department of Health and Social Care
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 Harmonisation of Technical Requirements for Pharmaceuticals 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
IP	Intellectual Property
ISO	International Organization for Standardization
ITU	International Telecommunication Union
MAS	Multi-agent Systems
MHRA	Medicines and Healthcare Products Regulatory Agency
ML	Machine Learning
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NIST	National Institute of Standards and Technology
OECD	Organisation for Economic Co-operation and Development
PACMP	Post-Approval Change Management Protocol
PMDA	Japanese Pharmaceuticals and Medical Devices Agency
QMS	Quality Management System
SAHPRA	South African Health Products Regulatory Authority
SaMD	Software as a Medical Device
SANAS	South African National Accreditation System

SAR	Special Access Route
SPIRIT-AI	Standard Protocol Items: Recommendations for Interventional Trials for AI
TGA	Therapeutic Goods Administration
TPLC	Total Product Lifecycle
US FDA	U.S. Food and Drug Administration
WG-RC	Working Group on Regulatory Considerations on Artificial Intelligence for Health
WHO	World Health Organization
WIPO	World Intellectual Property Organization

Executive Summary

The mission of the World Health Organization (WHO) is to promote health, keep the world safe and serve the vulnerable is articulated in its global strategy on digital health 2020–2025 *(1)*. At the heart of this strategy, 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 in order to prevent, detect and respond to epidemics and pandemics, developing infrastructure and applications. Many international organizations and global players are contributing to this area along with WHO, including The International Medical Device Regulators Forum (IMDRF), Global Harmonization Working Party (GHWP), the US Food and Drug Administration (U.S. FDA), Health Canada, the International Coalition of Medicines Regulatory Authorities (ICMRA), the International Organization for Standardization (ISO), the Organisation for Economic Co-operation and Development (OECD), the United Kingdom's Medicines and Healthcare Products Regulatory Agency (MHRA), the South African Health Products Regulatory Authority (SAHPRA), the European Commission (EC), Singapore's Health Sciences Authority (HSA), the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH), Japan's Pharmaceuticals and Medical Devices Agency (PMDA), Swissmedic and Australia's 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 and personalized care, as well as by creating more evidence-based knowledge, skills and competencies for professionals to support health care. Furthermore, with the increasing availability of health-care 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 health care and therapeutic development.

In order to facilitate the safe and appropriate use of AI technologies for the development of AI systems in health care, the International Telecommunication Union (ITU) and 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 health care and therapeutic development. The work of the WG-RC represents a multidisciplinary, international effort to increase dialogue and examine key concepts for the use of AI in health care.

This publication, which is based on the work of the WG-RC, aims to deliver an overview of *regulatory considerations on artificial intelligence for health* that covers the following six general topic areas: documentation and transparency, the 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 guidance or as a regulatory framework or policy. Rather, it is a discussion of key regulatory considerations and a resource that can be considered by all relevant stakeholders – including developers who are exploring and developing AI systems, regulators and policy-makers who 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 health care 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, namely: pre-market development management, post-market surveillance and change management. In addition, it is essential to consider a risk 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 and testing data through external analytical validation in an independent dataset. This external validation dataset should be representative of the population and setting in which it is intended to deploy the AI system and should be independent of the dataset used for developing the AI model during training and testing. Transparent documentation of the external dataset and performance metrics should be provided. Furthermore, it is important to 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, prospective validation can be considered in a real-world deployment and implementation trial which includes a relevant comparator that uses accepted groups. Finally, a period of more intense post-deployment monitoring should be considered through post-market surveillance and market surveillance for AI systems.

4) *Data quality:* Developers should consider whether available data are of sufficient quality to support the development of the AI system to 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 clause 4, such as biases and errors. Careful design or prompt troubleshooting can help identify data quality issues early and can prevent or mitigate possible resulting harm. Stakeholders should also consider mitigating data quality issues and the associated risks that arise in health-care data, 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 should ensure that the development process meets or exceeds such legal requirements. It is also important to consider implementing a compliance programme 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:* During development of the AI innovation and deployment roadmap it is important to consider the development of accessible and informative platforms that facilitate engagement and collaboration among key stakeholders, where applicable and appropriate. It is fundamental to consider streamlining the oversight process for AI regulation through such engagement and collaboration in order to accelerate practice-changing advances in AI.

Finally, the WG-RC has provided a forum for regulators and subject matter experts to discuss regulatory considerations 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 evolving rapidly and that the considerations in this deliverable may require expansion as technology and its uses develop. The working group 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 International Medical Device Regulators Forum (IMDRF) and the International Coalition of Medicines Regulatory Authorities (ICMRA) should continue to work on topics of AI for potential regulatory convergence and harmonization.

Table of Contents

ITU-T FG-AI4H Deliverable

DEL02 – Regulatory considerations on artificial intelligence for health

1 Introduction

The mission of the World Health Organization (WHO) to promote health, keep the world safe and serve the vulnerable is articulated in its Global strategy on digital health 2020–2025 *(1)*. At the heart of this strategy, 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 well-being in order to achieve the United Nation's health-related Sustainable Development Goals (SDGs) *(2)* and the triple billion targets of WHO's Thirteenth General Programme of Work, 2019–2023 *(3)*.

In addition to WHO's efforts, there is a wave of interest by many other international and regional organizations. Key players include the International Medical Device Regulators Forum (IMDRF) *(4)*, the Global Harmonization Working Party (GHWP), the International Coalition of Medicines Regulatory Authorities (ICMRA) *(5)*, the International Organization for Standardization (ISO) *(6)*, the Organisation for Economic Co-operation and Development (OECD) *(7)* and the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH). Moreover, there are national efforts sharing the same goal.[1]

The digital transformation of health care and therapeutic development, which includes exploring the uses of Artificial Intelligence (AI), 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" *(8)*. With the increasing availability of health-care data and the rapid progress in analytics techniques, AI has the potential to transform the health sector, which is one of the most important sectors for societies and economies worldwide.

2 Purpose

The International Telecommunication Union (ITU) is the United Nation's specialized agency for information and communications technology while WHO is the United Nation's specialized agency for health. These organizations partnered to establish an open group of experts to develop a generalizable benchmarking[2] 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

---

1 A non-exclusive list of national efforts: US Food and Drug Administration (US FDA), Health Canada, the Medicines and Healthcare Products Regulatory Agency (MHRA) of the United Kingdom, the South African Health Products Regulatory Authority (SAHPRA), the European Commission (EC), the Singapore Health Sciences Authority (HSA), Japan's Pharmaceuticals and Medical Devices Agency (PMDA), Swissmedic and Australia's Therapeutic Goods Administration (TGA).

2 This framework should not be confused with WHO's global benchmarking tool for the evaluation of national regulatory systems (https://www.who.int/tools/global-benchmarking-tools, accessed 25 July 2023).

technologies[3] for the development of AI systems[4] in health care 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 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 health care and therapeutic development.

This publication is a general, high-level and nonexclusive overview of key regulatory considerations in 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 health-care applications in general, 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 considerations and a resource for all relevant stakeholders – including 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 in medical devices, and health practitioners who deploy and use such medical devices and AI systems.

3 Definitions and fundamental concepts

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

4 Key artificial intelligence applications in health care and therapeutic development

AI is increasingly being explored to advance health care on multiple fronts. The blending of technology and medicine in research and development is facilitating a wealth of innovation that continues to improve *(9)*. Many health-related AI systems already exist or are being developed to meet a variety of stakeholders' needs in health care and therapeutic development. These solutions have wide-ranging uses across the spectrum of health-care delivery and therapeutic development. For instance, AI systems are being used in health care to support patients throughout the diagnosis and treatment of a disease, using solutions that support adherence to therapeutics and enhance communication capabilities with care providers.

Health care is becoming more patient-centric with personalized approaches to decision-making. This allows 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 increasingly being incorporated and utilized in the clinical setting. For instance, AI-enabled medical devices are being utilized to support clinical decision-making, and AI systems can facilitate clinical assessment of patients and care triaging. AI systems are also being used in the

[3] In the context of this publication, the term "AI technology" is used to refer to any AI technology (e.g., machine learning, deep learning, natural language processing, computer vision etc.) that is used to develop an AI system.

[4] An AI system is an AI-based system that is able to perform tasks such as visual perception, speech recognition, decision-making and translation between languages by using machine learning (ML) (including deep learning) or non-ML expert systems (based on rules such as decision trees). For example, an ML-enabled medical device uses ML, in part or in whole, to achieve its intended medical purpose and can therefore be considered an AI-based system.

development and evaluation of medical products, including during drug discovery to identify potential therapeutic candidates and in clinical research for patient enrichment. Figure 1 illustrates areas of AI research and development across the spectrum of health-care delivery and therapeutic development. The figure does not show 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.

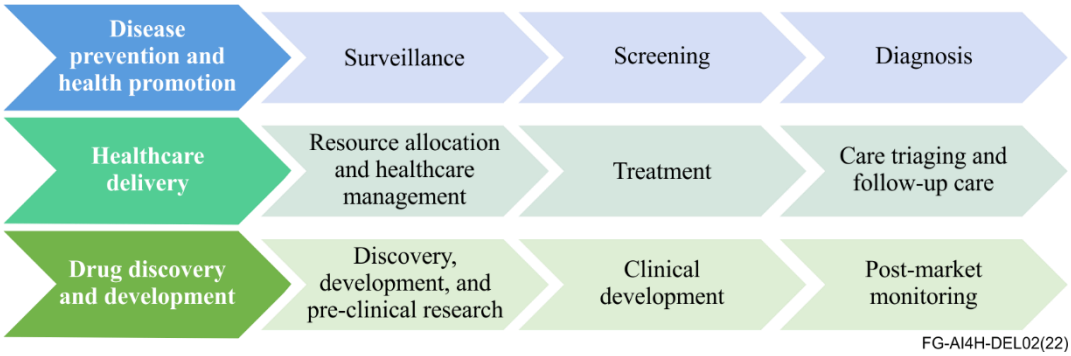

Figure 1 – A general spectrum of AI research and development in health-care delivery and therapeutic development

The spectrum in Figure 1 assists in determining what regulatory considerations may be applicable and how they can be implemented. This document describes a selection of key regulatory considerations and discusses topic areas that are relevant to many stakeholders in the current AI for health ecosystem.

5 Topic areas of regulatory considerations

AI systems may be utilized across all aspects of health care and therapeutic development. Regardless of the category of the AI system application, regulators are keen to ensure not only 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, manufacturers, regulators, users and patients – is highly advised as the AI community matures. Consequently, this publication aims to establish a common understanding of the use of AI systems in health that can be relevant to stakeholders.

An extensive literature review, which included current guidelines, allowed for the identification of a list of topic areas of regulatory considerations for the use of AI in health care 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 while also discussing the remaining sections of this publication. The working group was divided into six subgroups composed of subject matter experts who drafted a section on each topic area.

Table 1 – Six key topic areas of regulatory considerations

Topic Area No.	Topic Area Name
Topic Area 1	Documentation and transparency
Topic Area 2	Risk management and AI systems development lifecycle approaches
Topic Area 3	Intended use and analytical and clinical validation
Topic Area 4	Data quality
Topic Area 5	Privacy and data protection
Topic Area 6	Engagement and collaboration

The WG-RC stressed that this list is not a fully inclusive list of key considerations. The working group expects that the list will serve as a starting point for future deliberations and subsequent updates. For example, global systems for protecting intellectual property (IP) may be an important area to discuss as part of cross-jurisdiction regulations for some stakeholders (mainly AI system developers and manufacturers), and also in relation to, for instance, the protection of AI-related inventions by way of laws on patents and trade secrets. Although not addressed in this report, the World Intellectual Property Organization (WIPO) has already begun a dialogue on AI and IP *(10)*. Thus, WHO will engage in this work together with WIPO and other relevant stakeholders.

5.1 Documentation and transparency

Documentation and transparency are critical concepts that are essential for facilitating scientific and regulatory assessments of AI systems. They also help ensure trust not only in the AI system itself, but also between developers, manufacturers and end-users. Accurate and comprehensive documentation is essential to allowing a transparent evaluation of AI systems for health. This includes undertaking a total product lifecycle approach to pre-specifying and documenting processes, methods, resources and decisions made in the initial conception, development, 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 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 also apply to 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 instance, aspects requiring careful documentation include specifying the problem that developers are attempting to address, the context in which the AI system is proposed to function, and the selection, curation and processing of training datasets used in the development process.

Documentation should allow for the tracking, recording and retention of records of essential steps and decisions, including justifications and reasoning for deviating from pre-specified plans. Effective documentation may also help to show that developers and manufacturers are taking into consideration the full complexity of the context within which the AI system is expected to operate. Moreover, developers and manufacturers should describe how the AI system is addressing the needs of users and why widening the user base would be appropriate. Without transparent documentation, it becomes hard to understand whether the proposed approaches will generalize from the retrospective clinical evidence presented in the regulatory submission to real-world deployments in new settings, which may markedly reduce performance *(11)*. Figure 2 shows examples of essential steps and decision points that developers and manufacturers are encouraged to consider for documentation purposes.

Different entities with 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 are scientifically sound. Systems used to track and document the development processes and key decision points should record access and should be protected against data manipulation and adversarial attacks.

Documentation and transparency should not be seen 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 authorities prior to or in the 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. Steps to facilitate transparency include: publishing in peer-reviewed journals; sharing data and datasets; and making code available to foster mutual learning and facilitate additional studies. Academic institutions, medical journals, regulatory organizations and 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) *(12)* and Standard Protocol Items: Recommendations for Interventional Trials for AI (SPIRIT-AI) *(13)* – have given useful guidance about how to design studies to collect clinical evidence where AI systems are used, as well as how to publish the results. Transparency is not only an important consideration for building trust but can also be a useful tool for educating end-users. This can be achieved, if appropriate, by adapting communication strategies to the needs of end-users and other stakeholders such as patients and communities. 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.

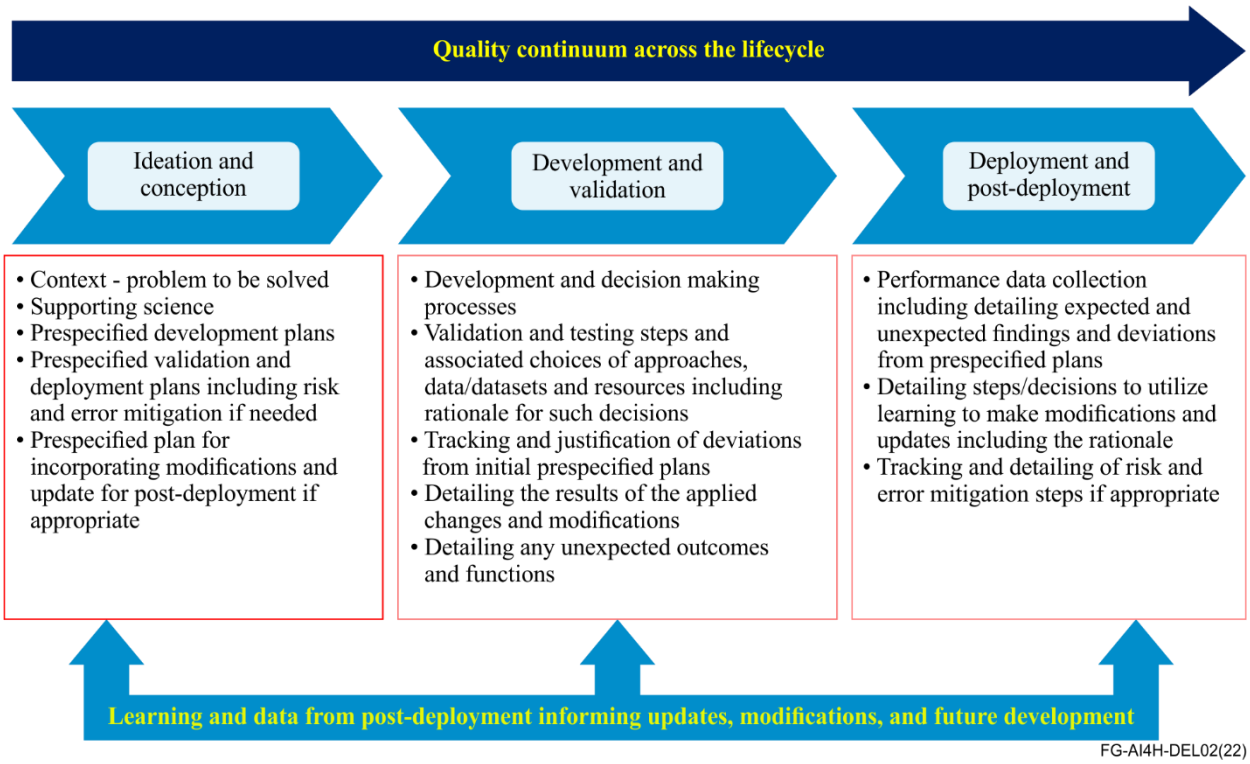

Figure 2 – Examples points of key development decision in the development of AI systems

The following are some elements that might be useful to consider in terms of documentation and record retention.

5.1.1 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 are identified and managed. Validation processes and benchmarking should be carefully documented – including the decisions for selecting specific datasets, reference standards, parameters and metrics to justify such processes. For example, careful consideration should be given to documenting how and why specific data or datasets are selected to train, externally validate and retrain the model (e.g., post-deployment retraining).

5.1.2 Pre-specification and documenting the medical purpose, clinical context and development

The intended medical purpose/function of the AI systems should be clearly documented. For instance, what is the problem that the AI system aims to resolve? This should take into consideration the full clinical and health contexts in which a tool is expected to function. For example, clinical care environments can be vastly complex and may involve several individuals with different roles and expectations. Documenting how the AI system should function in such active environments must 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 justifying any future changes.

5.1.3 Deployment and post-deployment

AI systems may be designed using data and datasets from specific populations. As with any therapeutics, once deployed, the AI systems will be utilized by a larger population and potentially variable end-users. Careful deployment plans and justification for targeting different end-users should be considered. Manufacturers should be obliged to carry out post-market surveillance, which is the systematic process for collecting and analysing experience gained from AI systems that are considered to be medical devices that have been placed on the market *(14)*. Deviations from pre-specified plans, updates or changes to the AI system, post-deployment performance, data capture and approaches to continued assessment of the system should also be documented. Such approaches will be increasingly relevant once there is a wider understanding that AI systems may change after deployment.

5.1.4 Risk-based approach and proportionality

Regulatory frameworks recommend a risk-based approach with processes in place to identify and mitigate errors, biases and other risks in a manner proportional to their importance. A risk-proportional approach should also be considered for the level of documentation and record-keeping for AI systems. Developers of AI systems should keep in mind that regulatory organizations have avenues for dialogue and discussion that can be used to shed light on regulatory requirements.

5.2 Risk management and Artificial Intelligence systems development lifecycle approach

AI systems fall into many categories – e.g., devices that rely on AI and are used as medical devices (commonly known as SaMDs, which is short for "Software as a Medical Device"). Such categories 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*"*(15)*. However, the regulatory considerations for such a category of AI systems are similar to those of typical software that are regulated as medical devices, with the addition of considerations that may include continuous learning capabilities, the level of human intervention, training of models, and retraining *(15)*. Furthermore, a holistic risk management approach that includes addressing risks associated with cybersecurity threats to an AI system, and the system's vulnerabilities, should be considered throughout the total product lifecycle. This topic area aims to present a holistic risk-based

approach to AI systems in general, and to those used as medical devices in particular, throughout their lifecycle, including during pre- and post-market deployment.

5.2.1 AI systems during the development and deployment process

Figure 3 illustrates the process of development and deployment of an AI system. Developers and implementers should establish measures to ensure responsible development of AI systems.

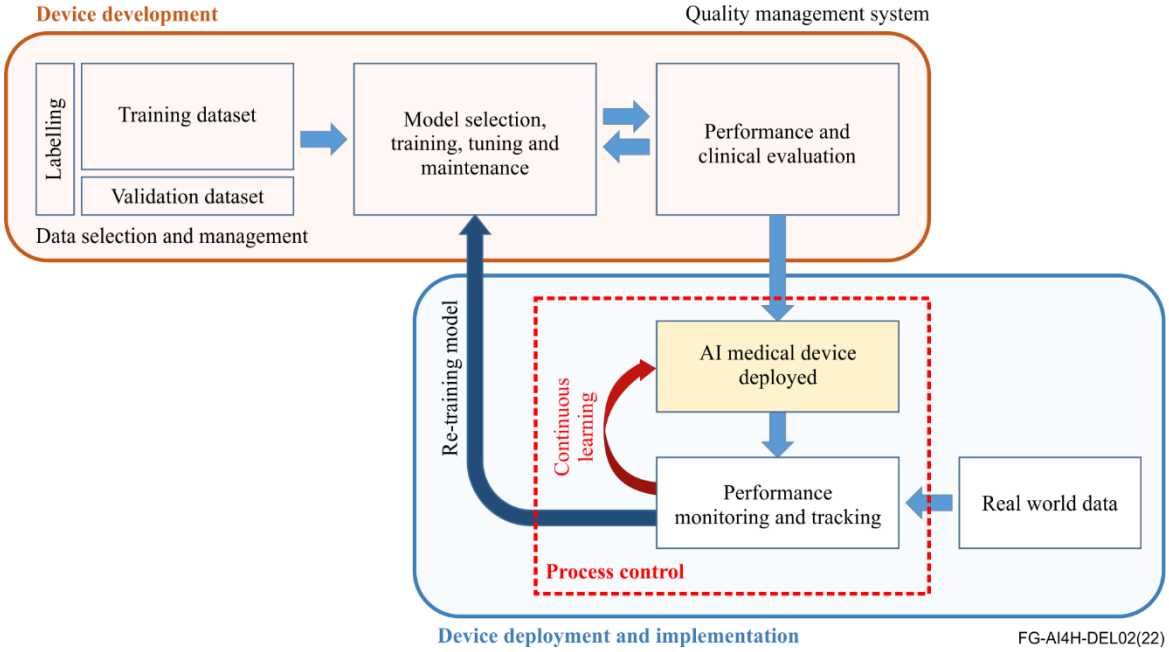

Figure 3 – The process of developing and deployment of the AI system *(16)*

Figure 3 shows that all activities related to the design, development, training, validation, retraining and deployment of AI systems should be performed and managed under a quality management system based on ISO 13485 *(16)*. For clinical endpoints, AI-specific monitoring dimensions include confidence *(17)*, bias and robustness *(18)*.

5.2.2 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 the development lifecycle approach involves appropriate development practices for the AI system. This approach could also potentially increase the trustworthiness, and assure performance and safety, of the AI system. An example is the Total Product Lifecycle (TPLC) approach *(4)* that could include the following four components (as illustrated in Figure 4):

- demonstration of a culture of quality and organizational excellence of the manufacturer of the AI systems;
- pre-market assurance of safety and performance;
- review of AI systems' pre-specifications and algorithm change protocol; and
- real-world performance monitoring.

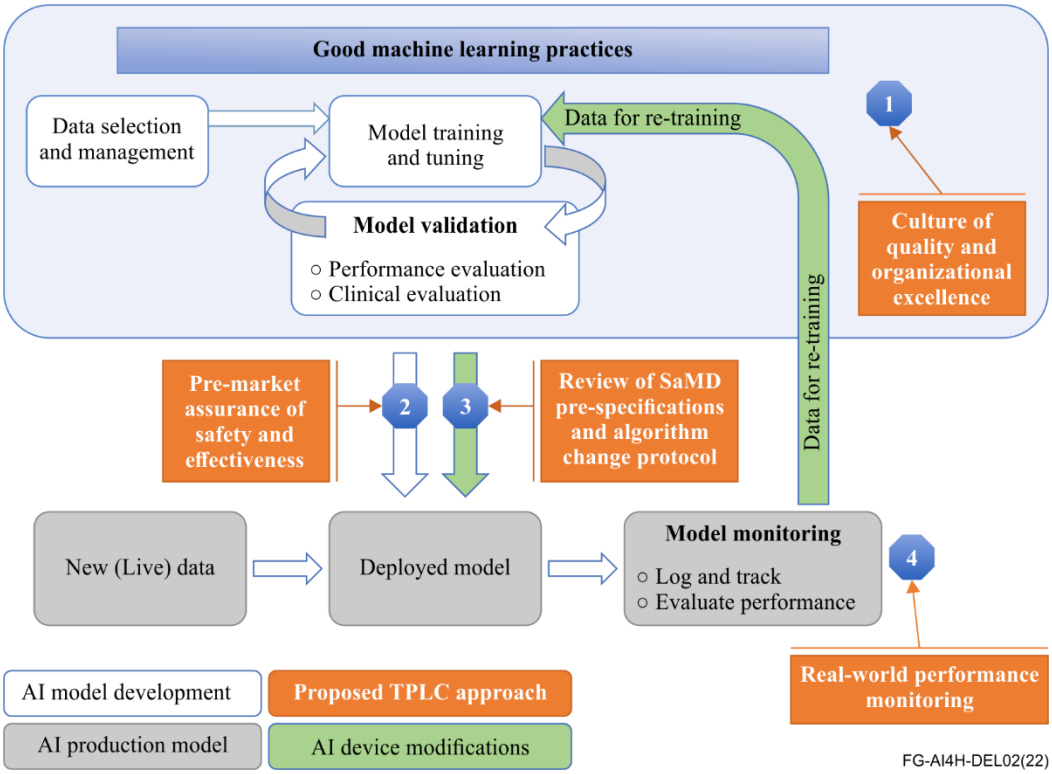

Figure 4 – AI system Total Product Lifecycle approach on AI workflow *(4)*

5.2.3 Holistic risk management

Holistic risk evaluation and management should be considered, taking account of the full context in which the AI system may be used. This could include not only the software or AI system that is 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 be considered throughout all phases in the life of a medical device. Consequently, manufacturers of AI systems should employ a risk-based approach to ensure that 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 the cybersecurity of the device by assessing risks and mitigations throughout the AI system's development life cycle. In order to achieve this, the IMDRF has published a security risk management process, as illustrated in Figure 5.

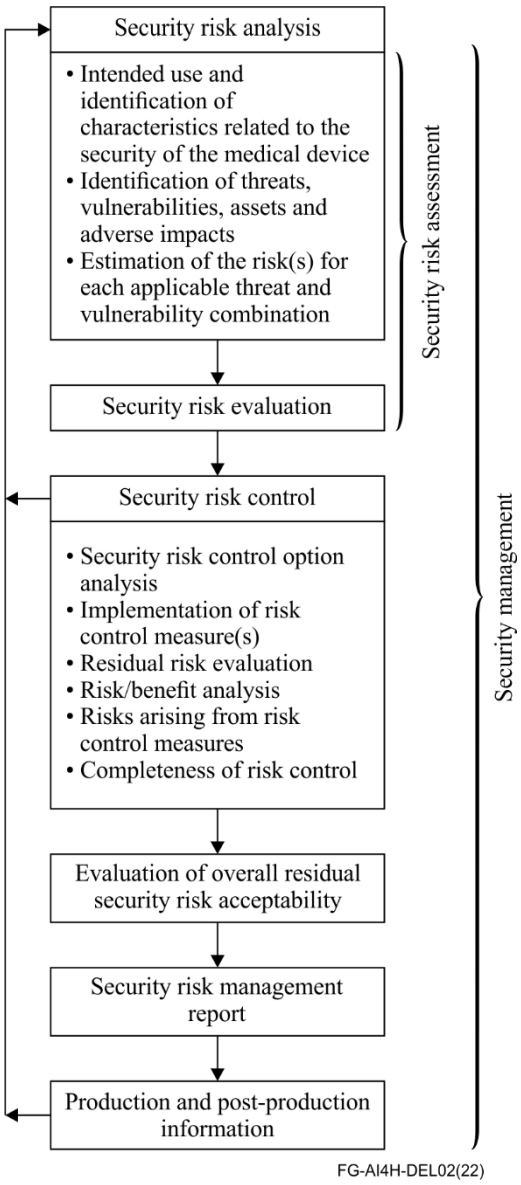

Figure 5 – IMDRF schematic representation of the security risk management process *(19)*

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

Figure 6 – General AI medical device risk management approach

Pre-market development management

There is a need for transparency regarding the functioning of any manufactured AI-based devices to ensure that users can have a better understanding of the benefits, risks and limitations of these AI-based systems *(20)*. In addition, the controls and measures put in place to ensure that a developed AI system functions as expected while minimizing risk of harm should be proportional to the risks

that could occur if the AI system were to malfunction. For instance, failure of an AI system that is designed to encourage adherence to a healthy diet is different from one that is designed to diagnose or treat certain diseases and pathologies. Therefore, developers should consider a risk-based approach through all processes to prioritize safety. Developers need to consider both the intended use of the AI system and the clinical context in order to evaluate the level of risk. For instance, the IMDRF risk framework for SaMD *(21)* identifies two major factors that may contribute to the impact or risk of an AI system. The first factor is the significance of the information provided by the AI system to the health-care 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 patient's health-care situation or condition – which is determined by the intended user, disease or condition, and the intended population for the AI system – i.e., critical, serious or non-serious health-care situations or conditions. Taken together, these factors relating to the intended use can be used to place the AI system into one of four categories from lowest risk (I) to highest risk (IV) to reflect the risk associated with the clinical situation and device use.

Table 2 – AI systems risk classification *(21)*

State of health-care situation or condition	**Significance of information provided by the AI system to the health-care 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 classification should be considered when testing different models and balancing trade-offs such as transparency and accuracy. In cases where training datasets are limited, simpler models, such as regression or decision-tree models, often provide equivalent or better results than more complex models and have the added benefit of more transparency and interpretability. On the other hand, in cases with larger and more complex datasets, 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 to build trust in the model through additional validation measures.

Furthermore, depending on the level of risk, some AI systems 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 on the basis of having stringent levels of surveillance in place with responsive back-up systems to handle any failure of the algorithm in order 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 implementers (i.e., users in health-care settings, such as medical practitioners). Appropriate documentation of risk management and proper auditing procedures are examples of ways that help assure transparency. In general, auditing of specific key components of the AI medical device should be considered (e.g., certain software, hardware, training data, failure cases). For instance, it is important to do version control with training data because more data are added with each update. If an algorithm suddenly deteriorates in performance after an update, an inspection of everything that contributed to the update may be desired. In most cases, the element 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-sense 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 health care. Such legislation includes 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). Compliance with relevant laws (local, 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.

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 the development, validation and clinical evaluation of the device. 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 *(14,16)*:

- any serious public health threat;
- death, serious deterioration in the state of health of patient, user or another person has occurred;
- death, serious deterioration in the state of health of patient, user or another person may 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).

Furthermore, manufacturers should proactively collect information (through scientific literature and other information sources such as publicly accessible databases of regulatory authorities, user training and surveys) as part of their post-market surveillance plan. The plan should outline how manufacturers will actively monitor and respond to evolving and newly-identified risks. Key considerations for the post-market surveillance plan include *(16)*: vulnerability disclosure, patching and updates, recovery and information-sharing. Additionally, 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 to relevant regulators.

In general there is a need for both post-market clinical performance follow-up and periodical safety checks to report any potential harm. The intensity of post-market surveillance by the manufacturer may be risk-proportionate (according to consequences of failure [creating potential risk of harm] and likelihood of early detection of such failure). Finally, post-market surveillance requires a minimum level of evaluation for each site in order to ensure that potential algorithm vulnerabilities due to variation in local environments can be detected.

For example, the AI Lab of the National Health Service (NHS) in the United Kingdom published guidance to accelerate a safe and effective adoption of AI in health *(22)*. The guide lists 10 questions in four categories to help buyers of AI products to make informed decisions, identify problems, assess products, and consider issues relating to implementation, procurement and delivery (Figure 7).

Figure 7 – The United Kingdom's National Health Service A buyer's guide to AI in health and care *(22)*

- **Change management**

In view of the character of AI systems, it is important that the regulatory system enables continuous modifications for improvement to be made throughout the AI system's development lifecycle. The term "change" refers to such modifications, including those performed during maintenance.

There are several proposed change management models and approaches for AI-based systems. Some consider change as part of the total development lifecycle (as in the TPLC approach) *(4)* (Figure 4). Other models focus on the change management process in the total lifecycle of medical device products which can be continuously improved. An example of this is the approach implemented by the Ministry of Health, Labour and Welfare of Japan and adapted in the Pharmaceuticals and Medical Devices Act as Post-Approval Change Management Protocol (PACMP) for medical devices *(23)* (Figure 8).

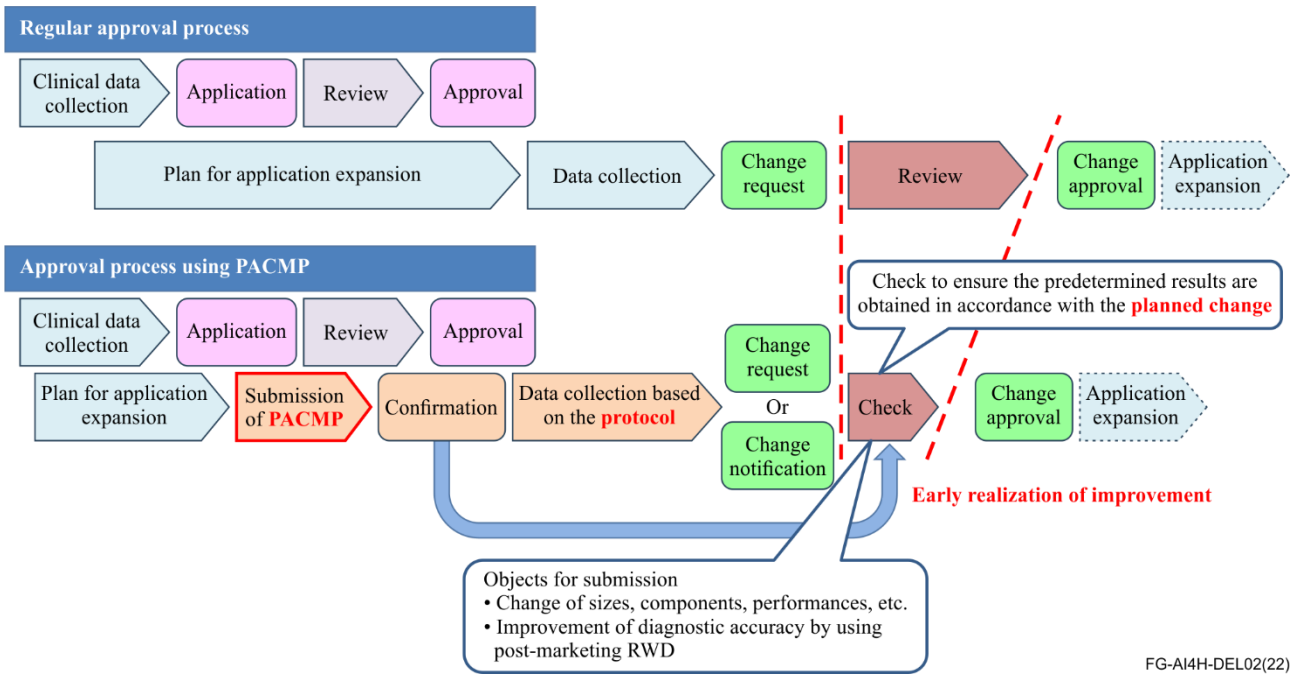

Figure 8 – Post-Approval Change Management Protocol for medical devices

5.3 Intended use and analytical and clinical validation

In principle, regulatory mechanisms are in place to answer the question: "Do the available data (included in the regulatory submission) support the conclusion that an investigational or experimental AI system is safe and performs sufficiently well to justify entry into the market and public access?" In addition to the principles discussed in clauses 5.1 and 5.2, one also must consider assessing if the use of the system is safe (i.e., it will not harm the user, the patient or other persons) and if the claims made about its performance can be verified (see Figures 9 and 10). Evaluation of these claims for AI systems requires a clear use case description, demonstration of analytical and clinical validation, and assessment of the potential for bias or discrimination in the AI system.

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

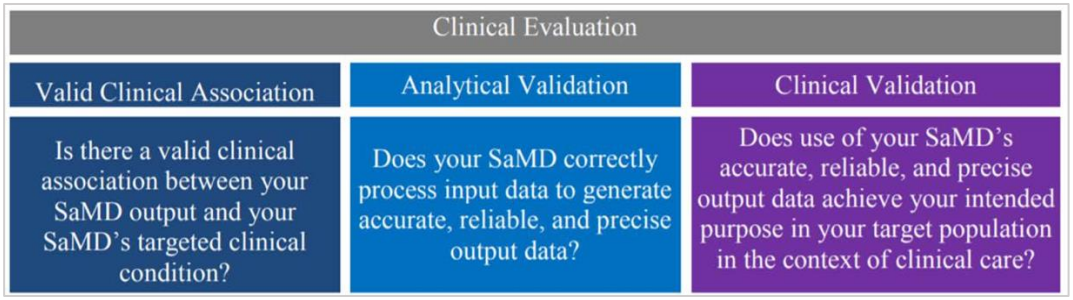

Figure 10 – IMDRF description of clinical evaluation components *(4)*

5.3.1 Use case description, analytical and clinical validation

Clinical evaluation is the review of evidence that demonstrates the safety and performance of a given product for a given intended use. For AI systems (especially devices that rely on AI and are used for medical purposes), guidance is useful for collecting evidence of analytical and clinical validation. The performance of AI systems can be changed rapidly – not only as a result of a code change but also to provide different or additional training/tuning data. Consequently, clinical evaluation that takes account of TPLC from development to analytical and clinical validation and to post-market surveillance should be considered for AI systems.

This topic area covers the considerations of use case descriptions (including statements of intended use) and analytical and clinical validation. These considerations follow the framework proposed by the WHO/ITU FG-AI4H Working Group on Clinical Evaluation (WG-CE) *(24)*. A full description of this framework can be found in the deliverable for the WG-CE. The following clause describes the key considerations and best practices, and builds on the important work of national and regional regulatory authorities and international bodies such as IMDRF. It is not intended to replace the work of these bodies. By outlining key considerations, this report draws attention to challenges that remain in this rapidly changing field. For instance, particular consideration is given to under-resourced settings which may have limited regulatory capacity at national level. The role of benchmarking in the evaluation of AI systems in health is also explored. Evaluation principles are applied to this topic area, and to the work of the WHO/ITU FG-AI4H in which benchmarking evaluation is a key component *(25)*.

5.3.2 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 into new clinical pathways altogether (e.g., into new telemedical pathways or as part of new triage tools). Therefore, for AI systems, 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 be utilized safely and appropriately and, in regulated AI systems (the statement of intended use), to allow regulators to assess whether the evidence of the analytical and clinical validation steps is appropriate and sufficient for the intended use.

When developing a health-related AI system, it is important to 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 (health-care provider or patient) and the clinical situation for which it is intended. Many interventions, tests and guidelines are prone to bias, and this is a particularly important consideration for AI systems which are highly sensitive to the characteristics of the data they were trained on and are prone to failure with unseen data types (such as a new disease feature or population type or context that was not previously encountered). Developers and manufacturers should also provide a clear clinical and scientific explanation of their tool's intended performance, including the populations and contexts for which it has been validated for use. Standardized reporting templates common to all stakeholders can help to communicate the intended use more effectively *(26, 27, 28)*. For some intended use cases there may be clear reasons why analytical performance of the tool would differ in different settings *(29)* (e.g., a symptom checker may perform differently in areas with a disease epidemiology that is different from 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 another very similar tool, has been shown to have similar analytical performance in a wider setting than those in which the tool was initially developed and validated *(30)* (e.g. retinal tools have been shown to have a similar performance in different populations *(31)*). Understanding of the generalizability of similar tools may be taken into account when providing a statement of the intended use or description of the use case *(32)*.

As part of the risk management process, regulators may wish to request evidence that developers have considered whether there are situations in which a tool should not be used (e.g., if there are 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 the intended settings.

5.3.3 Analytical validation (also referred to as technical validation)

For the purposes of this document, analytical validation refers to the process of validating the AI system 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 and manufacturers should provide a description of the datasets used in the AI system's training, tuning, testing and internal validation. The description of the intended use case (which can be on standardized reporting templates) should cover the size, setting, population demographics, intended user and clinical situation (with input and output data). Transparency and documentation on dataset selection and characteristics are critical to ensure that AI systems are used appropriately. Developers and regulators may expect that the AI system has been externally validated in a dataset different from that in which it was trained and tested in order to demonstrate the model's external validity and generalizability beyond the original dataset. The external validation dataset is expected to be representative of the setting and population that are described in the intended use (gender, race, ethnicity) in order to demonstrate robust performance in the intended setting. The validation dataset

should be of adequate quality, with appropriate robustness of labels. As part of the risk management process, it is important to identify any cases that are, or may be, high-risk *(28)*.

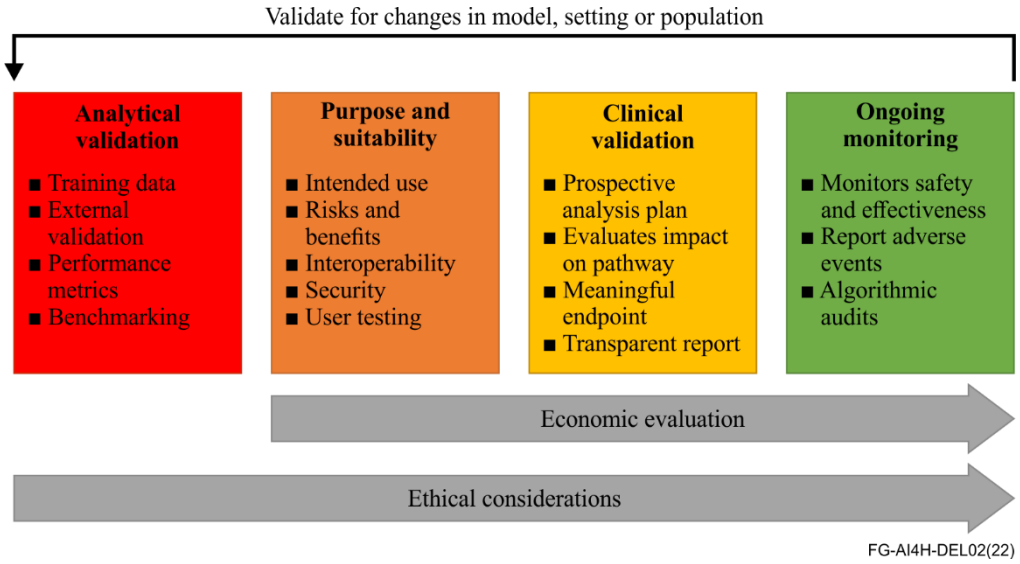

Figure 11 – Overview of framework for clinical evaluation of AI models in health developed by the WG-Clinical Evaluation

Although bias, errors and missing data are not unique to AI development, they are nevertheless 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., health-care settings). Reporting the gender, race and ethnicity of persons in the training and validation data cohorts, if feasible, can help to address the potential for bias and can avert its impact. For example, a better understanding of bias may help identify populations for which an AI system may not function as expected. Post-market surveillance can also provide insights into the impact of potential bias.

Obtaining datasets for training, testing and validation that are sufficiently representative and of sufficient quality can be difficult. Local, regional and national bodies interested in procuring AI systems could hold their own hidden dataset to enable external validation (e.g., a recent scheme of the United Kingdom's NHSX has nationally-representative datasets for some common use cases). Access to representative datasets for validation is a particular concern in many low- and middle-income countries. Where datasets are available in low-resource settings, there may also be limitations in 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 may be barriers that impede access to diagnosis and treatment. These major challenges – which have the potential not only to propagate inequality of access but also to compromise safety and performance of AI-based tools – are potential areas for future work. In this regard, the newly launched International Digital Health & AI Research Collaborative (iDAIR) *(33)* notes that collaborative, distributed and responsible use of data is at the heart of its strategic plan.

While most regulatory agencies have national or regional remits, some countries with limited regulatory capacity tend to rely on decisions made by other major regulators. The 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 dossiers provided to other bodies. However, the performance of AI-based systems is highly dependent on the context. In order to rely on regulatory review and decisions, many regulators (whether national or regional) could perform analytical validation as a second local validation step to ensure that the performance metrics obtained are consistent with those demonstrated in other regulatory jurisdictions.

This could be best prioritized through a needs-based approach – e.g., 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 an AI system, 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 (e.g., derived from logistic regression) with strong evidence-based or mandated standards of accuracy, sensitivity and specificity (such as for screening tools). Depending on the intended use case, the requirement for comparative performance may be more or less stringent (e.g., 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 *(34)*. 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 that AI system to understand potential similarity of performance. Benchmarking software is being developed as part of the work of the Open Code Initiative *(35)*. Platforms such as this may also be useful as ways 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 has been used only to understand comparative analytical performance. In addition, repeatedly using the same data for benchmarking multiple updated models (and thus, even if inadvertently, for training 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 *(36)* provides guidance on the methods, processes and software development for the analytical validation of health-related AI systems *(28)*.

5.3.4 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 for evaluation of other factors that may affect performance of the tool (e.g., user interaction, workflow integration, and unintended consequences of the 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 (including significance of the information provided by the tool and the state of the health condition) *(37, 38)*. The IMDRF document on clinical evaluation of SaMD (Table 2 *(21)*) proposes that devices in category I are the lowest-risk tools that have evidence of analytical validity, 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 on the basis of evidence of analytical validity. There is no universal agreement on the appropriate level of evidence of adequate clinical performance for a novel AI tool before deployment and this is the subject of a separate working group within the FG-AI4H (Working Group on Clinical Evaluation).

Randomized 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 *(12)*. However, currently there remain a small number of actively recruiting or completed randomized trials in this field *(39)*.

Randomized clinical trials have potential limitations that may make other options preferable (trials can be slow, or expensive, and may evaluate performance in specific groups under trial conditions). Where randomized evidence may not be necessary (e.g., the 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. Developers should provide a description of the steps taken to perform clinical validation in a context similar to that available in 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 WG-CE.

In some situations, as described below, special considerations apply. For instance:

5.3.5 Post-market monitoring

Post-market monitoring in some regulatory contexts relies heavily on reporting of adverse events. Recent WHO guidance recommends that proactive post-market surveillance must be carried out by the manufacturer. As part of a TPLC 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 combined with 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.

5.3.6 Changes to the AI tool

An update of an AI tool by a change of code, 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 with the concepts in this document and with the 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, 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.

5.3.7 Low- and middle-income countries

There is considerable variation in the implementation regulation for medical devices, and therefore also in deployed AI technologies and developed AI systems. Some countries 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 for regulatory bodies and other interested stakeholders, some of whom have aligned remits. Furthermore, there are important regulatory considerations 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 a relevant medical specialist

(e.g., interpreting retinal scans, histopathology slides or radiology images). In high-income countries, AI systems are more often positioned as an adjunct to medical professionals. Using an evaluation performed to support regulation in a high-income setting to inform how such AI systems are used in low- or middle-income settings may therefore not be appropriate. Thus, the full context of health-care 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 there is less attractive. Regulatory agencies in high-income countries could support this adaptation, which could also increase the generalizability and robustness of AI systems. However, this would require adaptive studies to ensure wider use in low- and middle-income countries or the use of incentives to encourage additional development, testing and validation. The availability of a range of representative datasets would support local analytical validation. Finally, AI systems for health 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-income to low-income counties, with the possibility of lower performance without appropriate evaluation or tuning with local data.

5.4 Data quality

5.4.1 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 are the most important ingredient for training AI/ML algorithms, and can be classified on the basis of format, structure, volume and many other factors. Data can take any form, including character, text, words, numbers, pictures, sound or video. Also, these data can be structured, semi-structured or unstructured *(9)*. Structured data are 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 or volume of the data, a more general classification can be based on the following 10 Vs of data *(9)* (as illustrated in Figure 12): Volume, Veracity, Validity, Vocabulary, Velocity, Vagueness, Variability, Venue, Variety and Value.

5.4.2 Good quality data in health AI systems

All AI tasks and solutions use some form of data, regardless of their characteristics, to facilitate machines to learn, adapt and improve on 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 10 Vs perspective is crucial to assess data quality in AI systems during development and even afterwards. Clause 5.4.3 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.

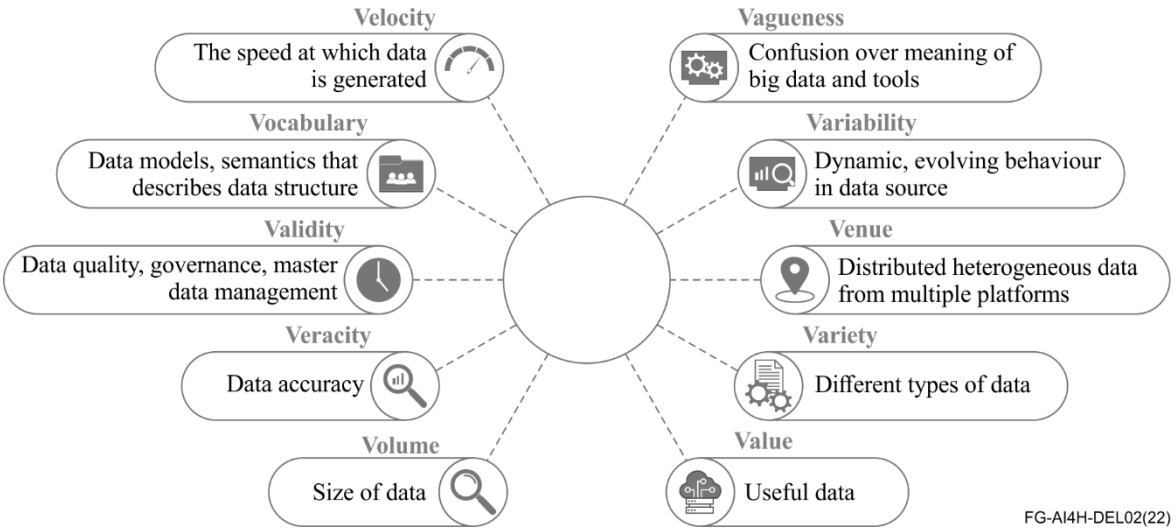

Figure 12 – The 10 Vs of data *(9)*

5.4.3 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 the purpose, and thus developers should determine if available data are of sufficient quality to support the development of systems that can achieve their intended goal. The lack of good-quality datasets for use in the development of AI systems may hinder their effectiveness and potential benefits. Data that are not of sufficient quality for the intended purpose can also lead to many problems, such as bias and errors. Some data quality issues that often arise when developing AI systems, and that need to be considered by all stakeholders, are discussed in this clause and summarized in Table 3. These issues and considerations can relate directly to dataset management, the ML model, the infrastructure used to manage the data, or general governance aspects, as follows:

- **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 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 in order to ensure that data quality and integrity are maintained throughout the lineage of the data. This is also essential for knowledge management and transfer in a highly evolving field. Further, in addition to the handling of the data, the capacity to plan for and conduct data analyses is also important.
- **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 and generate a coherent dataset for the purpose comprehensive analysis – which is especially challenging when using health-care data. For instance, 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).

- **Dataset selection and curation**. Knowing the source of data and making an initial assessment of the data quality can help to determine the potential for selection and information bias. Selection bias results when the data used to produce the model are not fully representative of the actual data that the model may receive or of the environment in which the model will function. In addition to selection bias, measurement bias is another relevant issue that results when the data collection device causes the data to be systematically skewed in a particular direction. Consequently, 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 the information 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, the mitigation of bias and assessment and mitigation of other risks to data robustness remain essential for a heterogeneous dataset.
- **Data usability**. It is essential to know whether the data used for development of the algorithm was intended for that training, so developers need to convey their full understanding of the dataset and why it was suitable for their purpose. For instance, 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 an AI model intended to diagnose a disease or condition.
- **Data integrity**. Data integrity can be defined as "the completeness, consistency, and accuracy of data" *(40)*. Lack of data integrity is an 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. Such checks should also be the first step in data preparation for any ML workflow.

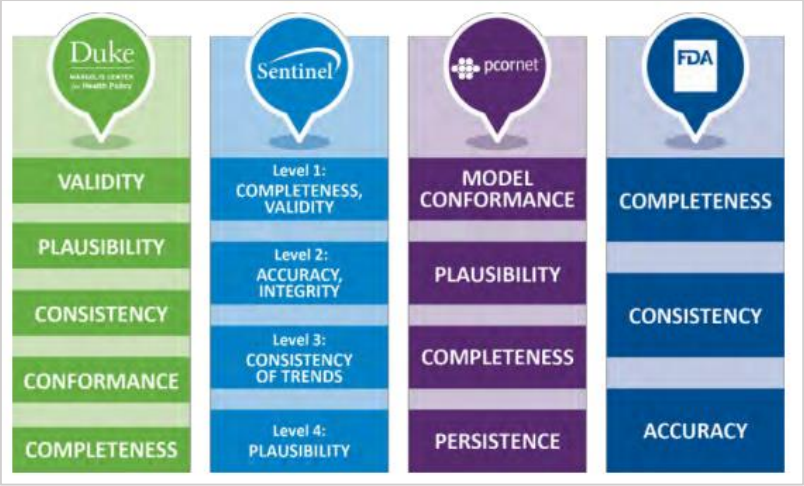

Figure 13 – Examples of quality check principles (41)

- **Model training**. AI algorithms are usually trained on a separate dataset (called the training dataset) and validated on a different dataset in order to measure the performance of the algorithm reliably. Training datasets should be well represented (e.g., by considering the prevalence of a disease/condition) to avoid "class imbalance". Medical record data is inherently biased, and therefore it is necessary to incorporate non-medical data such as the social determinants of health *(42)*. Furthermore, under-representation of important diagnostic features may limit the performance of the model and cause bias. This can be avoided by ensuring that inclusion and exclusion criteria at the patient level and the data input level do

not create a selection bias. Furthermore, when ensuring that the datasets reflect the setting in which the model will be applied, a lack of diverse data (age, race, geographical areas) could limit the generalizability and accuracy of a developed AI system. This is demonstrated by a recent study by Stanford University *(43)* which showed that 71% of patient data from just three US states train most of the AI diagnostic tools used in the United States of America.

- **Data labelling**. It is important to ensure consistent, reliable and accurate labelling of datasets for testing in line with 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 with patient outcomes, and whether the reference standard follows published metrics.
- **Documentation and transparency**. The algorithm and data supporting it are often not available or are not well documented for all AI system stakeholders. This makes it difficult to assess the quality of the underlying data. Transparency and careful documentation are important not only with regard to the methodology used in collecting data, but also for the selection and modifications of datasets used for training, validation and testing. Good documentation is fundamental to achieve transparency that enables verification and traceability. Transparency of methods should ensure data quality. Beyond the CONSORT-AI and SPIRIT-AI reporting guidelines, checklists have been devised by the machine learning community to report representativeness, completeness and other data quality characteristics *(44, 45)*.

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

The list in Table 3 summarizes the key data quality considerations for AI system safety and effectiveness.[5]

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 versus "cleaned" data
	Data wrangling and cleansing
	Parameters and hyperparameters
	Usability
	Characterization
	Labelling
	Dependencies

5 This list will be updated and harmonized with the work of the IMDRF.

Table 3 – General data quality considerations

Category	Data quality consideration item
	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 versus dynamic AI
	Open AI versus 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

5.5 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 standard of safety and security. Therefore, the strategy 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 the 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 clause discusses high-level considerations for privacy and data protection. For other ethical considerations, refer to the deliverable of the Working Group on Ethical Considerations on AI for Health[6] *(46).*

5.5.1 Current landscape

As the demand for health-related data increases, the protection of privacy is creating a unique challenge for all stakeholders wishing to benefit from the many opportunities created by 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, ensuring that large-scale datasets are secure from unauthorized access at each stage of the development process – collection, storage and management, transport, analysis, sharing and destruction – is an important consideration.

Some 145 countries and regions have data protection regulations and privacy laws that regulate the collection, use, disclosure and security of personal information *(47).* 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 European Union (EU) and elsewhere. These laws are often complex and may have conflicting obligations. When developing an AI system for therapeutic development or health-care applications, early in the development process the developers should consider gaining an understanding of applicable data protection regulations and privacy laws, including special regulatory provisions related to sensitive information such as genetic data. Developers should also consider national laws as well as regional ones. For instance, 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's 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 at least to evaluate the extent to which they are subject to it. Most privacy laws, including Singapore's Personal Data Protection Act, apply only to personal data processed within the country, whereas the GDPR[7] may apply to the personal data of EU citizens, regardless of the location where data are processed.[8] As a result, companies subject themselves to compliance obligations under the GDPR if they are located in the EU (including if any component of the organization is located in the EU), if they offer goods and services to individuals located in the EU, or if they monitor the behaviour of persons located in the EU.

---

6 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.

7 See also India's proposed Personal Data Protection Act.

8 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.

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 1 of the United Kingdom'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 United Kingdom's data protection law in the mistaken belief that the processing does not involve personal data *(48)*. Consent requirements also vary according to the jurisdiction. For instance, various jurisdictions may require "explicit consent", with heightened information requirements for the processing of health-related data (GDPR Article 9) *(49)*. Therefore, developers may wish to consider the varied legal contexts when documenting how they address privacy-related concepts, including measures taken to meet consent requirements, and 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 may have a significant impact on the way in which data are 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 whether a recipient country has thresholds of data protection laws and protections "essentially equivalent" or "substantially similar" to the jurisdiction from which the data were 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 the third country is considered by the EU Commission to have an "adequate" level of protection.[9] As of May 2023, the EU Commission had recognized only 13 countries as providing adequate protection *(50)*.

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

5.5.2 Documentation and transparency

Documentation and transparency are critical to facilitating trust with regard to privacy and data protection. Detailed privacy policy disclosures provide regulators with a benchmark by which to examine a company's handling of data. 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 – e.g., the types and sources of health data collected and processed; the identities 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 were collected (including whether adequate notice was provided to the data subject and how consent requirements were met); and technical and organizational information on the storage of data, including security measures.

---

9 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 in 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*").

Developers must 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 *(49)*.[10] Privacy protections should not be limited only 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) *(51)* 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 should develop their risk management practices accordingly (Figure 14).

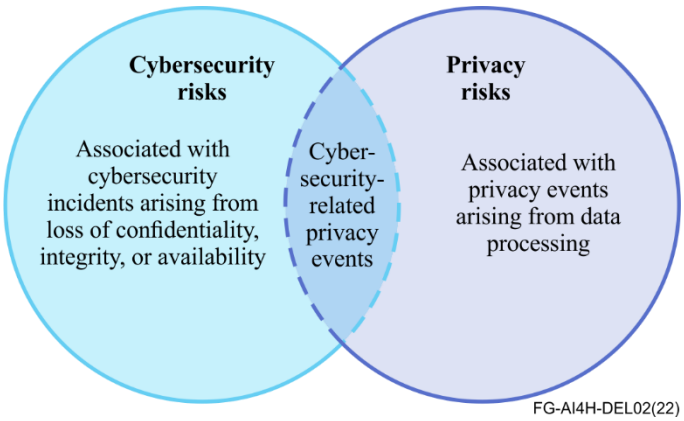

Figure 14 – NIST Privacy Framework – cybersecurity and privacy risk relationship *(51)*

A compliance programme should consider risks and should 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 account of 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 (GDPR Articles 25 and 32) *(49)* and India's proposed data privacy by design policy *(52)* – while others include the duty to implement and maintain reasonable security practices and procedures appropriate to the risk.[11] 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

[10] 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.

[11] For example: 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.

(GDPR Article 35) *(49)*[12], require companies to conduct these assessments.[13] Although United States of America's law does not require privacy impact assessments, the US Department of Commerce National Institute for Standards and Technology (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" *(51)*. 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 *(53)*. Such explanations and documentation should be available at different levels of detail, targeted at different audiences – regulators, managers, developers, operators and users. The nature of the information and explanations required may differ, but all 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 shows not only that controls have been applied but 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 consideration by developers of medical devices. Although an in-depth discussion of cybersecurity requirements is outside the scope of this clause, 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 wish 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 *(54)*. For instance, blockchain-based technologies may be one mechanism for protecting data privacy, security and integrity for AI in a traditionally fragmented health information systems ecosystem for national and regional contexts *(55)*.

5.5.3 AI regulatory sandboxes

The above regulatory challenges are recognized by regulatory authorities and policy-makers across the world *(56)*. As a result, over 50 countries are currently experimenting with sandboxes in a wide range of high-technology sectors – notably in the financial sector but sandboxes have also gained popularity for health and legal services *(57)*. The regulatory sandbox approach has gained considerable traction as a means of helping regulators to address the development and use of AI and other emerging technologies *(57)*. Regulatory sandboxes are generally regulatory tools that allow the flexibility to test innovative products or services with minimal regulatory requirements *(57)*. Consequently, regulatory sandboxes are considered an agile approach to innovation and regulation and thus regulatory authorities are increasingly favouring them. In the EU, regulatory sandboxes have been proposed for testing surveillance solutions in the fight against the COVID-19 pandemic, and for establishing a framework for EU-wide data access. In relation to AI regulations specifically, the first AI regulatory sandbox pilot presumably launched in 2023 by the Government of Spain with an aim

[12] "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".

[13] While risk assessments are quite common in information security standards and requirements, they are rarely seen in privacy rules in the United States of America. The GDPR, however, requires that an organization processing personal data must conduct a specific Data Privacy Impact Assessment or DPIA before beginning the processing.

to provide a guide to all EU Member States and the European Commission *(58)*. Although AI regulatory sandboxes raised a few concerns, they have the potential to bring many key benefits to AI system regulators, developers, manufacturers and even patients *(57)*. This is because such AI regulatory sandboxes can: 1) help enable a better understanding of the AI systems during the development phase and before they are placed on the market; 2) facilitate the development of adequate enforcement policies and technical guidance that can mitigate risks and unintended consequences; and 3) foster AI innovation by establishing a controlled experimentation and testing environment for innovative AI technologies, products and services for new and potentially safer AI systems.

5.6 Engagement and collaboration

Where applicable and appropriate, engagement and collaboration between developers, manufacturers, health-care practitioners, patients, patient advocates, policy-makers, regulatory bodies and other stakeholders can improve the safety and quality of an AI system. Many regulatory bodies have adopted engagement and collaborative approaches in this area, and this clause discusses the approaches of five of them: the United Kingdom's MHRA, the South African Health Products Regulatory Authority (SAHPRA), the European Commission, Singapore's HSA, and the U.S. 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.

Clause 5.6.2 examines two examples of engagement and communication between regulators and AI developers resulting in positive clinical outcomes (CURATE.AI and IDentif.AI). The last clauses consider the practical implications for engagement and collaboration in resource-limited settings and recommend ways that regulatory bodies can initiate this process even in countries without past experience in engagement and collaboration. 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, and 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 health care and therapeutic development

	With whom?	Why?	How?
1. Medicines and Healthcare Products Regulatory Agency (MHRA), United Kingdom	***Examples of stakeholders with whom the MHRA engages and collaborates:*** • **Patients/patient advocates** • **Academia** • **Health-care professionals** e.g., providers in the National Health Service (NHS) and private health-care providers. • **Industry** e.g., medical device and in vitro diagnostics industry, health technology industry. • **Domestic government partners** e.g., Department of Health and Social Care (DHSC), NHS England and Improvement, NICE, and Care Quality Commission (CQC).	***Examples of reasons why the MHRA engages and collaborates with stakeholders:*** • Alert users to problems with medical devices and medicines. • Answer enquiries about roles in regulation or raise awareness of safety issues. • Seek feedback on development of regulatory policy, managing adverse incidents and risks. • Interface with United Kingdom 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 health-care 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 *(59)*. • Working-level meetings with national stakeholders, bilateral meetings with other parts of NHS, government and international counterparts.

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

	With whom?	Why?	How?
2. South African Health Products Regulatory Authority (SAHPRA), South Africa	***Examples of stakeholders with whom SAHPRA engages and collaborates:*** • **Patients/patient advocates** • **Academia** • **Health-care 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 that the Conformity Assessment Body network is established in the country to certify the 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 TGA *(60)*.
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 *(61)* 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 elements linked to the digital transformation of health and care, such as financing and investment proposals and enabling technologies."
4. Health Sciences Authority (HSA), Singapore	***Examples of stakeholders with whom the HSA engages and collaborates:*** • **Academia** (e.g., research institutions). • **Health-care 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 *(16)*.	• Rapid, streamlined engagement portals are available for several facets of product regulation *(62)*. • Specific processes that can be straightforwardly addressed include Medical Device Information Communication System (for application submissions for licences, 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.

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

	With whom?	Why?	How?
		• To work with other agencies responsible for implementation and deployment of AI and software medical devices in the health-care system to facilitate greater adoption of innovative technologies in the health-care system.	• **Medical Device Development Consultation**: Online appointment booking system that allows innovators and developers to seek scientific and regulatory advice during the 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. Food and Drug Administration (FDA), United States of America	***Examples of stakeholders with whom the FDA engages and collaborates:*** • **Patients/caregivers/patient advocates** • **Academia** (e.g., research institutions). • **Health-care professionals** • **Industry** (e.g., developers, device manufacturers, drug companies, trade associations). • **National government partners** (e.g., National Institutes of Health [NIH], Office of the National Coordinator for Health Information Technology [ONC], Federal Communications Commission [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 obtain feedback from patients, industry and other stakeholders.

5.6.1 Discussion on strategies of profiled regulatory bodies

Table 4 shows the approaches of four national and one regional (in the case of the EC) regulatory body to foster engagement and collaboration. In the first category ("with whom?"), there are considerable similarities between these bodies. The shared targets for engagement and collaboration include health professionals (indicated by FDA, SAHPRA, MHRA, EC and HSA), academia (FDA, SAHPRA, MHRA, EC and HSA), industry (FDA, SAHPRA, MHRA, EC and HSA), patients or patient advocates (FDA, SAHPRA, MHRA and EC), domestic government bodies (FDA, SAHPRA and MHRA), media (national and trade press; FDA and MHRA), health providers (FDA and MHRA) and consumers (FDA and MHRA). Interestingly, the strategy paper by the US Department of Commerce's NIST also refers to academia and domestic government bodies as targets for engagement and collaboration.

In the second category ("why?"), SAHPRA notes the importance of communicating the benefits and intended use of devices, presumably to protect and promote public health (listed by the FDA and implied by MHRA). The FDA also stresses the importance of bilateral communication with stakeholders so that regulators are aware of developments in industry (or academia) and so that these stakeholders, in turn, are aware of 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 the EC, namely to facilitate approval of innovative solutions and support the digital transformation of health and care. The HSA acknowledges the importance of early engagement with innovators and developers to provide greater clarity in regulatory requirements and improve transparency in regulatory processes.

For the third category ("how?"), the 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 a 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 manufacturers, are not familiar with regulatory requirements *(60, 63).*

5.6.2 Two successful instances of engagement

To understand the value of engagement and collaboration between regulatory bodies and stakeholders, two real-world examples (Case 1 and Case 2) are described. 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 realized seamlessly. One aspect is in 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 *(64, 65, 66)*. In this context, integrating regulator accessibility with emerging innovation, sometimes in urgent circumstances, will ultimately result in life-saving outcomes. Importantly, these outcomes will not be confined to post-approval treatment but also to substantial patient benefit during the investigational stages of validation.

In Case 1, the developmental roadmap and validation of CURATE.AI and foundational technology of IDentif.AI were discussed with the Medical Devices Branch *(16)* of the 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 assess effectively the personalized medicine treatment outcomes and regulatory routes. A broader discussion on how clinical trial designs may evolve due to the emergence of AI was also conducted *(68, 69, 70).* 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 on the basis of (for example) the individualized dosage calibrations of the drug regimen (clinician-selected regimen), serial efficacy and toxicity measurements, efficacy-guided treatment protocols, 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) *(71)* for the potential rapid implementation of 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). Unlike traditional AI-based approaches, this strategy did not use 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 with potential trials being designed with clinical partners. With regard 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 part of the submission of a clinical validation programme because it stipulates the prerequisite of an initial assessment of device risk from the HSA *(72)*. The submission portal and portal interaction were particularly straightforward to navigate and were 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 was 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.

5.6.3 Recommended approaches for countries without past experience

For countries with limited experience in engagement and collaboration (and/or limited resources), it is important to establish: 1) what levels of engagement and collaboration are desired; 2) what steps can and should be taken to achieve those levels; and 3) what challenges are presented by the technology (e.g., AI explainability).

In many cases, it is desirable to adopt regulatory models that are adaptable, flexible, modular and scalable in order to account for the uncertainties of innovation through appropriate oversight and coordination. These features fit not only the specific challenges of emerging technologies but also of the regulatory approach of countries without past experience in this field or with scarce economic resources. On the one hand, priorities should be scalable so that growing amounts of work can be suitably addressed by adding resources to the regulatory model. On the other hand, however, priorities should be determined in accordance with the modular adaptability of the steps and levels of engagement. In ecology, adaptability applies to 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 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 and to align with society's preferences on emerging innovation, while allowing regulators to gain a growing understanding of technological challenges with increasing normative granularity *(73)*.

5.6.4 Narrative on using engagement tools based on practical experience

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

The Australian Government's recommended five-step model for engagement *(60)* 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 the engagement (based on what it is hoped to achieve) and identifying the relevant stakeholders. When planning 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 the stakeholder would be open to having). For instance, a digital health developer building an application (app) to support parents with children above a healthy weight may find that the primary health body concerned is an influential stakeholder which sets policies on managing children's weight. However, this 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. However, a developer will want to work with parents of children above a healthy weight to co-design the app and ensure that it fits their needs. It would, therefore, be important for the developer to "collaborate" with a representative group of parents and establish two-way or multi-way communication and shared learning and decision-making over the course of the project.

A similar approach for making sure that stakeholders are provided with the right information at the right time and are using optimal communication channels is outlined by one of the leading product development software companies *(74).* Within the stakeholder communication "play", importance is placed on who the stakeholders are, the desired method of communication and the frequency of communication. For instance, an internal government project developing a digital health product will have internal stakeholders (such as funders of the project and policy leads) and external stakeholders (such as leading academics). The communications plan should outline how each stakeholder group will be addressed (email, face-to-face conversation, video call, and/or social media) and how often there were will be 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 (for which examples of headings might be: method, audience/stakeholder, content to share, why, and frequency) for the whole development team to follow.

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

As demonstrated in Table 4 and discussed in the subsequent text, multiple regulatory bodies emphasize the importance of open (bilateral) communication with stakeholders so that regulators are aware of developments in AI-based technology and so that these stakeholders, in turn, are aware of changes in regulation. This is because AI-based technology is constantly changing and regulation needs to be able to keep pace. The development, deployment, post-market surveillance and iteration of AI products and services in health care should therefore be an ongoing conversation between developers and regulators.

It is recommended that regulators look at AI-based technology in health care from a mindset of accessible engagement that potentially, when applicable, facilitates 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), and 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. This means that a single, clearly marked pathway should be established and should be followed by an AI developer when ensuring the compliance of a product or service. Regulators need to collaborate with each other on issues such as clear messaging to developers and consistent levels

of engagement with developers at the right point, and by sharing what they learn from different engagements with developers.

If a country wishes to take an accessible engagement approach to the regulation of AI products and services, co-regulation could be explored. As outlined by Clarke *(75)*, in a co-regulation approach regulators outlined a regulatory framework based on required compliance to the legislative act(s). The details of how this is applied in practice are jointly developed by regulators and a representative sample of developers *(75)*. Similarly, when considering 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 by engaging them in the design and implementation of the regulatory process. The approach involves designing a regulatory process that reflects and acknowledges the needs of developers and not just those of the regulatory body and associated groups. Ultimately, however, regulators must remain fully independent of developers in order to make decisions that put the safety of the public first, as well as ensuring that public and private health-care resources are used only for technologies that meet independently developed standards of quality, safety and efficacy.

6 Recommendations for the way forward

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

Table 5 – Key recommendations for regulatory considerations on AI for health based on each of the six topic areas

Topic area	Recommendations
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. These 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 a risk 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	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.

Table 5 – Key recommendations for regulatory considerations on AI for health based on each of the six topic areas

Topic area	Recommendations
	3.2 Consider demonstrating performance beyond the training dataset through external, analytical validation in an independent dataset. This external validation 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 validation dataset and performance metrics should be provided. This external validation dataset should be appropriately independent of the dataset used for the development of the AI model during training and testing.
	3.3 Consider a graded set of requirements for clinical validation based on risk. Randomized clinical trials are the gold standard for the 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 are 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 early identification of data quality issues, which could potentially prevent or mitigate possible resulting harm.
	4.4 Consider mitigating data quality issues that arise in health-care 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 that the development process meets or exceeds such legal requirements.
	5.3 Consider implementing a compliance programme that addresses risks and develop privacy and cybersecurity practices and priorities that take into account potential harm and the enforcement environment.

Table 5 – Key recommendations for regulatory considerations on AI for health based on each of the six topic areas

Topic area	Recommendations
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 in order potentially to accelerate practice-changing advances in AI.

7 Conclusion

WHO recognizes the potential of 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 health care. Furthermore, with the increasing availability of health-care 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 health care and therapeutic development. For this reason, 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 systems in health care. The 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 health care and therapeutic development. This publication, which is based on the work of the WG-RC, is an overview of regulatory considerations on AI for health that covers the following six general topic areas: Documentation and transparency, Risk management and the AI Systems Development Lifecycle Approach, Intended use and analytical and clinical validation, Data quality, Privacy and data protection, and Engagement and collaboration. This overview is not intended as guidance, regulation or policy. Rather, it is a list of key regulatory considerations and is a resource that can be considered by all relevant stakeholders in medical devices ecosystems, including 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 in this area. The WG-RC recommends that stakeholders examine these key considerations and other potential ones as they continue to develop frameworks and best practices for the use of AI in health care and therapeutic development in relationship to the six topic areas.

The WG-RC recognizes that AI has been instrumental in rapidly advancing research in health care 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 considerations 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 are likely eventually to cross borders. Consistency of regulatory considerations for AI systems and technologies could improve standards and enable more rapid deployment. Third, 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 considerations in this deliverable may need to be expanded as the technology and its uses develop. The working group 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 AI topics for potential regulatory convergence and harmonization.

References

1 Global Strategy on Digital Health 2020-2025. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/handle/10665/344249, accessed 25 July 2023).

2 The 17 Goals – Sustainable Development (online). New York (NY): United Nations; 2020. (https://sdgs.un.org/goals, accessed 25 July 2023).

3 Thirteenth General Programme of Work 2019-2023. Geneva: World Health Organization (https://www.who.int/about/what-we-do/thirteenth-general-programme-of-work-2019---2023, accessed 25 July 2023).

4 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. Silver Spring (MD): US Food and Drug Administration; 2019 (https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf, accessed 25 July 2023).

5 Informal innovation network. Horizon scanning assessment report – Artificial Intelligence. International Coalition of Medicines Regulatory Authorities; 2021 (https://www.icmra.info/drupal/sites/default/files/2021-08/horizon_scanning_report_artificial_intelligence.pdf, accessed 25 July 2023).

6 ISO/IEC TR 24028:2020, Information technology – artificial intelligence – overview of trustworthiness in artificial intelligence (https://www.iso.org/standard/77608.html, accessed 25 July 2023).

7 Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449. Paris: Organisation for Economic Co-operation and Development; 2019 (https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449, accessed 25 July 2023.

8 Machine learning-enabled medical devices: a subset of AI-enabled medical devices: key terms and definitions. Proposed document posted for public consultation, 16 September 2021. International Medical Device Regulators Forum; 2021 (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, accessed 25 July 2023).

9 Panesar A. Machine learning and AI for healthcare. Big data for improved health outcomes. Coventry: Apress; 2019.

10 Artificial intelligence and intellectual property policy (online). Geneva: World Intellectual Property Organization; 2022 (https://www.wipo.int/about-ip/en/artificial_intelligence/policy.html, accessed 3 July 2023).

11 Wu E, Wu K, Daneshjou R, Ouyang D, Ho DE, Zou J. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. Nat Med. 2021;27(4):582–4.

12 Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. Nat Med. 2020;26(9):1364-74.

13 Rivera SC, Liu X, Chan A, Denniston AK, Calvert MJ. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. BMJ 2020;370:m3210.

14 Guidance for post-market surveillance and market surveillance of medical devices, including in vitro diagnostics. Geneva: World Health Organization; 2020 (https://apps.who.int/iris/handle/10665/337551, accessed 25 July 2023).

15 Software as a Medical Device (SaMD): key definitions. International Medical Device Regulators Forum; 2013. (http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf, accessed 25 July 2023).

16 Regulatory guidelines for software medical devices – a lifecycle approach (online). Singapore: Health Sciences Authority; 2022 (https://www.hsa.gov.sg/docs/default-source/hprg-mdb/guidance-documents-for-medical-devices/regulatory-guidelines-for-software-medical-devices---a-life-cycle-approach_r2-(2022-apr)-pub.pdf, accessed 25 July 2023

17 Oala L, Heiß C, Macdonald J, März M, Kutyniok G, Samek W. Detecting failure modes in image reconstructions with interval neural network uncertainty. Int J Comput Assist Radiol Surg. 2021;16(12):2089–97.

18 Oala L, Johner C, Goldschmidt P.G., Balachandran P. Good Practices for Health Applications of Machine Learning: Considerations for Manufacturers and Regulators. In: Proceedings of the ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H) – Meeting O; 2023 (https://www.itu.int/dms_pub/itu-t/opb/fg/T-FG-AI4H-2022-2-PDF-E.pdf , accessed 25 July 2023).

19 Principles and practices for medical device cybersecurity. International Medical Device Regulators Forum; 2019 (http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-ppmdc.pdf, accessed 25 July 2023).

20 Artificial intelligence/machine learning (AI/ML)-based software as a medical device (SaMD). Action plan. US Food and Drug Administration; 2021 (https://www.fda.gov/media/145022/download, accessed 25 July 2023).

21 Software as a medical device: possible framework for risk categorization and corresponding considerations. International Medical Device Regulators Forum; 2014 (https://www.imdrf.org/sites/default/files/docs/imdrf/final/technical/imdrf-tech-140918-samd-framework-risk-categorization-141013.pdf, accessed 25 July 2023).

22 A buyer's guide to AI in health and care. London: NHSX; 2020 (https://www.nhsx.nhs.uk/ai-lab/explore-all-resources/adopt-ai/a-buyers-guide-to-ai-in-health-and-care/, accessed 25 July 2023).

23 Notification No.0831-14, 31 August 2020 (Chinese). Handling with applications for confirmation of PACMP for medical devices, PSEHB/SD (in Japanese). Tokyo: Ministry of Health, Labour and Welfare; 2020 (https://www.mhlw.go.jp/content/11120000/000665757.pdf, accessed 25 July 2023).

24 Workshop on clinical evaluation of AI for health. Geneva: International Telecommunication Union; 2020 (https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/ws/2010.aspx, accessed 25 July 2023).

25 Schörverth E, 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.

26 Sendak M-P, Gao M, Brajer N, Balu S. Presenting machine learning model information to clinical end users with model facts labels. NPJ digital medicine. 2020;3(1):1–4.

27 Verks B, Oala L. Data and artificial intelligence assessment methods (DAISAM) Audit Reporting Template. In: Proceedings of the ITU/WHO Focus Group on Artificial Intelligence for Health (FG-AI4H) – Meeting J, 2020.

28 Oala L, Fehr J, Gilli L, Balachandran P, Leite AW, Calderon-Ramirez S et al. ML4H Auditing: from paper to practice. In: Proceedings of Machine Learning for Health (ML4H) NeurIPS Workshop. Proceedings of Machine Learning Research. 136:280-317. (https://proceedings.mlr.press/v136/oala20a.html, accessed 25 July 2023).

28 Willis K, Oala L. Post-hoc domain adaptation via guided data homogenization. (https://arxiv.org/abs/2104.03624, accessed 25 July 2023).

30 Calderon-Ramirez S, Oala L. More than meets the eye: semi-supervised learning under non-IID data. Presented as a RobustML workshop paper at International Conference on Learning Representations (ICLR), 2021 (https://arxiv.org/abs/2104.10223, accessed 25 July 2023).

31 Bellemo V, Lim ZW, Lim G, Nguyen QD, Xie Y, Yip MYT et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. Lancet Digit Health. 2019;1(1):e35–e44.

32 Macdonald J, März M, Oala L, Samek W. Interval neural networks as instability detectors for image reconstructions. In: Palm C, Deserno TM, Handels H, Maier A, Maier-Hein K, Tolxdorff T, editors. Bildverarbeitung für die Medizin. Informatik aktuell (Image processing for medicine. IT update). Wiesbaden: Springer Vieweg; 2021.

33 International Digital Health and AI Research Collaborative (I-DAIR) (online) (http://i-dair.org/, accessed 25 July 2023).

34 Salim M, Wåhlin E, Dembrower K, Azavedo E, Foukakis T, Liu Y et al. External evaluation of 3 commercial artificial intelligence algorithms for independent assessment of screening mammograms. JAMA Oncol. 2020;6(10):1581–8.

35 FG-AI4H Open Code Initiative (OCI). International Telecommunication Union; 2022 (https://www.itu.int/en/ITU-T/focusgroups/ai4h/Pages/opencode.aspx, accessed 16 March 2023).

36 AI audit.org (website) (https://aiaudit.org/, accessed 25 July 2023).

37 Software as a medical device (SaMD): clinical evaluation. International Medical Device Regulators Forum; 2016 (http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-samd-ce.pdf, accessed 25 July 2023).

38 Evidence standards framework for digital health technologies. London: National Institute for Health and Care Excellence (NICE); 2019 (https://www.nice.org.uk/Media/Default/About/what-we-do/our-programmes/evidence-standards-framework/digital-evidence-standards-framework.pdf, accessed 25 July 2023).

39 Topol EJ. Welcoming new guidelines for AI clinical research. Nat Med. 2020;26(9):1318-20.

40 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. Silver Spring (MD): US Food and Drug Administration; 2021 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory, accessed 25 July 2023).

41 Determining real-world data's fitness for use and the role of reliability. Durham (NC): Duke-Margolis Center for Health Policy; 2019 (https://healthpolicy.duke.edu/sites/default/files/2019-11/rwd_reliability.pdf, accessed 25 July 2023

42 Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366(6464):447–53.

43 Shana L. The geographic bias in medical AI tools. Ethics and Justice, Healthcare, Machine Learning. Stanford (CA): Stanford University Human-Centered Artificial Intelligence News and Announcements, 21 September 2020 (https://hai.stanford.edu/news/geographic-bias-medical-ai-tools, accessed 25 July 2023).

44 The dataset nutrition label (online). The Data Nutrition Project (https://datanutrition.org/, accessed 25 July 2023).

45 Gebru T, Morgenstern J, Vecchione B, Vaughan JW, Wallach H, Iii HD et al. Datasheets for datasets. Communications of the ACM. 2021;64(12):86–92.

46 Ethics and governance of artificial intelligence for health: WHO guidance. Geneva: World Health Organization; 2021. (https://apps.who.int/iris/handle/10665/341996, accessed 25 July 2023).

47 Greenleaf G. Global tables of data privacy laws and bills, seventh edition (February 11, 2021) 169 Privacy Laws & Business International Report; 2021:6–19. (https://ssrn.com/abstract=3836261 or http://dx.doi.org/10.2139/ssrn.3836261, accessed 25 July 2023).

48 Introduction to anonymisation: draft anonymisation, pseudonymisation, and privacy enhancing technologies guidance. London: Information Commissioner's Office (ICO); 2021 (https://ico.org.uk/media/about-the-ico/consultations/2619862/anonymisation-intro-and-first-chapter.pdf, accessed 25 July 2023).

49 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance).

50 Adequacy decisions: how the EU determines if a non-EU country has an adequate level of data protection. Brussels: European Commission (https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en, accessed 25 July 2023).

51 NIST privacy framework: a tool for improving privacy through enterprise risk management. Washington (DC): National Institute of Standards and Technology (NIST), US Department of Commerce; 2020 (https://www.nist.gov/system/files/documents/2020/01/16/NIST%20Privacy%20Framework_V1.0.pdf, accessed 25 July 2023).

52 India's Personal Data Protection Act. Chapter VI, 22(1)(e), 24(1).

53 West DM, Allen JR. Turning point: policymaking in the era of artificial intelligence. Washington (DC): Brookings Institution Press; 2020.

54 Framework for improving critical infrastructure cybersecurity. Washington (DC): National Institute of Standards and Technology (NIST), U.S. Department of Commerce; 2018 (https://www.nist.gov/cyberframework, accessed 25 July 2023).

55 Alsalamah SA, Alsalamah HA, Nouh T, Alsalamah SA. *HealthyBlockchain* for global patients. Computers, Materials & Continua. 2021;68(2):2431–49.

56 Attrey A, Lesher M, Lomax C. The role of sandboxes in promoting flexibility and innovation in the digital age. Going Digital Toolkit Note, No. 2. Paris: Organisation for Economic Co-operation and Development; 2020. (https://goingdigital.oecd.org/data/notes/No2_ToolkitNote_Sandboxes.pdf, accessed 25 July 2023

57 Madiega T, Van De Pol AL. Artificial intelligence act and regulatory sandboxes. European Parliamentary Research Service, June 2022., PE 733.544; (https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI(2022)733544_EN.pdf, accessed 25 July 2023).

58 First regulatory sandbox on artificial intelligence presented. European Parliamentary Research Service, June 2022. Brussels: European Commission (https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented, accessed 25 July 2023).

59 How should we engage and involve patients and the public in our work? London: Medicines and Healthcare products Regulatory Agency (MHRA); 2020 (https://www.gov.uk/government/consultations/how-should-we-engage-and-involve-patients-and-the-public-in-our-work, accessed 25 July 2023).

60 Stakeholder Engagement Framework. Canberra: Government of Australia, Department of Health and Aged Care; 2017 (https://www.health.gov.au/resources/publications/stakeholder-engagement-framework, accessed 25 July 2023).

61 Communication on enabling the digital transformation of health and care in the Digital Single Market; empowering citizens and building a healthier society. Brussels: European Commission; 2018 (https://digital-strategy.ec.europa.eu/en/library/communication-enabling-digital-transformation-health-and-care-digital-single-market-empowering, accessed 25 July 2023)

62 E-services. Singapore: Health Sciences Authority (HAS) (https://www.hsa.gov.sg/e-services, accessed 25 July 2023).

63 International Association for Public Participation (IAP2) Spectrum. International Association for Public Participation; 2007 (https://www.iap2.org/, accessed 20 March 2023).

64 Ho D. Artificial intelligence in cancer therapy. Science. 2020;367(6481):982–3. (https://science.sciencemag.org/content/367/6481/982, accessed 20 March 2023).

65 Ho D. Addressing COVID-19 drug development with artificial intelligence. Adv Intell Syst. 2020;2(5):2000070 (https://onlinelibrary.wiley.com/doi/full/10.1002/aisy.202000070, accessed 25 July 2023).

66 Blasiak A, Lim JJ, Seah SGK, Kee T, Remus A, Chye DH et al. IDentif.AI: Rapidly optimizing combination therapy design against severe acute respiratory syndrome Coronavirus 2 (SARS-Cov-2) with digital drug development. Bioeng Transl Med. 2020;6(1):e10196 (https://aiche.onlinelibrary.wiley.com/doi/10.1002/btm2.10196, accessed 25 July 2023).

67 Regulatory guidelines for software medical devices – a lifecycle approach. Singapore: Health Sciences Authority; 2019 (https://www.hsa.gov.sg/docs/default-source/announcements/regulatory-updates/regulatory-guidelines-for-software-medical-devices--a-lifecycle-approach.pdf, accessed 25 July 2023).

68 Ho D, Quake SR, McCabe ERB, Chng W J, Chow E K, Ding X et al. Enabling technologies for personalized and precision medicine. Trends Biotechnol. 2020;38(5):497–518. (https://www.cell.com/trends/biotechnology/fulltext/S0167-7799(19)30316-6, accessed 25 July 2023).

69 Shah P, Kendall F, Khozin S, Goosen R, Hu J, Laramie J et al. Artificial intelligence and machine learning in clinical development: a translational perspective. NPJ Digit Med. 2019;2:69. (www.nature.com/articles/s41746-019-0148-3, accessed 25 July 2023).

70 Harrer S, Shah P, Antony B, Hu J. Artificial intelligence for clinical trial design. Trends Pharmacol Sci. 2019;40(8):577–91 (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, accessed 25 July 2023).

71 Special access routes (medical devices). Import and supply of unregistered medical devices by request of qualified practitioners. Singapore: Health Sciences Authority; 2019 (https://www.hsa.gov.sg/medical-devices/registration/special-access-routes/qualified-practitioner-request, accessed 25 July 2023).

72 Complementary health products (CHP) classification tool. Singapore: Health Sciences Authority (https://www.hsa.gov.sg/CHP-classification-tool, accessed 25 July 2023).

73 Pagallo U, Casanovas P, Madelin R. 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. 2019;7(1):1–25.

74 Stakeholder communications. Sydney: Atlassian (https://www.atlassian.com/team-playbook/plays/stakeholder-communications-plan, accessed 25 July 2023).

75 Clarke R. Regulatory alternatives for AI. Computer Law & Security Review. 2019;35(4):398–409.

Annex

Terms and fundamental concepts

The FG-AI4H is proposing a new deliverable titled: "FG-AI4H terms and definitions" which aims to establish a new deliverable for the FG-AI4H with a glossary with agreed terminology in AI for health. The objectives of the new deliverable are the consistent use of terms across various deliverables, including WG-RC, and the promotion of harmonized use of important AI for health terms across the different disciplines involved in this cross-disciplinary field. However, this clause 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 the FG-AI4H terms and definitions deliverable.

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 *(1)*.

2 Trustworthiness

Trustworthy AI in the context of this document refers to AI systems and technologies that meet the 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 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, and 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 others.

4 Documentation

For the purpose of this document, the term "documentation" refers to processes and methods used to document, retain and pre-specify critical development ideas, including the initial conception, validation, deployment and post-deployment plans – as well as relevant key decisions, 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 the 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 pre-specified approaches and protocols are tracked, recorded and justified. Although effective documentation is only one element that supports transparency, it is a key regulatory principle.

5 Privacy

Privacy is a broad and multidimensional concept. It is a universally accepted fundamental human right.[14] 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 *(2)*. Privacy risks include reidentification and the release of unwanted inferences about a data subject (e.g., whether they have a certain disease *(3)*.

6 Data integrity

Data integrity can be defined as "the completeness, consistency, and accuracy of data"*(4)*.

7 Data protection

Data protection is a more technical issue under the broader umbrella of privacy which includes more domains beyond the protection of an individual'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 *(2)*.

8 Health data

Health data is personal data relating to a person's physical or mental health, and includes the provision of health-care services and information regarding a person's health status *(5)*. Health data are often considered to be a special category of personal data, or "sensitive" personal data, because of the nature and influence such data has on human lives and the impact on their fundamental rights and freedoms.

9 Sources of health data

Sources of health data include data acquired from digital health and medical technologies *(6)*, 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, smartphones and other electronic devices. 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, they will inform how the data are processed and for what purpose.

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"*(7)*.

[14] 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 honour and reputation."

11 AI system

The IMDRF *(1)* 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 that influence 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.

12 AI technology

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

References

1 Machine learning-enabled medical devices: a subset of AI-enabled medical devices: key terms and definitions. Proposed document posted for public consultation, 16 September 2021. International Medical Device Regulators Forum; 2021. (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, accessed 25 July 2023).

2 What is privacy? International Association of Privacy Professionals (IAPP); 2020 (https://iapp.org/about/what-is-privacy/, accessed 25 July 2023).

3 Kearns M, Roth A. The ethical algorithm: the science of socially aware algorithm design. New York (NY): Oxford University Press, 2019.

4 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. Silver Spring (MD): US Food and Drug Administration; 2021 (https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory, accessed 25 July 2023).

5 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance).

6 Vayena E, Dzenowagis J, Brownstein JS, Sheikh A. Policy implications of big data in the health sector. Bull World Health Organ. 2018;96(1):66–8. doi:10.2471/BLT.17.197426. (https://pubmed.ncbi.nlm.nih.gov/29403102/, accessed 25 July 2023)

7 Software as a Medical Device (SaMD): Key definitions. International Medical Device Regulators Forum; 2013. (http://www.imdrf.org/docs/imdrf/final/technical/imdrf-tech-131209-samd-key-definitions-140901.pdf, accessed 25 July 2023).