COCIR White Paper on MDR Article 61(10)

Use of performance data for the clinical evaluation of medical devices

Table of Contents

1 Introduction and context

MDR Article 61(1) states that confirmation of conformity with relevant general safety and performance requirements (GSPRs) under the normal conditions of the intended use of the device, and the evaluation of the undesirable side-effects and of the acceptability of the benefit-risk ratio, shall be based on clinical data providing sufficient clinical evidence, including where applicable relevant data as referred to in Annex III. The manufacturer shall specify and justify the level of clinical evidence necessary to demonstrate conformity with the relevant GSPRs. That level of clinical evidence shall be appropriate in view of the characteristics of the device and its intended purpose.

To that end, Article 61(10) allows the sole use of non-clinical data for demonstration of conformity with GSPRs. Adequate justification for the approach shall be given based on the results of the manufacturer's risk assessment, benefit-to-risk profile, and on consideration of the specific nature of the interaction between the device and the human body, the intended use, and the manufacturer's claims.

For devices where a clinical evaluation may be conducted under Article 61(10), non-clinical testing is an essential data source within a clinical evaluation of any medical device placed on the European Union market under the MDD or MDR. By utilizing standardized testing protocols and controlled environment, the underlaying ground truth can be established and the potential bias due to uncontrolled data reduced so that the boundary conditions can be evaluated, and the validity of the conclusions verified. Controlled design is required to assure integrity and robustness of the data. The performance established in the controlled testing environment is verified under the uncontrolled real clinical use environment with variable patient populations, users, and use environments within the device's PMCF.

Rapid technical development does not only expand the capabilities of medical devices, but also their testing environment. Considering this, digital twinning, curative databases, computer modelling, use of physical or digital phantoms or generation of artificial patients may provide controlled and scientifically valid concept to be utilized as non-clinical data within the device's clinical evaluation. Use of retrospective human or patient data depositories may constitute an additional controlled data source.

As the possibilities are practically unlimited, the focus on the assessment within the clinical evaluation should be on scientific validity of the testing methodology, test case design and the output, whether the data can be extrapolated to the expected clinical use of the device and in the intended clinical use environment, and whether the non-clinical data solely or in addition to available clinical data is sufficient to cover all clinically relevant characteristics and claims made on the device by the manufacturer, and thus demonstrate the conformity of the device with the applicable GSPRs.

Note: MDR uses the term "pre-clinical data" to describe non-clinical data generated during the product development and may be applied to substantiate devices conformity with the GSPRs. This may include testing conducted within product verification and validation processes, such as engineering, laboratory, biocompatibility, simulated use, usability, computer modelling, and utilization of animal models (Annex II, 6.1).

2 Purpose of this document

MDR Article 61 (10) is creating uncertainty on its interpretation and correct application, especially for medical devices falling into the low to moderate risk class (Class IIa) and in the moderate to high (class IIb) risk class, where the requirement to perform a clinical investigation for the demonstration of conformity with the GSPRs is not imposed by the legislation. Uncertainty on interpretation leads to disputes between various stakeholders.

The following sections are intended to clarify the concept of using the results of non-clinical testing methods, as scientific evidence to substantiate device's intended performance and safety. Understanding of the concept is essentially important to generate a harmonized regulatory framework able to adequately adapt to the variety of medical devices and their development cycles.

This document is not intended to substitute valid clinical data as the pre-requisite for the initial CE-marking of the device and in the device's post-market phase, but to illustrate the options for a certain group of devices where non-clinical data and use of retrospective human data depositories[1] can serve as valid approach for clinical performance testing.
This document does not apply to class III and implantable devices.

The hypothesis made on the expected performance and safety of the device in the (long-term) widespread clinical routine use that were based on non-clinical data, clinical data of the equivalent device(s), and/or clinical investigation(s) conducted with the device itself, must be verified within the device's post-market clinical follow-up (PMCF) (Annex XIV, Part B).

Based on this White Paper, we recommend the elaboration of a guidance document by the Medical Device Coordination Group. COCIR and the experts from our membership would gladly contribute to such work.

[1] *Use of human data is subject to legal and ethical regulations pertinent to clinical research and personal data protection.*

3 Abbreviations

GSPR	General Safety and Performance Requirements
MDSW	Medical Device Software
PMCF	Post-Market Clinical Follow up
US	Ultrasound

4 Definitions

'Performance Evaluation or Quality Control (QC) test': a series of distinct technical procedures ensuring the diagnostic device satisfactory produces high-quality images (Ref. ACR, American College of Radiology, Magnetic Resonance Imaging Quality Control Manual, 2015)

'Phantom': An object of known geometric and material composition mimicking the responses of human tissues under specified conditions

'Digital phantom': A digital model of human anatomy and physiology based on an image or digital dataset that can be projected for viewing, reproducible evaluation of reconstruction algorithms or simulation of anatomical or physiological functions.

'Qualified healthcare professional' A person qualified and entitled by national law to provide medical care in the specific fields of responsibilities. Within this document used synonymously with 'physician' and 'practitioner'.

'Retrospective data': Human data that were collected prior to the start of the data use, such as, data available in the hospital patient registry or hospital information system.

5 Clinical and non-clinical data

According to MDR Article 2, clinical data means information concerning safety or performance that is generated from the use of a medical device. Although not explicitly mentioned in the definition, the sources of clinical data imply to the use of the device in human (a patient and/or a healthy volunteer) and may include clinical investigations, PMCF studies, other PMCF data, post-market surveillance data, and reports on other clinical experience, such as clinical case studies or humanitarian use of non-approved device. The data may be generated by the manufacturer or sourced from the public data sources, such as scientific literature or vigilance databases for both, the device under evaluation or for a device where equivalency has been demonstrated

Consequently, data relevant to the safety or performance of the device that do not involve patients and/or healthy volunteers are considered as non-clinical data. Examples of non-clinical data sources include engineering or laboratory tests, animal testing, biocompatibility testing, phantom studies, reader studies utilizing artificial (patient) data, software verification and validation, and/or simulated use modelling.

Among the non-technical and prospective clinical data sources, retrospective evaluation of patient data, such as diagnostic images, patient health records, or other values in the hospital patient registry, generated within the clinical routine or other purposes (e.g., within previous studies), may constitute one of the major data sources for certain devices, such as medical device software (MDSW) for post-processing of diagnostic images or for clinical decision support. This data pool is a controlled source to reach the study objectives to produce accurate, reliable and reproducible results.

Typical use cases for retrospective human data are, for example

a) data utilization for algorithm training and validation (within development),

b) assessment of long-term performance during a retrospective cohort study (PMCF), and
c) evaluation of a diagnostic image quality or correlation of the new MDSW output compared to other available devices to achieve the same intended use.

While these kinds of devices must be able to provide accurate medical information on individuals, the final clinical outcome for the patient is dependent on further diagnostic and/or therapeutic interventions. Therefore, among non-clinical data, patient data from retrospective data sources is considered as the major body of clinical evidence for the purpose of this document.

6 Clinical evaluation based on performance data: Case studies

The case studies provided in this document are intended to provide practical examples of implementation of the clinical evaluation in accordance with MDR Article 61 (10). They do not address the entire clinical evaluation of the device which is performed in accordance with the MDR and applicable guidance (MEDDEV 2.7 / 1 Revision 4 and MDCG 2020-1 for clinical evaluation of medical device software).

6.1 Universal Image Viewer (MDSW)

Intended purpose

A software application is used for reference and diagnostic viewing of medical imaging and non-imaging data with associated reports and documents within an enterprise imaging solution. It enables healthcare professionals including (but not limited to) physicians, surgeons, and nurses to receive and view patient images, documents and data from multiple departments and organizations within one multi-disciplinary viewer and to perform common basic image manipulations and measurements (e.g., window/level, rotation, zoom, distance line and markups). The device supports diagnostic reading on medical grade monitors that are CE-marked for such use. When images are reviewed and used as an element of diagnosis, it is the responsibility of the trained physician to determine if the image quality is suitable for their clinical application.

The viewer facilitates diagnostic reading by a radiologist.

Clinical benefits

Medical images are the starting point for diagnosis of a variety of clinical situations such as cancer (e.g., bones, breast, and lungs cancer), rheumatoid arthritis, osteoporosis and fractures, conditions affecting lungs or teeth. Natural consequences of the medical conditions concerned, if not diagnosed and treated, can be physical disabilities or death. Diagnostic images are usually obtained during a period of greater clinical uncertainty, where the etiology of patient symptoms is unknown, and a broad scope of imaging may be necessary to elicit a differential diagnosis or confirm a clinical suspicion. Universal Image Viewer enables clinical users to get an overview of all the available image data of a patient as well as related medical information within a single application. Elimination of the need to switch between various applications results in faster turnaround time, increased productivity and reduces the risk of errors. When used in appropriate viewing settings, displayed images can be used directly for making a clinical diagnosis.

As the device is a workflow support tool, clinical benefits cannot be expressed in terms of measurable, patient-relevant clinical outcomes. The clinically relevant output, leading to benefits for the patient (displaying images for diagnosis), is achieved through

- diagnostic image quality
- implementing appropriate toolsets for image processing and measurements (per relevant technical standard)
- regular validation of medical grade monitors in the viewers setting
- predictable, accurate and reliable performance of the device,
- user Interface, which fosters fatigue-proof and high-performance viewing

Conclusion: Since the device has no direct influence on clinical outcomes, and the clinical benefits cannot be specified through measurable, patient-relevant clinical outcome(s), the clinical benefit is indirect.

Risks Associated with the Clinical Use of the Device

The worst hazardous situation associated with the diagnostic use of a medical image viewer is that of a misdiagnosis, delays in diagnosis and incorrect treatments which are likely to occur when the decisions are based on:

- an image of suboptimal quality,
- incomplete information, either because information is not available or because gathering related information requires additional navigation or information is not updated within a reasonable time, e.g., due to interoperability issues,
- data that are corrupted or deleted e.g., incomplete image sets,
- wrong data (incorrect patient or study or inaccurate measurements).

Healthcare decisions taken based on the images managed, displayed, and processed can impact patient care, but the risk to health is expected to be low when the product meets end users' requirements and is technically performing as intended by the manufacturer.

Risks associated with technology are linked to suboptimal image quality, inaccurate or unreliable technical performance and human errors. They can be mitigated by

- complying to technical standards,
- verification of technical performance
- sufficient in-built checks and constraints to prevent human errors
- validation of the diagnostic image quality by radiologists
- validation of the user interface and overall fitness for use by all foreseen user groups
- training and context dependent help function for end users in how to operate the system.

Technical performance

The technical performance of the MDSW is verified and validated in accordance with the appliable standards IEC 62304 and IEC 82304-1 for software.
Human factors and usability engineering is applied as defined in IEC 62366-1.
Due to the deterministic nature of the computer algorithm, the performance of the product is not expected to change during the device lifetime.

Assessment of the expected clinical performance and safety for the initial clinical evaluation

In medical imaging, information about the patient and possible abnormalities is transferred to the radiologist in two major steps: (i) data acquisition and image formation, and (ii) processing and display. Consequently, it depends critically on the design and the technical performance of the equipment, characteristics of the display and the conditions under which the image is viewed. Direct determination of clinical performance of a diagnostic viewer can be difficult as this involves the overall value of the image to the patient's diagnosis in terms of diagnostic accuracy and eventually the value of diagnosis to treatment. A diagnostic image quality can be assessed as a surrogate for clinical performance, and this can be achieved through task-oriented observer experiments. This is supported by the state-of-the-art literature where evaluation of clinical performance of diagnostic workstations includes quantitative (physical properties) and qualitative image quality assessment (reader studies), and do not measure endpoints related to diagnostic or therapeutic impact or patient outcomes.

Verification of the long-term performance and safety in the clinical routine use

The long-term performance and safety of the MDSW in the widespread clinical routine use is evaluated within the device's PMCF. Due to the indirect diagnostic purpose of the MDSW, PMCF is based on surveillance of published scientific literature for changes in the clinical state-of-the-art. Among PMCF activities, active trending of non-serious and serious complaints indicative of inadequate performance or safety issues due to the MDSW output is implemented.

6.2 Ultrasound diagnostic Imaging (US)

Intended Purpose

The diagnostic ultrasound device is based on high-frequency sound waves to produce still images of the internal anatomical structures as well to capture the real-time movement of the body's internal organs, and blood flowing through the blood vessels. The system utilizes ultrasonic transducers that, according with their technology and shape, can be distinguished in linear, convex, and phased array probes.

The ultrasound device assists primary intended users, i.e., the ultrasound practitioners, in diagnostic decision-making[2] of a variety of medical and physiological conditions.

Common ultrasound imaging diagnostic procedures include, but are not limited to:

- Abdominal ultrasound (to visualize abdominal tissues and internal organs)
- Breast ultrasound (to visualize breast tissue)
- Doppler ultrasound (to visualize blood flow through a blood vessel, organs, or other structures)
- Echocardiogram (to view the heart)
- Fetal ultrasound (to view the fetus during pregnancy)

It is known, by the physics of ultrasound waves propagation, that sound waves are subjected to refraction phenomena and energy loss/attenuation during their travelling across the human body. Consequently, the use of device in obese patients (BMI ≥ 30) may result sub-optimal due to thickness of subcutaneous fat and the higher sound-attenuating properties of fat itself.

Furthermore, it is also known that air may represent a natural barrier to ultrasound waves propagation, consequently, Lungs or Bowel air/gas may prevent the sonographic imaging of the underneath organs (e.g., the heart and of abdominal organs).

All the above-mentioned aspects are intrinsic limitations of ultrasound imaging.

Therefore, healthcare professionals are the primary and ultimate responsible for determining if the image quality, achieved during the diagnostic procedure, is appropriate for their clinical decision making in individual situations during their daily scope of practice.

Nature of the interaction of the device with the human body

Ultrasound medical diagnostic devices are active devices. They come in contact with the human body by means of ultrasonic transducers. A layer of ultrasound gel is interposed between the probe surface and the patient's skin to allow the proper propagation of the ultrasound acoustic waves to the patient's body part under investigation.

The contact duration is *transient*; normally intended for continuous use for less than 60 minutes (Annex VIII Section 1.1). In particular, when those devices are utilized for imaging organs and body structures with linear, convex and phased array ultrasonic probes, they come into contact with the patient's body though intact skin.

[2] *Definition: "Clinical decision making is a contextual, continuous, and evolving process, where data are gathered, interpreted, and evaluated in order to select an evidence-based choice of action.", Enhancing Clinical Decision Making: Development of A Contiguous Definition and Conceptual Framework" , Tiffen J et al. J Prof Nurs. 2014 Sep-Oct;30(5):399-405.*

Another possible mean of contact with the human body is through natural orifices. In such a case endocavitary (e.g. endorectal and/or endovaginal) and transesophageal transducers are utilized.

Manufacturer's claim

Ultrasound scan is a diagnostic procedure that uses high-frequency sound waves to form an anatomical image of body's internal organ in order to aid the primary intended users of the device to monitor, evaluate or diagnose several medical conditions as well to provide a real-time imaging support during the so called "ultrasound-assisted" minimally invasive procedures.

The primary intended users of ultrasound medical diagnostic devices are ultrasound practitioners[3], i.e., healthcare professionals holding recognized qualifications in medical ultrasound and that are able to competently perform ultrasound examination falling within their personal scope of practice.

Clinical benefits

The diagnostic ultrasound applications are designed to assist clinical decision-making of various diseases and physiological conditions.
Additionally, ultrasound imaging may be used to assist minimally invasive procedures where the device is used to assist other medical devices (e.g., biopsy needle, ablation therapy devices, etc.) to achieve their own intended purposes.
Conclusion: the clinical benefit associated with an ultrasound diagnostic imaging device use can be define as indirect: in other words, the device itself does not directly achieve the positive impact on patients, as in it not intended to treat or diagnose directly, rather it assists primary intended users in clinical decision-making, or other devices to achieve their own intended purposes (as in the case of ultrasound assisted minimally invasive procedures).

Risks Associated with the Clinical Use of the Device

Ultrasound imaging does not utilize ionizing radiation.

Ultrasound energy has the potential to produce biological effects on the body, such as heating the tissues or producing gas in body fluids or tissues.

According to the American Institute of Ultrasound in Medicine (AIUM), "*diagnostic ultrasound has been in use since late 1950s and no independently confirmed adverse effects caused by exposure from present diagnostic ultrasound instruments have been reported in human patients in absence of contrast agents*"[4]. .
However, in consideration of those potential biological effects, the use of diagnostic ultrasound for non-medical purposes is discouraged, especially within fetal ultrasound procedures.

Inadequate image quality, incorrect measurements, or unavailability or incompleteness of the image data may contribute to, or result in, wrong or delayed clinical decision-making .

Ultrasound transmission gel is applied to the skin is required to allow a properly propagation of the ultrasound waves from the transducers head to the patient body.

[3] *Ref. Guidelines for professional ultrasound practice Rev5 - SCor (The Society & College of Radiographers) and BMUS (British Medical Ultrasound Society)- 2020, December]*
[4] *https://www.aium.org/officialStatements/34?__sw_csrfToken=433806de*

The use of ultrasound gel may induce allergic reactions, or it has been associated in the past to cases of microbial contaminations that have led to serious clinical infections to patients.

The risks associated with the contrast agents utilized with contrast enhanced ultrasound applications, such as potential hypersensitivity to active substances present in the medical product, uncontrolled systemic hypertension, skin erythema, bradycardia, dyspnea, loss of consciousness, cardiac/cardio-respiratory arrest, anaphylactic reaction, or anaphylactic shock are evaluated by the manufacturer of the respective medicinal products and specific indications, contraindications and precautions are reported in the Ultrasound Contrast Agent IFUs.

Technical performance

Ultrasound medical diagnostic imaging represent a well-established use of the technology in medicine.
As such the technical performance of the diagnostic ultrasound device is verified and validated in accordance with the appliable harmonized standards for ultrasound imaging:

- IEC 60601-2-37 for basic safety and essential performance of ultrasonic diagnostic imaging,
- IEC 61157 for reporting of the acoustic output of medical diagnostic ultrasonic equipment,
- IEC TS 62736 to verify stability of an imaging system's elementary performance,
- IEC TS 61206 for continuous wave doppler systems,
- IEC TS 61390 for real-time pulse-echo systems, and IEC TS 62791 for utilization of low-echo sphere phantoms and method for performance testing of gray-scale medical ultrasound scanners applicable to a broad range of transducer types.
- ISO 10993 for biological evaluation of medical devices
- IEC 62366-1 to evaluate the human factors and usability engineering aspects related to a medical device

Clinical performance and safety

Expected performance and safety of Ultrasound Medical Device in clinical use is evaluated as specified in the MDR, applicable MDCGs related to clinical investigation and evaluation and applicable parts of the MEDDEV 2.7/1 revision 4.

Assessment of the expected clinical performance and safety for the initial clinical evaluation.

Since the device has no direct influence on clinical outcomes, image quality can be assessed as a surrogate for clinical performance.
The image quality can be effectively evaluated utilizing imaging Performance Evaluation or Quality Control (QC) test based on the use of ultrasound phantoms.

Phantoms, in medical imaging, are imaging specimens of known geometric and material composition and, in particular, *tissue-mimicking phantom* emulates important properties of biological *tissue* for the purpose of providing a more clinically realistic imaging test.

The use of imaging phantoms allows:

1. **Absolute and relative imaging capability**

The device's absolute and relative imaging capabilities are evaluated by utilizing established ultrasound medical imaging phantoms of known geometric and material composition. Tissue-mimicking phantoms emulate properties of specific biological tissues and anatomical structures.
The expected clinical performance is evaluated under standardized conditions utilizing physical ultrasound phantoms and technical testing parameters.

Absolute imaging capabilities are defined by repetitive measurements against affirmed quantifiable standards which provide consistent objective test ensuring that the device performance in terms of image quality and stability fulfill the requirements for the expected diagnostic performance.

Relative imaging capabilities are defined as parallel measurements with the new ultrasound device and devices representing the current clinical state of the art in the respective ultrasound imaging applications.

Within the testing procedures, test cases are designed to evaluate and verify:

- Image uniformity and artifact survey
- System sensitivity: visual determination of the maximum depth of visualization of speckle patterns or phantom targets, and quantitative measurements of signal-to-noise ratio
- Geometric accuracy: measurement of known distances between the phantom test targets in the axial and lateral directions,
- Contrast resolution: the use of anechoic and low contrast echogenic targets as well as 2-D cylindrical and 3-D spherical targets. The use of larger 2-D targets emphasizes contrast resolution performance, whereas the use of small targets also tests spatial resolution capabilities.
- Spatial resolution: measurements in the axial, lateral, and elevational directions, including visual interpretation of groups of phantom pin/fiber targets and measurement of pin target dimensions, and elevational resolution measurements with special phantom and multipurpose phantoms
- Fidelity of the display device(s) used for primary interpretation: when used for diagnostic purposes, the electronic displays on the scanner and any modality workstations are considered as primary diagnostic devices, for scanner applications used exclusively as an aid to guide therapeutic procedures, no workstation are considered in the primary device set-up. Display characteristics that are evaluated include grayscale response, presence of pixel defects, and overall image quality. These evaluations are performed using specialized test pattern images and involve the use of photometric equipment.
- Doppler functionality:
  Qualitative visual testing procedures are applied to evaluate:
  (A) Spectral doppler mode with test cases relevant for the (1) positioning of the doppler sampling volume, (2) specification of doppler angle, (3) doppler spectral display, and (4) directionality of flow, and lack of velocity signal where no flow is present
  (B) Color flow imaging mode with test cases relevant for color map and flow direction and color signal superimposition on the grayscale image. Quantitative test procedures are applied to evaluate
  (C) Doppler sensitivity as a function of depth in attenuating media (i.e. determination of the lowest detectable flow) with test cases relevant for (1) verification of velocity measurement accuracy over a clinical range, including pathologies, such as stenosis;

(2) Verification of correct directional discrimination; (3) Accuracy of angle correction, (4) Assessment of gate/sample volume registration, and (5) Verification of volume flow measurement accuracy

- Elastography functionality :
  Qualitative and quantitative testing procedures are applied to evaluate (1) Assessment of stiffness measurement accuracy as a function of depth in attenuating media, and (2) contrast-detail assessment of elastography imaging performance.

2. Contrast enhanced imaging

Contrast enhanced imaging is obtained when an ultrasound medical device is used, within its intended purposes, in conjunction with a medical product (e.g., an ultrasound contrast-enhancement agent).

In this case, the ultrasound medical device is enabling the Contrast Enhancement agent in use to achieve its own intended therapeutic indications of enhancing the echogenicity of the blood, or of fluids which results in an improved signal to noise ratio in ultrasound imaging.

In this context, performance test of the ultrasound device can be conducted by injecting the targeted contrast agent in doppler ultrasound flow phantoms filled blood mimicking fluid (BMF) simulating the acoustic and physical properties of human blood. The test is designed to evaluate the enhancement of the echogenicity of the blood in ultrasound imaging and the delineation of the blood-vessel-simulating, ultrasound-compatible tube present inside the phantom due to the injection of the contrast media in the pumping system of the phantom.

It should also be added that the contrast media are subjected rigorous clinical testing for safety and efficacy before approval by the Competent Authorities for free sale is granted.

As such, risks associated with the contrast agent use, such as potential hypersensitivity to active substances present in the medical product, uncontrolled systemic hypertension, skin erythema, bradycardia, dyspnea, loss of consciousness, cardiac/cardio-respiratory arrest, anaphylactic reaction, or anaphylactic shock are carefully evaluated and investigated by the manufacturer of the respective medicinal products.

3. Ultrasound-assisted minimal invasive procedures

Ultrasound-assisted minimally invasive procedures include real-time visualization and monitoring of procedures conducted with (1) biopsy needles and guides to obtain core biopsy samples from soft tissue such as kidney, liver, prostate, and various soft tissue masses, (2) fine needles intended for fine needle aspiration procedures where a small amount of tissue or fluid is removed from a suspicious area with a thin, hollow needle to confirm a diagnosis or guide treatment, (3) ablation systems for percutaneous ablation of tissue using radiofrequency or thermal energy, chemical products (e.g. injection of concentrated ethanol alcohol), or laser light to induce cell death.

In this case, the ultrasound medical device is used, within its original intended purposes, in conjunction with other medical devices (e.g., fine needles, core biopsy needles, needle

guidance kits, thermal, laser or chemical ablators etc.) to help them achieving their own intended purposes

To evaluate the capability of the ultrasound device with superficial probes to provide images of the internal body and tissue structures and their suitability to assist and provide a support in minimally invasive procedures, specific tissue mimicking imaging phantom incorporating cists or solid masses/ lesions are utilized.

Verification of the long-term performance and safety in the clinical routine use

The long-term performance and safety of the ultrasound device in the widespread clinical routine use is evaluated within the device's PMCF.

General and specific PMCF methodology is applied.

As ultrasound devices represent well-established use of the technology in medical imaging field, PMCF is generally based on surveillance of published scientific literature for changes in the clinical state-of-the-art. Among PMCF activities, pro-active monitoring of non-serious and serious complaints on the device in question and/or on devices belonging to the same generic device group is implemented aiming the identification of possible inadequate performances or of possible news risks in some specific or new use cases..

References

- [1] AIUM Routine Quality Assurance of Clinical Ultrasound Equipment, version 2.0, 2020
- [2] AIUM Routine Quality Assurance "Cookbook", a companion guide to the reference above
- [3] ACR–AAPM Technical Standard for Diagnostic Medical Physics Performance Monitoring of Real Time Ultrasound Equipment – revision 2021 https://www.acr.org/-/media/ACR/Files/Practice-Parameters/US-Equip.pdf
- [4] ACR Ultrasound/Breast Ultrasound Quality Control accreditation requirements – Revised 3-2-2021 https://accreditationsupport.acr.org/support/solutions/articles/11000064162-quality-control-ultrasound-and-breast-ultrasound
- [5] Guideline for Technical Quality Assurance (TQA) of ultrasound devices (B-Mode)--version 1.0 (July 2012): EFSUMB Technical Quality Assurance Group--US-TQA/B, Ultraschall in Med 2012; 33: 544–549, DOI http://dx.doi.org/
- 10.1055/s-0032-1325347
- [6] BMUS QA Working Party. BMUS guidelines for the regular quality assurance testing of ultrasound scanners by sonographers. Ultrasound 22, 8-14, 2014
- [7] BMUS QA working party. The BMUS guidelines for regular quality assurance testing of ultrasound scanners. Ultrasound. 2014 Feb;22(1):6-7. doi: 10.1177/1742271X13515557
- [8] Institute of Physics and Engineering in Medicine (IPEM). Quality Assurance of Ultrasound Imaging Systems. Report No. 102. York: IPEM, 2010 (https://www.ipem.ac.uk/ScientificJournalsPublications/IPEMReportSeries/AvailablePublications.aspx)

6.3 Clinical Decision Support for Treatment Planning (MDSW)

Intended Purpose

Clinical Decision Support and Treatment Planning is a novel MDSW. The MDSW is intended to support clinical decision making and therapy planning. The software provides a multi-modality digital platform to incorporate relevant patient data available from the patient's health records as well as laboratory tests and diagnostic images to one user interface so that a personalized treatment plan can be established by the physician or hospital tumor board.

The MDSW utilizes natural language processing to screen patient's health records, hospital information system, and associated Picture Archive Communication System (PACS) for the information on patient's disease and treatment status that is displayed on the user interface. The relevant information includes diagnostic results from radiology, pathology, genetics, and laboratory testing.

The agglomeration of data is correlated with the established cancer scoring indices that enable estimation of the tumor aggressiveness and the course of the cancer. Additionally, the algorithm is mapping the patient information aggregate with the relevant physician's guidelines on the treatment of cancer. Based on the individual disease state of the patient and relevant clinical practice guidelines, MDSW provides suggestions on the treatment pathways and options. If relevant information is missing, this is indicated on the user interface.

The MDSW does not make diagnostic decisions but supports the user in decision making. Individual recommendations generated as the MDSW output must be confirmed or rejected on the user interface by the physician.

Clinical benefits

The MDSW is designed to improve workflow efficiency and optimize decision making for multidisciplinary teams involved in the tumor boards. This is based on the transparent and structured information content presented on the user interface as well as ongoing adaptation to the applicable evidence-based medical guidelines.

The natural language processing enables subtraction of relevant data sets from the data pool and reduction of inter-user variability and treatment errors due to missing information or human errors.

Risks Associated with the Clinical Use of the Device

Wrong or delayed treatment decision due to wrong or unclear MDSW input or output
Delayed treatment decision due to unavailability of the MDSW

Technical performance

The technical performance of the MDSW is verified and validated in accordance with the appliable good manufacturing standards IEC 62304 and IEC 82304-1 for software.
Human factors and usability engineering is applied as defined in IEC 62366-1.

Clinical performance and safety

Expected performance and safety of the MDSW in clinical use was evaluated as specified in the MDCG 2020-1.

Assessment of the expected clinical performance and safety for the initial clinical evaluation

(1) Natural language processing

The natural language processing capability of MDSW is tested in standardized environment by using a curated database with large-scale sets of artificial records mimicking medical notations in the unstructured health data. Within this, generalizability of the MDSW output in terms of semantic variations in medical terminology is correlated with the expected values.
To analyze algorithm performance among the real use environment, a specific registry study is conducted utilizing raw health data of the cancer patients that are admitted to the hospital in the defined time frame. The quality of data mining and extracted data is assessed in terms of accuracy, reliability, trueness, and precision. The data rate, availability, confidentiality, and integrity of the MDSW output, and potential cybersecurity vulnerabilities are evaluated at the study sites with varying IT infrastructure environments.

(2) Treatment pathway

The user interface and expected clinical performance of the MDSW is assessed within a simulated use environment. In the study, standardized sets of artificial patients are created simulating distribution of pathological patient values and expected normal values in the general population, as well as implausible and incomplete diagnostic findings. The readers participating in the study are qualified in the groups based on their professional experience in the respective field of oncology: in foundation training, mid-term professionals, and experienced. The study design is double blinded. The readers perform cross-reading of the artificial patient data sets utilizing MDSW. The automated reading results, including staging (stages primary/in situ to secondary/metastatic cancer) and treatment pathways as MDSW output, are correlated with annotated scoring and treatment pathways defined in the physicians' guidelines. Among correlation of the MDSW output with the manual reading, the time-to-result is recorded for automated and manual treatment decision.

Verification of the long-term performance and safety in the clinical routine use

The long-term performance and safety of the MDSW in the widespread clinical routine use is evaluated within the device's PMCF. As the MDSW represented innovative technology, specific PMCF methodology is applied. Representative reference sites are selected based on the number of oncological examinations performed and pre-defined criteria to minimize bias. The MDSW is equipped by a software application that records the manual corrections and time-to-result made by the physician on the MDSW output over the specified timeframe. Additionally, after each use of the MDSW, the physician is requested specific information over the questionnaire on the user interface on the user's professional experience as well as potential issues with the performance or safety of the MDSW. The PMCF results from the reference sites are monitored by the manufacturer for any indication of inadequate performance of the MDSW. Among specific PMCF activities, active trending of non-serious and serious complaints indicative of wrong or delayed treatment decision due to the MDSW output is implemented.

References

- IEC 62304 Medical device software - Software life cycle processes
- IEC 62366-1 Medical Devices – Part 1: Application of usability engineering to medical devices
- IEC 82304-1 Health software – Part 1: General requirements for product safety

- MDCG 2020-1 Guidance on Clinical Evaluation (MDR)/ Performance Evaluation (IVDR) of Medical Device Software, revision March 2020

6.4 Clinical Decision Suport System for MR-based Lesion Detection and Characterization (MDSW)

Intended Purpose
The clinical decision support system is a post-processing MDSW that performs analysis of contrast-enhanced magnetic resonance (MR) images of brain (diffusion weighted imaging, diffusion tensor imaging, and perfusion weighted imaging). MDSW performs automatic segmentation of different brain regions and transfers the output to digital tissue data so that the brain anatomy can be described numerically instead of a visual analysis. The MDSW extracts quantitative features from MR images and marks deviations indicative of lesions suspected of cancer.

The MDSW does not make diagnostic decisions but supports the user in decision making. MDSW output must be confirmed or rejected on the user interface by the physician.

Clinical benefits
The MDSW is designed to improve diagnostic, prognostic, and predictive accuracy of oncological decision-making. High-resolution 3D MR imaging enables a high contrast-to-noise ratio of lesions. AI-assisted interpretation of imaging data reduces inter-reader variability, improves detection sensitivity of discernible structures, reduces wrong diagnostic findings due to human error, and optimizes the differential diagnosis of brain pathologies. Ultimately, this leads to improved patient outcomes and reduced healthcare costs.

Risks Associated with the Clinical Use of the Device
Wrong or delayed treatment decision due to wrong or unclear MDSW input or output.
Delayed treatment decision due to unavailability of the MDSW.

Technical performance
The technical performance of the MDSW is verified and validated in accordance with the appliable standards IEC 62304 and IEC 82304-1 for software.
Human factors and usability engineering is applied as defined in IEC 62366-1.

Clinical performance and safety

Expected performance and safety of the MDSW in clinical use was evaluated as specified in the MDCG 2020-1.

Assessment of the expected clinical performance and safety for the initial clinical evaluation

The analytical performance of the MDSW is assessed by using a four-dimensional (4D) digital phantom that has been designed to provide a realistic model of human anatomy. The phantom is based on human diagnostic imaging datasets obtained by CT and MR imaging and allows modelling different anatomical variations. The phantom is generated as there is no ground-truth standard available for comparison of quantitative image

reconstruction abilities of the MDSW in-vivo due to variability of human tissues and dynamic contrast agent perfusion (depending on heart rate, respiration, etc.).
The accuracy of the automatic organ segmentation is evaluated by simulating different sizes of the different brain regions, and different physiological variations such as patient ages and ethnicities. The lesion detection capability in different brain regions is measured by introducing artificial lesions with varying sizes, shapes, and configurations in the different brain regions, and by simulating different dynamic contrast-enhancement (perfusion) patterns for the same patient geometry.

The AI-based segmentation and characterization abilities of MDSW are assessed within a retrospective reader study. In the study, radiologists read curated data sets of brain MR examinations acquired according to standard protocols with already diagnosed and histologically confirmed tumor diseases with and without the aid of the MDSW. The data sets are annotated in the diagnostic classes (affected brain region, volume, cancer stage) by independent neuroradiologists. Neuroradiologists' reading performance in AI-aided and unaided scenarios will be compared using receiver-operating characteristics (ROC) analysis in a multi-reader, multi-case study. Effectiveness of the AI-assisted characterization is demonstrated when the detection accuracy of brain cancer, based on histological results and patient follow-up, improves when MDSW is used.

Verification of the long-term performance and safety in the clinical routine use

Reference sites are enclosed in a 2-arm prospective PMCF study to assess the performance and safety of the MDSW among oncological routine settings. The primary and secondary endpoints address inter-user variability, workflow efficiency, and accuracy of the MDSW output compared with the standard of care.
A retrospective hospital registry data analysis is conducted at representative sites selected based on the number of oncological brain examinations per annum. The DICOM data stored in the hospital information system is reassessed to compare AI-assisted diagnostic accuracy with the information on the patient follow-up history in the hospital registry.

References

- IEC 62304 Medical device software - Software life cycle processes
- IEC 62366-1 Medical Devices – Part 1: Application of usability engineering to medical devices
- IEC 82304-1 Health software – Part 1: General requirements for product safety
- MDCG 2020-1 Guidance on Clinical Evaluation (MDR)/ Performance Evaluation (IVDR) of Medical Device Software, revision March 2020

6.5 CT Volumetry Clinical Decision Support System (MDSW)

Intended use

The CT Volumetry Clinical Decision Support System is a post-processing MDSW that is intended for automated volumetric analysis and segmentation of different organ structures. The software performs three-dimensional (3D) analysis of objects on medical images generated by the computed tomography (CT). The MDSW may be used for oncological applications, including primary and metastatic cancer assessment, therapy planning, and evaluation of response to the therapy. The MDSW may also be used for diagnosis and management of other diseases, such as ischemic or hemorrhagic stroke or pulmonary diseases.
The analysis results are automatically compared to a normative database and deviations indicated. The MDSW enables automatic segmentation organs by differentiation of tissue

characteristics and marking of suspected areas by the physician. Additionally, data can be recorded and monitored over time or additional analyses performed retrospectively.

Clinical benefits

The MDSW is designed to predict clinical outcomes and to improve clinical decision-making of various diseases and monitoring therapy progress. Asymmetric and irregular lesions can be accurately quantified. Automatic calculation of pathologic lesion volume reduces inter-user variability, optimizes workflow efficiency, and enables monitoring of changes in the lesions over the time. The visualization of lesion changes and differentiation of tissue characteristics enables stratification of a treatment or prediction of clinical patient outcomes.

Risks Associated with the Clinical Use of the Device

Wrong or delayed treatment decision due to wrong or unclear MDSW input or output
Delayed treatment decision due to unavailability of the MDSW

Technical performance

The technical performance of the MDSW is verified and validated in accordance with the appliable good manufacturing standards IEC 62304 and IEC 82304-1 for software. Additionally, IEC 606061-2-44 is applied.
Human factors and usability engineering is applied as defined in IEC 62366-1.

Clinical performance and safety

Expected performance and safety of the MDSW in clinical use was evaluated as specified in the MDCG 2020-1.

Assessment of the expected clinical performance and safety for the initial clinical evaluation

The performance of the MDSW is initially assessed by using a physical four-dimensional (4D) and a digital phantom.
In the first step, image datasets are generated by the repeated scanning of standardized physical phantoms mimicking human anatomy and different pathologies. The scan protocols are used to prove capabilities of the MDSW for lesion delineations, as well as differentiation of organ structures and tissue types under standardized and controlled testing environment.
In a second step, digital phantoms providing computational models of the patient anatomies and motion are used. The phantoms are based on human diagnostic imaging datasets that model anatomical variations and moving targets. The sizes, shapes, and configurations of the object structures are modelled to simulate non-spherical changes in the space-occupying lesions over time.
Repeated measurements are conducted to evaluate the analytical performance of the MDSW, including accuracy and limits of detection of the volume measurements in different anatomical structures. Additionally, correlation and the predictive values of the MDSW output with the expected values in the normative databases are analyzed.
Clinical sensitivity and specificity of the MDSW is verified within a retrospective study utilizing curated and annotated data sets of DICOM images of patients with diverse pathologies (tumor, ischemic stroke, hemorrhagic stroke, pulmonary disease) and classes

(affected organ, volume, disease stage) collected during the clinical routine at medical care centers. The MDSW output is correlated with the manual reconstruction of cross-sectional areas as regions of interest (ROI), and software-aided contour tracing of the lesions as the standard of care.
For assessment of clinical accuracy of the MDSW reading in different tissues with varying characteristics and differential diagnostics of lesions, a standardized ground-truth is required. This is determined by agglomerating patients' clinical follow-up information (confirmed primary diagnosis and follow-up) and retrospective analysis of imaging databases. Ground truth cannot be determined in a prospective imaging study as relevant information only becomes available or is confirmed by further procedures (e.g. interventional biopsies or surgical procedures) during the patients' follow-up.

Verification of the long-term performance and safety in the clinical routine use

Long-term performance and safety of the MDSW is evaluated in two phases within the device's PMCF.
In the first phase, reference sites are enclosed in a 2-arm prospective PMCF study to assess the performance, safety, and efficacy of the MDSW compared with the standard of care among clinical routine settings in the oncological, neurological, and emergency care units. The study design includes no additional procedures that are invasive or burdensome to the study participants.
In the second phase, two prospective multi-center PMCF studies are conducted.
A randomized study is designed to compare clinical patient outcome of ischemic and hemorrhagic stroke in terms of recurrence, functional outcome, mortality, survival rate, and morbidity with a diagnostic procedure supported by the MDSW and manual reading as the standard of care.
The primary endpoint of the second study is to evaluate whether automated CT volumetry may replace the semi-automated pre- and post-procedural staging of cancer as the standard of care. Over a defined period, the staging based on the MDSW output is compared with the conventional diagnostic procedures for detection of primary cancer and follow-up of treatment progress.

References

- IEC 62304 Medical device software - Software life cycle processes
- IEC 62366-1 Medical Devices – Part 1: Application of usability engineering to medical devices
- IEC 82304-1 Health software – Part 1: General requirements for product safety
- IEC 60601-2-44 Medical electrical equipment – Particular requirements for the basic safety and essential performance of X-ray equipment for computed tomography
- MDCG 2020-1 Guidance on Clinical Evaluation (MDR)/ Performance Evaluation (IVDR) of Medical Device Software, revision March 2020