Questionnaire „Artificial Intelligence (AI) in medical devices“

(Version 5, 15.12.2023)

Preliminary remarks

- This questionnaire was compiled by the German Notified Bodies Alliance (Interessengemeinschaft der Benannten Stellen für Medizinprodukte in Deutschland - IG-NB) and is intended to serve as orientation for Notified Bodies, manufacturers and interested third parties.
- Created by Dr. Abtin Rad (TÜV SÜD), Dr. Andreas Purde (TÜV SÜD), Dr. Andreas Schwab (TÜV Rheinland), Volker Sudmann (mdc medical device certification), Markus Bianchi (DNV Medcert), Annika Chill (DNV Medcert), Martin Tettke (Berlin Cert), Michael Bothe (DQS Medizinprodukte), Mark Küller (TÜV-Verband / IG-NB).
- This questionnaire follows the idea that the safety of AI-based medical devices can only be achieved through a process-oriented approach, whereby all relevant processes and phases of the life cycle must be considered. Accordingly, the guideline does not place specific requirements on the products, but on the processes.
- This questionnaire is based in part on the „Guideline for AI for medical devices“ by Prof. Christian Johner, Christoph Molnar et al. (ai-guideline/Guideline-AI-Medical-Devices_EN.md at master · johner-institut/ai-guideline · GitHub).
- The document makes no claim to completeness or mandatory application.
- The focus of the assessment results from the intended use.
- The questions should be understood as a reference to best practice from the authors' point of view. The authors substantiate this with references to laws, standards or other guidelines.
- In the absence of standards specifically for medical devices, reference is also made in some cases to horizontal standards. However, these are only applicable to a limited extent.
- Questions regarding cybersecurity of medical devices can be found in IG-NB's questionnaires on cybersecurity for Medical Devices (https://www.ig-nb.de/veroeffentlichungen).

References

- REGULATION (EU) 2017/745 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC (2017/745/EU) - *in its current version*
- REGULATION (EU) 2017/746 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU (2017/746/EU) - *in its current version*
- 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 - GDPR) (2016/679/EU) - *in its current version*

- COMMISSION IMPLEMENTING REGULATION (EU) 2021/2226 of 14 December 2021 laying down rules for the application of Regulation (EU) 2017/745 of the European Parliament and of the Council as regards electronic instructions for use of medical devices (2021/2226/EU) - *in its current version*
- COUNCIL DIRECTIVE of 20 December 1979 on the approximation of the laws of the Member States relating to units of measurement and on the repeal of Directive 71/354/EEC (80/181/EEC) - *in its current version*
- MDCG 2019-16 Rev. 1 - Guidance on Cybersecurity for medical devices
- MDCG 2020-1 - Guidance on Clinical Evaluation (MDR) / Performance Evaluation (IVDR) of Medical Device Software
- MEDDEV 2.7/1 revision 4 (MEDDEV 2.7/1)
- EN ISO 13485:2016-08 + AC:2018 + A11:2021 Medical devices - Quality management systems - Requirements for regulatory purposes
- EN ISO 14971:2019 + A11:2021 Medical devices - Application of risk management to medical devices
- BS/AAMI 34971:2023, Application of ISO 14971 to machine learning in artificial intelligence
- ISO/IEC 25010:2011-03 Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models
- EN ISO 20417:2021 Medical devices – Information to be supplied by the manufacturer
- EN 62304:2006 + Cor:2008 + A1:2015 Medical device software - Software life-cycle processes
- EN 62366-1:2015 + COR1:2016 + A1:2020 Medical devices - Application of usability engineering to medical devices
- EN 82304-1:2017 Health Software – Part 1: General requirements for product safety
- EN ISO 14155:2020 Clinical investigation of medical devices for human subjects — Good clinical practice
- IEC 60601-1, 2006-10, A1:2013-10 - Medical electrical equipment – Part 1: General requirements for basic safety and essential performance
- IEC 80001-1:2021, Edition 2.0 - Application of risk management for IT-networks incorporating medical devices - Part 1: Safety, effectiveness and security in the implementation and use of connected medical devices or connected health software

Supplementary References

- ISO/IEC TR 24028:2020, Edition 1.0 - Information technology - Artificial intelligence - Overview of trustworthiness in artificial intelligence
- ISO/IEC TS 4213:2022, Information technology – Artificial intelligence – Assessment of machine learning classification performance
- ISO/IEC FDIS 23894 Information technology — Artificial intelligence — Guidance on risk management
- ISO/IEC 5338 Information technology — Artificial intelligence — AI system life cycle processes
- ISO/IEC FDIS 5339 Information Technology – Artificial Intelligence - Guidance for AI applications
- ISO/IEC DIS 5259-2 - Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 2: Data quality measures

- ISO/IEC DIS 5259-3 - Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 3: Data quality management requirements and guidelines
- ISO/IEC 5259-4 - Artificial intelligence — Data quality for analytics and machine learning (ML) — Part 4: Data quality process framework
- ISO/IEC 12207 - Systems and software engineering - Software life cycle processes

Terms and Definitions

- In this document, the term medical device is frequently used. Whenever the term medical device is mentioned, both types are meant, medical devices and in vitro diagnostic medical devices.

Changes to last version

- Revision/concretization of chapter A.1 (Certifiability of AI)
- Supplementation of missing references, addition of further references
- Linguistic and content adjustment of a variety of questions

Table of contents:

A) General requirements

1. Certifiability of AI

Medical devices using artificial intelligence/machine learning (AI/ML) are certifiable in principle under the assumption, that an appropriate conformity assessment procedure has been conducted.
For all medical software, including AI devices, regulatory requirements set limits for certification.
According to Regulation 2017/745/EU Annex II (4.) or 2017/746/EU Annex II (4.), manufacturers must comply with the general safety and performance requirements for their products.
Demonstration of conformity is challenging for AI-based products, especially products that use sophisticated stochastic machine learning techniques, rather than for products whose software contains explicit, deterministic algorithms. Therefore, this guideline is intended to provide AI-specific action guidance in demonstrating conformity.
Hence, the possibility of certification requires a review by the Notified Body and is a case-by-case decision.
As a general rule, a software device has to be validated prior to being placed on the market. The proof of a commensurate validation is part of the technical documentation and will be assessed by the Notified Body. In the case of "learning" software systems, the process of learning generally changes the performance of the device. Such changes, if they exceed a certain level, have to be considered as significant changes and require a new conformity assessment, which again has to be performed before placing the device on the market.
Practice has shown that it is difficult for manufacturers to sufficiently prove the conformity for AI devices, which update the underlying models using in-field self-learning mechanisms. Notified bodies do not consider medical devices based on these models to be "certifiable" unless the manufacturer takes measures to ensure the safe operation of the device within the scope of the validation described in the technical documentation.

2. Processes

The manufacturers should cover all aspects listed below either in the procedural instructions or in the relevant plans to ensure that the safety of the product is systematically guaranteed. Normally, the following standard operating procedures or plans are affected:

- Development
- Risk management
- Data management (customer property)
- Verification or validation (if not part of development)
- Post-market surveillance and vigilance
- Software life cycle, including service, installation and decommissioning
- Customer communication
- Purchasing, supplier and supply-chain control
- Management review (ISO 13485 requires consideration of “applicable new or revised regulatory requirements”.)

If the manufacturer outsources processes, the requirements apply accordingly. Examples would be a (software) development service provider or contract research organization to be required to consider the relevant chapters of this guideline.

3. Competences in development

1.	Does the manufacturer create a list of all roles that are directly or indirectly concerned with AI?	• 13485, 6.2 • 14971, 4.2, 4.3 • 62304, 5.1 • 82304
2.	Does the manufacturer identify AI-related skills for each role (e.g. developers, statisticians, modellers, etc.)?	• 13485, 6.2 • 14971, 4.3
3.	Does the manufacturer has adequate records of education, training and competences to conclude that the persons actually have these competences?	• 13485, 6.2 • 14971, 4.3
4.	Does the (software) development plans lay out the product-specific competences (beyond or deviating)?	• 13485, 7.3.2 • 82304, 6.1
5.	Is the integration of external competences done according to the rules on outsourced processes? How are outsourced competences recorded/documented?	• 13485, 4.1.5, 7.3.2 • 62304, 5.1.9

4. Documentation

1.	Does the manufacturer document compliance with the requirements for AI as part of the general safety and performance requirements?	• 2017/745/EU, Annex I, 17.2., 17.4., 23.4 • 2017/746/EU, Annex I 16.2., 16.4., 20.4.1 • 13485, 7.3.6, 7.3.7
2.	Does the manufacturer define appropriate retention periods for the specific data sets and information used for the AI model and implemented measures to protect against loss of those information?	• 2017/745/EU Art. 10 (8), Annex II 3. • 2017/746/EU, Art. 10 (7), Annex II 8. • 13485, 4.2.4, 4.2.5

B) Requirements for product development

1. Intended use and stakeholder requirements

a) Intended use

1.	Does the manufacturer determine for which medical purpose (diagnosis, therapy, monitoring, predictions) the device is to be used and for which parts of the intended use an AI is to be used?	• 13485, 4.2.3, 7.3.3 • 82304, 4.1 • 13485, 7.3.3 a)
2.	Does the manufacturer characterize the patients to be diagnosed, treated or monitored with the medical device? Does this characterization includes indications, contraindications and associated diseases?	• 2017/745/EU, Annex II 1.1. c) • 2017/746/EU, Annex II 1.1. c) • 62366, 5.1, 5.3 • 13485, 7.3.3 a)
3.	Does the manufacturer specify on which body locations the product will be used or from which body location the data originate?	• 2017/745/EU, Annex II, 1.1. • 2017/746/EU, Annex II, 1.1. • 62366, 5.1.3 • 13485, 7.3.3 a)
4.	Does the manufacturer identify measuring functions and defines sufficient accuracy, precision and stability for their intended purpose, based on appropriate scientific and technical methods? Does this information disclosed in the Instruction for Use accordingly? Are these measuring units expressed in legal units as per Council Directive 80/181/EEC?	• 2017/745/EU, Art 110, Annex I 14.6., 15., 23.4. (h), Annex II 6.2. (f) • 2017/746/EU, Art 103, Annex I 13.7., 14., Annex II 6.5 (c) • 13485, 7.3.3 a)
5.	Does the manufacturer specify the learning objectives for machine learning methods applied? Note: Examples are classification, regression and clustering.	• 2017/745/EU Annex II 1.1. • 2017/746/EU Annex II 1.1. • 62366, 5.1 • 13485, 7.3.3 a)
6.	Does the manufacturer implement an AI life cycle process throughout the entire product life span including post-market activities?	• 60601, 4.4 • 82304, 8. • 62304, 6.

b) Intended user, intended context of use

1.	Does the manufacturer characterize the intended users, e.g. - using demographic features (age, gender), - regarding the training and experience in medical domains, - regarding technical knowledge, - using physical and mental limitations, linguistic skills and cultural background?	• 2017/745/EU, Annex I 1., 5. (b), Annex II 1.1 • 2017/746/EU, Annex II 1.1., 5. • 62366, 5.1
2.	Does the manufacturer characterise the intended use environment (also with regard to the social environment, influenced by stress, shift work, frequently changing colleagues, etc.)?	• 62366, 5.1, second to last paragraph
3.	Is the device intended to be used by lay person? Does the manufacturer maintain evidence that the device performs appropriately according to its intended purpose when used by lay persons?	• 2017/745/EU Annex I 22. • 2017/746/EU, Annex I 19.
4.	Does the manufacturer assess the risk arising from unbalanced datasets based on the intended target patient population of the device?	• 2017/745/EU, Annex I 1.,3.,5. • 2017/746/EU, Annex I 1.,3.,5. • 14971, 5.3, 5.4, 5.5
5.	Does the manufacturer consider risks arising from Bias of the model?	• 2017/745/EU, Annex I 1., 3., 5. • 2017/746/EU, Annex I 1., 3., 5. • 14971, 5.3, 5.4, 5.5
6.	Does the manufacturer consider risks arising from differences in demographic, anthropometric, anatomical and physiological differences among target patient population when establishing datasets?	• 2017/745/EU, Annex I 1., 3., 5. • 2017/746/EU, Annex I 1., 3., 5. • 14971, 5.3
7.	Does the manufacture identify residual risks and define appropriate controls to judge the acceptance?	• 2017/745/EU, Annex I 4. • 2017/746/EU, Annex I 4. • 14971, 6, 7, 8
8.	How is the single fault condition assured for the device implementing or using AI? (AI module is either part of the medical device or a device on its own right)? Does the AI model output used for triggering alarm conditions, or used in any means to implement risk control measures?	• 2017/745/EU, Annex I 18.1. • 2017/746/EU, Annex I 17.1.

c) Stakeholder requirements

1.	Does the manufacturer identify the stakeholder requirements and translate them accordingly into the performance specifications?	• 13485, 7.3.3 • 62304, 5.1.1, 5.1.3
2.	Does the manufacturer define all markets and all relevant regulatory requirements there (e.g. CMDE Guideline for Assessments for China)?	• 13485, 7.2.1 c), 7.3.3 (b) • 62304, 5.2.2 (l)
3.	In the case of standalone software: Does the manufacturer determine the run-time environment of the product in terms of hardware (screen size, screen resolution, memory, network connection, etc.) and software (e.g. operating system, browser, run-time environments such as Java Run-time Environment or .NET including their version)?	• 2017/745/EU, Annex I 17.3., 17.4. • 2017/746/EU, Annex I 16.3., 16.4. • 62304, 5.2.2

d) Input for risk management and clinical evaluation

1.	Does the manufacturer list alternative methods to AI and evaluated them with regard to benefit, safety and performance?	• 2017/745/EU, Annex I 1. • 2017/746/EU, Annex I 1. • MEDDEV 2.7/1, 10.2
2.	Does the manufacturer demonstrate that any additional risks related to the AI methods are outweighed by the clinical benefit, when compared to conventional methods with same intended use?	• 2017/745/EU, Annex I 1. • 2017/746/EU, Annex I 1. • MEDDEV 2.7/1, 10.2
3.	Does the manufacturer drawn up a list of risks specifically arising from the use of AI techniques?	• 14971, 5.3, 5.4
4.	Does the manufacturer analyse the risks that arise when users other than the specified users use the product?	• 14971, 5 • 62304, 5.1 -5.6
5.	Does the manufacturer analyse the risks posed by inputs that do not meet the specified formats and/or have not been generated according to the specified prerequisites?	• 14971, 5.2
6.	Does the manufacturer analyse the risks that arise if the outputs do not meet the specified quality criteria?	• 14971, 5.6.8 • 82304, 4.1
7.	Does the manufacturer assess the risks if the system is used in a different patient population than specified?	• 14971, 5.4, 5.5 • 82304, 4.1, 4.8 • 62304, 4.1, 6.2
8.	Does the manufacturer analyse the risks arising from use in an environment other/different than the specified use environment?	• 14971, 5 • 62366, 5.1 -5.6

	Does the manufacturer ensure that the data collection is performed under pre-defined conditions?	
9.	Does the manufacturer derive the quantitative quality criteria based on the state of the art? Does the manufacturer define operational limits (e.g. dose limits) within which the AI system may operate? Does the manufacturer define how to ensure that these operational limits are not exceeded?	• 2017/745/EU, Annex I 1., 17.2. • 2017/746/EU, Annex I 1, 1., 16.2. • 13485, 4.1 • MEDDEV 2.7/1
10.	Does the manufacturer assess the risks arising from a training dataset, that does not match the actual patient population?	• 14971, 5
11.	Does the manufacturer assess the risks if the system is not available?	• 2017/745/EU, Annex I 14.1., 17.1. • 2017/746/EU, Annex I 13.1., 16.1. • 62304, 5.2.3
12.	Does the manufacturer assess the risks through the specific choice of target platform?	• 2017/745/EU, Annex I 14.1., 17.4. • 2017/746/EU, Annex I 13.1., 16.4. • 14971 • 82304, 3.2, 7.2
13.	Does the manufacturer assess the risks related to splitting the data into training, validation and test data?	• 2017/745/EU, Annex II. 6.1 b • 14971, 7.1 • 13485, 3.7, 7
14.	Does the manufacturer evaluate to what extent the results achieved are based on causal relationships (explainable AI)?	• 13485, 7.3.7, 7.3.6 • 62304, 5.7.4, 5.8.1 • 82304, 4.6, 6.1
15.	Does the manufacturer assess the risks by making predictions that themselves change the predicted outcomes if possible? Note: This phenomenon applies, when the model switches from being an observer to an actor. It is refereed to as "performative prediction". Manufacturers should investigate the possible effects on people or systems and describe them, e.g. with a DAC ("directed acyclic graph"), observe a possible distribution shift and, if necessary, an unaffected	• 82304, 8.4, 8.3 • 13485, 7.3.9 • 62304, 5.7.3

	control group, and take action if necessary, such as choosing a different model or re-training the existing model.	
16.	Does the manufacturer assess the risks from use errors? Note: These risks should also take into account that users do not recognise or misunderstand the explanation of the outputs ("explain ability").	• 62366, 5.3 f)

2. Software requirements

a) Functionality and performance

1.	Does the manufacturer derive quantitative quality criteria or requirements for the software or/and the algorithm from the intended use in a comprehensible way, which corresponds to the state of the art including available benchmarks?	• 2017/745/EU, Annex I 17.2. • 2017/746/EU, Annex I 16.2. • 13485, 7.3.3 • 62304, 5.2
2.	Does the manufacturer consider the quantitative quality criteria or requirements such as: - for classification problems: accuracy (mean or balanced accuracy), positive predictive value (precision) or specificity and sensitivity? - for regression problems: mean absolute error and mean square error?	• 2017/745/EU, Annex I 15.1., 15.2. • 2017/746/EU, Annex I 14.1., 14.2. • 13485, 7.3.3, 7.3.4 • 62304, 5.2
3.	Are these qualitative criteria sufficient and inline with the clinical outcome / performance evaluation parameters and scientifically justified?	• 2017/746/EU, Annex XIV Part A 1. a) • 2017/746/EU, Annex XIII Part A 1.
4.	Does the manufacturer specify the expected value ranges of the outputs? How is ensured that the output remains within validated range?	• 13485, 7.3.3, 7.3.4 • 62304, 5.2 • 82304, 4.5
5.	Does the manufacturer specify the requirements regarding repeatability and reproducibility of requirements?	• 2017/745/EU, Annex I 17.1. • 2017/746/EU, Annex I 16.1. • 13485, 7.3.3, 7.3.4

6.	Does the manufacturer specify how the system will behave if the inputs do not meet the specified requirements?	• 25010 • 62304, 5.2 • 80001 • 82304, 4.5, • 13485, 7.2.3
7.	What requirements must be met in order to be able to detect misconduct, e.g. by means of self-tests? If the manufacturer uses self-tests: Does he explain which of the specified quality criteria are checked with it and which risks are thereby controlled? Is it specified how the system behaves in the event of negative results?	• 13485, 7.3.3 • 62304, 5.2 • 82304, 4.5 • 14971, 5.3 • 80001
8.	Does the manufacturer specify how fast the system must generate the outputs?	• 2017/745/EU, Annex I 17.1. • 2017/746/EU, Annex I 16.1. • 13485, 7.3.3 • 62304, 5.2 • 82304, 4.5
9.	Does the manufacturer specify requirements for the availability of the medical device?	• 25010 • 62304, 5.2 • 14971, 4.3 • 13485, 7.3.3 • 80001 • 82304, 4.5
10.	Does the manufacturer define interoperability, and combination with other device, related characteristics, interfaces etc.?	• 2017/745/EU, Annex I 14.1., 17.3., 17.4., Annex II 6.2. g) • 2017/746/EU, Annex I 13.1., 16.3., Annex II 6.5. g) • 62304, 5.2 • 82304, 4.5

b) User interface

1.	Does the manufacturer specify what the user interface must display if the requirements are not met in order to operate the system safely (e.g. inputs not valid or not expected)?	• 2017/745/EU, Annex I, 6. • 2017/746/EU, Annex I 6. • 62366, 5.2
2.	Does the manufacturer determine whether a quality of output needs to be provided to the user?	• 2017/745/EU, Annex I 5.

	If so, how is the quality indicated to the user?	• 2017/746/EU, Annex I 5. • 62366, 5.2, 5.3
3.	Does the manufacturer determine whether an instruction for use and training materials are required?	• 2017/745/EU, Annex I 23. • 2017/746/EU, Annex I 20. • 13485, 4.2.3 • 62366, 5.8

c) Additional software requirements

1.	Does the manufacturer specify the data interfaces, including the formats and, in the case of images, their specific properties (size, resolution, colour coding)?	• 62304, 5.2.2 • 82304, 4.2
2.	Does the manufacturer specify the input data requirements?	• 2017/745/EU, Annex I, 5. • 2017/746/EU, Annex I, 5. • 62366, 5.2 • 13485, 7.3.3, 7.3.4 • 62304, 5.2 • 82304, 4.2
3.	Does the manufacturer specify which requirements the system must fulfil in order to be able to detect a failure of the system?	• 2017/745/EU, Annex I 17., 18., 23.4. • 2017/746/EU, Annex I 16., 17., 20.4. • 62304, 5.2, 5.3, 7.1 • 14971, 5.4 • 82304, 4.2
4.	Does the manufacturer specify the requirements for IT security? (Use guidelines on cybersecurity for this: https://owasp.org/www-project-ai-security-and-privacy-guide/, https://www.ig-nb.de/veroeffentlichungen) Has the manufacturer considered AI-specific cybersecurity risks?	• 2017/745/EU, Annex I 17.4., 23.4. aa) • 2017/745/EU, Annex I 16.4., 20.4.1. ah) • MDCG 2019-16

d) Security risks of artificial intelligence

Note: In addition to already known cybersecurity risks for software-assisted medical devices and software medical devices (see IG-NB's Questionnaires on Cybersecurity for Medical Devices), there are also AI-specific attacks. These are fundamentally different from conventional cyberattacks, which are mostly due to „bugs“ or human errors in the code. Cyberattacks against AI are usually directed against inherent vulnerabilities in the underlying algorithms, which cannot be fixed or can only be fixed with difficulty. So-called adversarial attacks aim to manipulate the decision/classification of the AI.

1.	Does the manufacturer identify the cybersecurity risks applicable to the AI, such as poisoning attacks, evasion attacks or model extraction etc.?	• 2017/745/EU, Annex I 3. b) • 2017/746/EU, Annex I 3. b) • 13485, 7.1
2.	Does the manufacturer search and document sources (such as Adversarial ML Threat Matrix, MAUDE database and others) for identifying threats against AI models?	• 2017/745/EU, Annex I 3. e) • 2017/746/EU, Annex I 3. b) • 13485, 8.2, 8.4 • 14971, 5.3, 5.4, 7.2, 10 • MDCG 2019-16 • 62304, 9.6 • 82304, 8
3.	Does the manufacturer consider and assess the identified security risks in its risk management?	• 2017/745/EU, Annex I 3. c), 17.2, 17.3 • 2017/746/EU, Annex I 3. c) • 13485, 7.3.3 c. • 14971, 5.3
4.	Does the manufacturer define risk mitigation measures for the identified risks?	• 2017/745/EU, Annex I 3. c. • 2017/746/EU, Annex I 3. c) • 14971, 7.2
5.	Does the AI lifecycle take into account an appropriate security lifecycle?	• 2017/745/EU, Annex I 17.2. • 2017/746/EU, Annex I 16.2. • 62304, 4.3, 5.1.1 e) • MDCG 2019-16

6.	Have measures been implemented and taken into account hardening the algorithms against adversarial attacks?	• 2017/745/EU, Annex I 1., 4. • 2017/746/EU, Annex I 1., 4. • 14971, 10.2
7.	Does the manufacturer consider the entire infrastructure and supply chain within the security risk assessment regarding the AI model and its deployment strategy? (For example back-end software running in the cloud or utilizing cloud based services for certain functions during the development and product life-cycle.)	• 2017/745/EU, Annex I 1., 4. • 2017/746/EU, Annex I 1., 4. • 62304, 5.1.1, 4.3
8.	Has the manufacturer identified and evaluated the gaps between 23894 and 14971 in his risk management documentation considering the device under assessment?	• 14971 Supplementary references: • 23894

3. Data management

Data can generally be divided into training, validation and test data, which can be subject to different requirements. Insofar as not further specified in this chapter, the term ‘data‘ includes all three types.

a) Collection of the training, validation and test data sets

1.	Does the manufacturer specify the number of records and given a justification as to why this is sufficient?	• 13485, 7.3.7 Supplementary references: • 4213, 5
2.	Does the manufacturer characterise the inclusion and exclusion criteria of data using relevant attributes?	• 34971, 5.3.2, 5.3.3 Supplementary references: • 4213, 5
3.	Does the manufacturer specify technical inclusion and exclusion criteria for data?	• 34971, 5.3.2, 5.3.3 Supplementary references: • 4213, 5
4.	Does the manufacturer describe the procedure to ensure that records that do not meet the inclusion criteria or are to be excluded are in fact excluded?	• 34971, 5.3.2, 5.3.3

		Supplementary references: • 4213, 5
5.	Does the manufacturer describe the collected data using descriptive statistics?	• 34971, 5.3.2, 5.3.3 Supplementary references: • 4213, 6
6.	Does the manufacturer justify where it collects e.g. training, test and validation data and why it is representative of the target population? Where appropriate, has the manufacturer compared these with data from the Federal Statistical Office, scientific publications and registries?	• 34971, 5.3.2, 5.3.3 Supplementary references: • 4213, 5
7.	Does the manufacturer justify that the size of the dataset(s) sufficiently represent(s) the target patient groups / sub-groups (minimizes the risk of unbalanced dataset), for example if there are race, regional and other differences that might adversely impact the clinical safety and performance of the device?	• 34971, 5.3.2, 5.3.3 Supplementary references: • 4213, 5
8.	Does the manufacturer list and discuss factors that could cause a bias of the validation and test data?	• 34971, 5.3.3 Supplementary references: • 4213, 5.3
9.	Does the manufacturer analyse what influences the type and location of data collection has on the data?	• 34971, 5.3.2, 5.3.3 Supplementary references: • 4213, 5.3
10.	Does the manufacturer establish a procedure to anonymise or pseudonymise data before training and testing?	• 34971, 5.3.4
11.	Does the manufacturer investigate and rule out the possibility of label leakage?	• / Supplementary references: • 4213, 5.3
12.	For systems that process patient data: How is it ensured that patient data is adequately protected? (Including prevention from model extraction)	• 34971, 5.3.4 • GDPR
13.	Does the manufacturer establish procedures for sufficient handling of customer property, including patient data, health records or other data	• 13485, 7.5.10 • 34971, 5.3.4

	that is used for the AI model development and otherwise can be considered as customer property?	
14.	Does the manufacturer establish procedures for data storage and retention according to the applicable regulatory requirements?	• 13485, 4.2.4, 4.2.5
15.	Does the manufacturer establish procedures to collect patient data based on well know practices such as Good Clinical Practice (GCP)? (e.g. consent of participants are collected and handled accordingly, etc.)	• 14155

b) Labelling of data

1.	Does the manufacturer derive the labels from the intended use for which the training data is understood and justify this choice?	• /
2.	Does the manufacturer specify a procedure for labelling, if no labels were yet present in the data?	• /
3.	Does this procedure specify quantitative/qualitative classification criteria for labelling? Has the manufacturer justified the choice of these criteria?	• /
4.	Does this procedure specify the requirements for the number, training and competence of the persons responsible for labelling?	• /
5.	Does this procedure specify how the competence of the persons responsible for labelling is checked?	• 34971, 4.3
6.	Does this procedure specify how the persons responsible for labelling are trained and how the success of this training is evaluated?	• 34971, 4.3
7.	Does this procedure specify how the correctness of the labels is systematically reviewed? Has the manufacturer documented the choice of this rationale?	• 34971, 5.3.2
8.	Does this procedure specify how it is monitored that the persons responsible for labelling are also permanently capable and willing to perform during labelling?	• /

c) Procedure for (pre-)processing of data

1.	Has the manufacturer set a procedure describing the (pre-)processing of the data?	• / Supplementary references: • 5259-3, 7.1., 7.2.2.5, 7.3, 7.3.2 f), 7.3.3 e)
2.	Does this procedure describe the individual processing steps such as conversions, transformations, aggregations, normalisation, format conversions, calculation of features and conversion of numerical data into categories (augmentation)?	• 13485, 7.3.3 e) Supplementary references: • 5259-3, 7.2.2.6

3.	Does the procedure describe how the correctness of the intermediate steps and the final results is checked? Are these checks carried out on a risk basis?	• 13485, 7.3.2, 7.3.3 c), 7.3.4 a), 7.3.4 c), 7.3.6, 7.5.6 Supplementary references: • 5338, 6.4.3.3 • 23894, 6.4.2.6
4.	Does this procedure specify how values with different measurement scales or units are recognised and processed (normalisation of data)?	• 13485, 7.3.3 e) Supplementary references: • 5259-4, 12.5.1
5.	Does this procedure specify how values determined with different measurement methods are detected and processed?	• 13485, 7.3.3 e) Supplementary references: • 5259-3, 7.3.4.3f, 10.2c
6.	Does this procedure specify how values or metadata with the same names (such as in column headers) are detected and processed?	• 13485, 7.3.3 e) Supplementary references: • 5259-4, 7.5.9.3.3
7.	Does this procedure specify how missing values, outliers and unusable data within data sets are detected and processed? Has the manufacturer justified this specification?	• 13485, 7.3.7 Supplementary references: • 5259-4, 12.5.2
8	Does the manufacturer establish procedures to control changes to the pre-processing steps/algorithms?	• 13485, 7.3.9 • 62304, 8.2 Supplementary references: • 5259-3, 8.3.3

d) Documentation and version control

1.	Does the manufacturer document all points from sections 3. a (collection of the training, validation and test data sets), 3. b (labelling of data) and 3. c (procedure for (pre-)processing of data) in a comprehensible way?	• 13485, 4.2.3 • 62304, 5.1.8, 5.8.6 Supplementary references:

		• 5338, 6.4.8.2 b, c 6.4.8.3 b • 5339, 10.2
2.	Has the manufacturer all software for data processing, including the libraries used in the process, documented and under version control?	• 13485, 4.1.6, 4.2.4, 7.5.6
3.	Does the manufacturer has the training, validation and test data set under version control?	• 13485, 4.2.5
4.	Does the manufacturer describe all data sources (e.g. clinics, devices)?	• 13485, 7.3.3, 7.4.2 Supplementary references: • 5259-4, 12.2
5.	Is the dataset used scientifically justified for appropriateness based on the intended purpose, e.g. using descriptive statistics?	• 13485, 7.3.7 Supplementary references: • 5339, 7.4.5 c, 7.3.2.3
6.	Does the manufacturer establish procedures for data storage and retention according to the applicable regulatory requirements?	• 13485, 4.2.4, 4.2.5 Supplementary references: • 5339, 7.3.8.3

4. Model development

a) Preparation

1.	Does the manufacturer justify the selection of the parameters considered during training?	• 13485, 7.3.2, 7.3.3
2.	Has the manufacturer established procedure(s) to identify parameters, AI model architecture and ensures that changes to theses are handled accordingly pre/post-market deployment?	• 12207, 6.4.4 • 62304, 5.3, 8.2.1, 9.4 Supplementary references: • 5338, 6.4.4
3.	Does the manufacturer describe the interdependence of the parameters, especially in the case of tabular data?	• 13485, 7.3.2, 7.3.3

		Supplementary references: • 5259-4, 7.5.9.3.3
4.	Does the manufacturer document and justify the ratio in which it divides the data into training, validation and test data?	• 13485, 7.3.2 c), 7.3.7 Supplementary references: • 5259-3, 7.3.2.3
5.	Does the manufacturer document the stratification used to divide the data into training, validation and test data?	• 13485, 7.3.2 c), 7.3.7 Supplementary references: • 5259-3, 7.3.2.3
6.	How does the manufacturer ensure the consistency of training data (e.g. dispersion of a specific parameter due to accidental transfer to the general public)?	• / Supplementary references: • 5259-2, 6.2.3.2
7.	Does the manufacturer document how it ensures that test data has not been used in both training and validating the model?	• 13485, 7.3.2 e) and f) Supplementary references: • 5338, 6.4.8.2 b, c
8.	If the manufacturer recodes the data specifically for the model or specifically for the library: Does he describe the procedure?	• 13485, 7.1, 7.3 • 62304, 5.1 Supplementary references: • 5338, 6.4.11

b) Training

1.	Does the manufacturer determine, document and justify the quality metrics based on the intended use for which he wants to optimise the model?	• 13485, 7.3.3 a), 7.3.7 Supplementary references: • 5259-3, 7.3.7.3 h

2.	Has the manufacturer trained and compared several model types (including simpler and interpretable models), where appropriate?	• 62304, 5.3.1, 5.4.2 Supplementary references: • 5338, 6.4.9.3

c) Evaluation

1.	Does the manufacturer document the quality metrics for the different models, e.g. for a binary classification, with the help of a confusion table?	• / Supplementary references: • 4213, 5, 6, 7
2.	Does the manufacturer assess and document the quality metrics for the different models not only globally, but also separately for different features, if applicable?	• / Supplementary references: • 4213, 5, 6, 7
3	Does the manufacturer not only globally assess and document the quality measures for the different models, but also separately for different features, if applicable?	• / Supplementary references: • 4213, 5, 6, 7
4	Does the manufacturer identify means to reduce the risk of training procedure related effects such as overfitting?	• 24028, 9.8.2.23
5.	Does the manufacturer examine the data sets that predicted particularly well and those that predicted particularly poorly?	• /
6.	How have the boundaries of safe operation been determined? Does the manufacturer examine the data sets for which the model decision is particularly safe or particularly unsafe? (Has the effect of operation outside the specified acceptance range been assessed? E.g. via worst case scenario analysis, out of boundary analysis)	• /
7.	Does the manufacturer justify the final choice of model on the basis of the quality criteria and the intended use, and in particular explain when simpler and more interpretable models were not used?	• /
8.	For tabular data in particular, does the manufacturer consider displaying, for individual data sets, the features that particularly drove the model to make the decision (Explainable AI)?	• /
9.	For tabular data in particular, does the manufacturer consider evaluating how and to what extent individual features would have to change for the model to arrive at a different prediction?	• /

10.	For tabular data in particular, does the manufacturer consider analysing / visualising the dependence (strength, direction) of the predictions on the feature values?	• /
11.	Does the manufacturer consider synthesising data sets that particularly activate the model?	• /

d) Documentation

1.	Does the manufacturer has the model and/or training code under version and configuration control? Does the manufacturer keep records to demonstrate which data sets are used for training, validating and testing the model?	• 13485, 4.1.6, 4.2.4, 7.5.6
2.	Can the manufacturer reproduce the test and validation results?	• 13485, 7.3.6, 7.3.3
3.	Does the manufacturer has the SOUP (libraries and frameworks) under version and configuration control?	• 62304, 8.1.2
4.	Does the manufacturer document the architecture of the model and the model itself including its hyper parameters?	• 13485, 4.2.3, 4.2.5
5.	Does the manufacturer describe when it has worked with a "pretrained model" and shown why this "pretraining" is suitable for the task?	• /
6.	Does the manufacturer document the quality of the models based on the quality metrics?	• 13485, 4.2.3, 4.2.5
7.	In particular, for tabular data: Does the manufacturer document within which limits (e.g. feature values) the model achieves the requirements for the quality metrics?	• 13485, 4.2.3, 4.2.5
8.	Does the manufacturer assess the validity of confidence intervals for quality parameters depending on input data?	• /
9.	Does the manufacturer evaluate the performance of different models and multiple sets of parameters to control the learning process? Note: This is deemed necessary to meet the requirement of 14971 to maximise the benefit-risk ratio.	• 14971, 7.1

5. Product Development

a) Software development

1.	Has the manufacturer carried out and documented all required activities?	• 62304 • 82304
2.	If the manufacturer has implemented the model in another language or for another runtime environment: Has he made a plan which of the activities he has to repeat?	• 62304 • 82304

3.	Does the manufacturer check the performance (response times, resource consumption) on the target hardware (e.g. browser, mobile device)?	• 2017/745/EU, Annex I 17.1., 17.3. • 2017/746/EU, Annex I 16.1., 16.3.
4.	Does the manufacturer describe how to verify all SOUP or OTS components?	• 62304

b) Accompanying materials

1.	Do the instructions for use identify the version of the product with sufficient precision?	• 2017/745/EU, Annex I 23.1., 23.4. • 2017/746/EU, Annex I 20.4.1., a)
2.	Do the instructions for use describe how the product is to be used?	• 2017/745/EU, Annex I 23.4. • 2017/746/EU, Annex I 20.4.1.
3.	Do the instructions for use describe the intended use of the product including the expected medical benefit?	• 2017/745/EU, Annex I 23.4. b) • 2017/746/EU, Annex I 20.4.1. c)
4.	Do the instructions for use identify the intended patient population on the basis of indications, contraindications and - where relevant - other parameters such as age, gender, concomitant diseases or availability of information?	• 2017/745/EU, Annex I 23.4. b) • 2017/746/EU, Annex I 20.4.1. c)
5.	Does the manufacturer identify measuring functions and defines sufficient accuracy, precision and stability for their intended purpose, based on appropriate scientific and technical methods? Is this information disclosed in the Instruction for Use accordingly? Are these measuring units expressed in legal units as per Council Directive 80/181/EEC?	• 2017/745/EU, Annex I 14.6., 15., 23.4. (h), Annex II 6.2. (f) • 2017/746/EU, Annex I 14., Annex II 6.5. (c)
6.	Do the instructions for use explicitly state the patients / data / use cases for which the product may not be used?	• 2017/745/EU, Annex I 23.4. b) • 2017/746/EU, Annex I 20.4.1. c)
7.	Do the instructions for use document the requirements for the input data (including formats, resolutions, range of values, etc.)?	• 20417, 6.6.2 c)
8.	Do the instructions for use specify the intended user groups according to the intended use?	• 2017/745/EU, Annex I 23.4. b)

		• 2017/746/EU, Annex I 20.4.1. e)
9.	If the device is intended to be used by lay person: Does the manufacturer maintain evidence that the device performs appropriately according to its intended purpose when used by lay persons?	• 2017/745/EU, Annex I 22. • 2017/746/EU, Annex I 19.
10.	Do the instructions for use describe what other prerequisites the product assumes (e.g. runtime environment, usage environment)?	• 2017/745/EU, Annex I 23.4. f) • 2017/746/EU, Annex I 20.4.1. j)
11.	Do the instructions for use describe the residual risks?	• 2017/745/EU, Annex I, 23.4. g) • 2017/746/EU, Annex I, 20.1. g) • 14971, 8
12.	If useful: Do the instructions for use specify the data with which the model was trained? Note: For example depending on the use context.	• 2017/745/EU, Annex I, 23.4. h)
13.	If useful: Do the instructions for use describe the model or the algorithms? Note: For example depending on the use context.	• 2017/745/EU, Annex I, 23.4. h
14.	If useful: Do the instructions for use specify the quality criteria? Note: For example depending on the use context.	• 2017/745/EU, Annex I 23.4. e) • 2017/746/EU, Annex I 20.4.1. w)
15.	Do the instructions for use list the factors that can have a negative impact on the quality criteria?	• 2017/745/EU, Annex I 23.4. s) • 2017/746/EU, Annex I 20.4.1. n)
16.	Do the instructions for use describe how updates are made?	• 2017/745/EU, Annex I 23.4. a) • 2017/746/EU, Annex I 20.4.1. ad)
17.	Do the instructions for use identify the manufacturer and the channels through which inquiries can be made?	• /
18.	Do the instructions for use name the URL where the latest versions of the instructions for use can be found?	• 2021/2226/EU

c) Usability validation

1.	As part of the usability validation: Does the manufacturer assess whether the users understand the instructions for use?	• 62304
2.	As part of the usability validation: Does the manufacturer assess whether users blindly trust the product or check the results?	• 62366
3.	As part of the usability validation: Does the manufacturer assess whether the users correctly recognise and understand the results?	• 62366, 5.7, 5.8, 5.9

d) Clinical evaluation

1.	Does the manufacturer demonstrate in the clinical evaluation that the expected clinical benefit is achieved, considering the given quality parameters and predefined acceptance criteria?	• 2017/745/EU, Annex XIV, Annex XV • 2017/746/EU, Annex XIII, Annex XIV • MEDDEV 2.7/1
2.	As part of the clinical evaluation: Does the manufacturer demonstrate that the expected clinical benefit corresponds to the state of the art?	• 2017/745/EU, Annex XIV, Annex XV • 2017/746/EU, Annex XIII, Annex XIV • MEDDEV 2.7/1
3.	For devices intended for diagnosis or aid to diagnosis of patients, has the technical/analytical performance been verified and validated in the intended computing- and use environments, using appropriate outcome parameters?	• MDCG 2020-1, 4.3
4.	Does the manufacturer clearly identify and justify whether a clinical investigation and/or PMCF study is required or not?	• 2017/745/EU, Annex XIV, Annex XV • 2017/746/EU, Annex XIII, Annex XIV
5.	Does the manufacturer establish procedures to perform clinical investigations and/or PMCF studies to obtain sufficient evidence for the AI model clinical safety and performance, when required according to the applicable regulatory requirements?	• 2017/745/EU, Annex XV • 2017/746/EU, Annex XIV
6.	Does the manufacturer clearly define the clinical outcome parameters regarding the intended purpose of the device and outputs of the AI model?	• 2017/745/EU, Annex XIV • 2017/746/EU, Annex XIII

7.	Does the manufacturer establish procedures to collect patient data based on well know practices such as Good Clinical Practice (GCP)? (e.g. consent of participants are collected and handled accordingly, etc.)	• 2017/745/EU, Annex XIV • 2017/746/EU, Annex XIII • 14155
8.	Does the manufacturer scientifically and sufficiently justify the transferability of clinical data among the different patient groups, including differences arising from region or ethnicity, physiological differences, etc.?	• 2017/745/EU, Annex XIV • 2017/746/EU, Annex XIII • 14155

6. Product release

(essential points, not an exhaustive list)

1.	Does the manufacturer document the models and data used against the above criteria?	• 2017/745/EU, Annex II, 1.1. j) • 13485, 7.3.5
2.	In risk management: Does the manufacturer assess the risks as acceptable and document that all of the activities specified in the risk management plan were performed?	• 14971, 8, 9
3.	Are residual risks communicated to customers?	• 2017/745/EU, Annex I, 23.4. g) • 2017/746/EU, Annex I, 20.1. g)
4.	Does the manufacturer prepare a post-market surveillance plan?	• 2017/745/EU, Art. 84 • 2017/746/EU, Art. 79

C) Requirements for the post development phases

1. Production, distribution, installation

1.	Does the manufacturer describe how it is ensured that only exactly the intended artefacts (files) are delivered in exactly the intended version in the product or as a product?	• 62304, 5.8.8
2.	Does the manufacturer describe how the people responsible for the installation will know which is the latest version and how mix-ups during installation can be avoided?	• 13485, 7.8.3, 8.3 • 62304, 5.8.4
3.	Does the manufacturer describe how it will be ensured during installation that the requirements specified in the accompanying materials (see above) are actually fulfilled?	• 13485, 7.5.3
4.	Has the manufacturer established procedures to ensure that it can communicate with the operators and users of its products in a timely manner?	• 13485, 7.2.3, 8.3.3 • 82304, 8.4
5.	Does the manufacturer specify and communicate minimum requirements regarding hardware, IT network characteristics and IT security measures, including protection against unauthorised access? Note: This can be part of the IFU documentation for some devices.	• 2017/745/EU, Annex I, 23.4. ab) • 2017/746/EU, Annex 1, 20.4.1. ah) • 80001

2. Post-Market Surveillance

1.	Has the manufacturer prepared a Post-Market Surveillance (PMS) Plan?	• 2017/745/EU, Art. 84 • 2017/746/EU, Art. 79
2.	Does the manufacturer specify in this PMS plan the data he intends to collect and evaluate?	• 2017/745/EU, Annex III 1.1. • 2017/746/EU, Annex III 1.1.
3.	Does the manufacturer specify in the PMS plan at which quality criteria and thresholds it considers action necessary, in particular a reassessment of the risk-benefit balance?	• 2017/745/EU, Annex III 1.1. • 2017/746/EU, Annex III 1.1.
4.	Does the manufacturer describe in the PMS plan how it collects and analyses what information on adverse medical effects?	• 2017/745/EU Annex III 1.1. 2017/746/EU Annex III 1.1.

5.	Does the manufacturer describe in the PMS plan how and which information on (adverse) behavioural changes or (predictable) misuse is collected and how these information are assessed?	• 2017/745/EU Annex III 1.1. • 2017/746/EU, Annex III 1.1.
6.	Does the manufacturer describe in the PMS plan how it collects and assesses information on additional "adverse effects"?	• 2017/745/EU Annex III 1.1. • 2017/746/EU Annex III 1.1.
7.	Does the manufacturer describe in the PMS plan how and which information is collected to assess whether the data in the field is consistent to the expected data or training data?	• 2017/745/EU Art. 83 and 84 Annex III • 2017/746/EU Art. 78 and 79, Annex III
8.	Does the manufacturer describe in the PMS plan how and how often it will collect information on whether the product is still state-of-the-art?	• 2017/745/EU Annex III 1.1. • 2017/746/EU Annex III 1.1.
9.	Does the manufacturer describe in the PMS plan how and how often it will collect information on whether the ground truth or gold standard is still current?	• 2017/745/EU Art. 83 and 84 Annex III • 2017/746/EU Art. 78 and 79, Annex III
10.	Are the required post-market reports (PSUR, PMSR) available? Note: This can only be checked for products already on the market.	• 2017/745/EU Art. 85 and 86 • 2017/746/EU Art. 80 and 81
11.	Does the manufacturer identify in the PMS potential model drift effects (concept drift, data drift), changes to intend purpose overtime and foreseeable misuse?	• 2017/745/EU Art. 83 and 84 Annex III • 2017/746/EU Art. 78 and 79, Annex III

3. Decommissioning

1.	Does the manufacturer prepare a decommissioning plan before withdrawing its product from the market?	• / Supplementary references:

	Note: Such a plan specifies, for example, whether and how the software must be uninstalled, whether data must be backed up or exported, how the confidentiality of the data is guaranteed, who is responsible for these activities, how the progress of the decommissioning is monitored and ensured, and which organisations are to be informed and how.	• 24028
2.	Does the manufacturer identify, assess and manage the risks arising from decommissioning? Note: This should be assessed in the risk management file. Risks from the unavailability of the product, from usage errors and from an impact on other products should be considered.	• 2017/745/EU Annex I 3. • 2017/746/EU Annex I 3. • 14971, 5.4 Supplementary references: • 24028

D) Other References

- MDCG 2019-11 - Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR
https://health.ec.europa.eu/system/files/2020-09/md_mdcg_2019_11_guidance_qualification_classification_software_en_0.pdf

- MDCG 2023-4 Medical Device Software (MDSW) – Hardware combinations Guidance on MDSW intended to work in combination with hardware or hardware components
https://health.ec.europa.eu/system/files/2023-10/md_mdcg_2023-4_software_en.pdf

- U.S. Food and Drug Administration - Good Machine Learning Practice for Medical Device Development: Guiding Principles
https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

- INTERNATIONAL ELECTROTECHNICAL COMMISSION - TECHNICAL COMMITTEE 62: MEDICAL EQUIPMENT, SOFTWARE AND SYSTEMS - Supporting document listing the standards suitable to be used for development of AI/ML standards in the medical area (confirmed by IEC TC62 “medical equipment, software and systems”) - Standards suitable to be used for development of AI/ML standards in the medical area (confirmed by IEC TC62 “medical equipment, software and systems”)
https://assets.iec.ch/public/tc62/62-440-INF.pdf?2023060521