MDCG 2020-1

Guidance on Clinical Evaluation (MDR) / Performance Evaluation (IVDR) of Medical Device Software

March 2020

This document has been endorsed by the Medical Device Coordination Group (MDCG) established by Article 103 of Regulation (EU) 2017/745. The MDCG is composed of representatives of all Member States and it is chaired by a representative of the European Commission.

The document is not a European Commission document and it cannot be regarded as reflecting the official position of the European Commission. Any views expressed in this document are not legally binding and only the Court of Justice of the European Union can give binding interpretations of Union law.

Guidance on Clinical Evaluation (MDR) / Performance Evaluation (IVDR) of Medical Device Software

Table of Contents

1. Purpose 3
2. Scope 3
3. Background 4
   3.1. Abbreviations 5
   3.2. Formats used within this document 5
   3.3. Definitions 5
4. General principles of the MDSW CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) process 9
   4.1. Introduction 9
   4.2. Determination of the valid clinical association / scientific validity 12
   4.3. Technical Performance /Analytical Performance 12
   4.4. Clinical Performance 13
      4.4.1. Clinical investigations and clinical performance studies 14
      4.4.2. Where demonstration of conformity based on clinical data is not deemed appropriate 15
   4.5. Final analysis and conclusion of the clinical evaluation (MDR) / performance evaluation (IVDR) 15
   4.6. Continuous update of the clinical evaluation (MDR) / performance evaluation (IVDR) 15
   Annex I – Methodological principle for generation of CLINICAL EVIDENCE 17
   Annex II – Examples of CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) strategies 18
      a) MDSW intended to analyse sleep quality data 18
      b) MDSW intended for image segmentation 19
      c) MDSW intended to detect inflammatory bowel diseases (IBD) 20
      d) Active devices containing MDSW to enable their intended purpose 21
      e) MDSW which provides an additional user-interface to control an insulin pump 21
      f) MDSW intended to analyse exhaled CO2 in a life-sustaining device in order to control ventilator settings 21

1. Purpose

The purpose of this guidance is to provide a framework for the determination of the appropriate level of CLINICAL EVIDENCE required for MEDICAL DEVICE SOFTWARE (MDSW) to fulfil the requirements set out in Regulation (EU) 2017/745 – Medical Devices Regulation (MDR) and Regulation (EU) 2017/746 – *In Vitro* Diagnostic Medical Devices Regulation (IVDR).[1]

In order to promote global convergence, this document takes into account certain concepts outlined in International Medical Device Regulators Forum (IMDRF) guidance documents (such as N41).[2]

2. Scope

This guidance should be applied to MDSW. For the purpose of this guidance, MDSW is software that is intended to be used, alone or in combination, for a purpose as specified in the definition of a "medical device" in the medical devices regulation or *in vitro* diagnostic medical devices regulation.

It should be noted that software can be associated[3] with another medical device, by driving or influencing its use. The guideline MDCG 2019-11 clarifies that software which is driving or influencing is covered by the medical devices regulations[4] either as a part/component of a device or as an accessory for a medical device.

Software developers should refer to MDCG 2019-11 for guidance on the appropriate qualification and classification of software prior to such software being introduced into the market. The same principles of CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) apply to all MDSW. Conceptually, the following models of software can be understood (whereas combinations may be possible, refer to Table 1):

a) Software for which the manufacturer claims a specific medical intended purpose. Such software has a CLINICAL BENEFIT and requires CLINICAL EVIDENCE within its own conformity assessment.

b) Software for which the manufacturer does not claim any medical intended purpose. Such software is intended to drive or influence a medical device. The CLINICAL EVIDENCE is provided within the context of the driven or influenced device and is therefore out of the scope of this document.

It should be recognised that the concept of a CLINICAL BENEFIT for MDSW may deviate from that which applies in the case of pharmaceuticals or other medical devices, since the benefit of MDSW may lie in providing accurate medical information on patients, where appropriate, assessed against medical information obtained through the use of other diagnostic options and technologies, whereas the final clinical outcome for the patient is dependent on further diagnostic and/or therapeutic options which could be available.

[1] Depending on the device in question, the level of Clinical Evidence may differ and shall be assessed on a case by case basis.
[2] International Medical Device Regulators Forum – IMDRF/SaMD WG/N41FINAL:2017 – Guidance on Software as a Medical Device (SaMD): Clinical Evaluation
[3] Associated medical device may be software or hardware.
[4] The use of "The Medical Devices Regulations" from here on out refers to both Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR.

Model of Software	CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) - scope
MDSW (with independent intended purpose and claimed CLINICAL BENEFIT)	MDSW only
MDSW (with intended purpose and claimed CLINICAL BENEFIT related to driving or influencing a medical device for a medical purpose)	MDSW and the driven or influenced medical device Notes 1,2
Software driving or influencing the use of a medical device (with no independent intended purpose or independent claimed CLINICAL BENEFIT)	Driven or influenced medical device including the software (component or accessory)

Table 1 Different MDSW and CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) requirements

Note 1: If a software is driving/ influencing more than one medical device, an independent CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) *is required for each foreseen and clinically viable software – device combination.*

Note 2: Out of scope of this guidance (See MDCG 2019-11 for examples).

3. Background

Article 61 (1) of the MDR and Article 56 (1) of the IVDR state the following:

> 'The manufacturer shall specify and justify the level of CLINICAL EVIDENCE necessary to demonstrate conformity with the relevant general safety and performance requirements. That level of CLINICAL EVIDENCE shall be appropriate in view of the characteristics of the device and its intended purpose.'

Article 2 (51) of the MDR and Article 2 (36) of the IVDR define 'CLINICAL EVIDENCE' as:

> 'clinical data and CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) results pertaining to a device of a sufficient amount and quality to allow a qualified assessment of whether the device is safe and achieves the intended CLINICAL BENEFIT(S), when used as intended by the manufacturer.'

In order to provide guidance relating to the level of CLINICAL EVIDENCE required for MDSW and as set out in recital (5) of the MDR and IVDR, this guidance takes into account internationally converged principles adopted by an international group of regulators, IMDRF (http://www.imdrf.org). Adoption of these principles provides European regulators an initial framework when further developing MDR/ IVDR-specific regulatory approaches and expectations for regulatory oversight.

While this document describes a converged approach to CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) **for MDSW**, it should be read in conjunction with other documents that aim to provide

horizontal guidance for the CLINICAL EVALUATION of medical devices or PERFORMANCE EVALUATION of *in vitro* diagnostic medical devices.[5]

Note: Please be advised that this document is subject to revision upon the publication of the aforementioned horizontal guidance.

Clinical expertise and judgments are required at every step of the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR), including literature search and appraisal. Each indication and claimed CLINICAL BENEFIT that is part of the intended purpose should be assessed individually and have the supporting CLINICAL EVIDENCE. Systematic and explicit approach for the appraisal of supporting data allows achieving confident, scientifically substantiated conclusions and facilitates transparency of these judgments.

3.1. Abbreviations

GSPR	General Safety and Performance Requirements
IMDRF	International Medical Device Regulators Forum
IVDR	*In Vitro* Diagnostic Medical Devices Regulation; EU 2017/746
MDCG	Medical Device Coordination Group
MDR	Medical Devices Regulation; EU 2017/745
MDSW	Medical Device Software
PMCF	Post Market Clinical Follow-up
PMPF	Post Market Performance Follow-up
PMS	Post Market Surveillance
RWE	Real-World Evidence
SaMD	Software as a Medical Device
SOTA	State-of-the-Art
SSCP	Summary of Safety and Clinical Performance
SSP	Summary of Safety and Performance

3.2. Formats used within this document

Cursive	A note to a text
CAPITALIZED	Terms defined in this document or the Regulations
subscript	References

3.3. Definitions

The definitions elaborated within this section and utilised within this document are **intended to apply solely to Medical Device Software (MDSW)** according to the MDR and IVDR.

[5] These guidance documents are under development and will be published on the Commission's Medical Devices website.

CLINICAL BENEFIT	Article 2 (53) MDR defines CLINICAL BENEFIT as the positive impact of a device on the health of an individual, expressed in the terms of a meaningful, measurable, patient-relevant clinical outcome(s), including outcome(s) related to diagnosis, or a positive impact on patient management or public health; whereas Article 2 (37) IVDR defines CLINICAL BENEFIT as the positive impact of a device related to its function, such as that of screening, monitoring, diagnosis or aid to diagnosis of patients, or a positive impact on patient management or public health.[6] Source: EU 2017/745 (MDR), Article 2 (53); EU 2017/746 (IVDR), Article 2 (37) and IVDR recital (64)
CLINICAL DATA	(MDR) Information concerning safety or performance that is generated from the use of a device and is sourced from the following: - clinical investigation(s) of the device concerned, - clinical investigation(s) or other studies reported in scientific literature, of a device for which equivalence to the device in question can be demonstrated, - reports published in peer reviewed scientific literature on other clinical experience of either the device in question or a device for which equivalence to the device in question can be demonstrated, - clinically relevant information coming from post-market surveillance, in particular the post-market clinical follow-up; Source: EU 2017/745 (MDR) (IVDR) Clinical Data, in particular: - from relevant peer-reviewed scientific literature and available consensus expert opinions or positions from relevant professional associations relating to the safety, performance, clinical benefits to patients, design characteristics, scientific validity, clinical performance and intended purpose of the device and/or of equivalent or similar devices; or - other relevant clinical data available relating to the safety, scientific validity, clinical performance, clinical benefits to patients, design characteristics and intended purpose of similar devices, including details of their similarities and differences with the device in question - clinically relevant information coming from post-market surveillance, in particular the post-market performance follow-up; Source: Adopted from EU 2017/746 (IVDR) Annex XIV (2.4) and Annex VII (4.10) and (4.11)

[6] IVDR recital (64) states : It should be recognised that the concept of clinical benefit for *in vitro* diagnostic medical devices is fundamentally different from that which applies in the case of pharmaceuticals or of therapeutic medical devices, since the benefit of *in vitro* diagnostic medical devices lies in providing accurate medical information on patients, where appropriate, assessed against medical information obtained through the use of other diagnostic options and technologies, whereas the final clinical outcome for the patient is dependent on further diagnostic and/or therapeutic options which could be available.

CLINICAL DEVELOPMENT PLAN (MDR)	A plan indicating progression from exploratory investigations, such as first-in-man studies, feasibility and pilot studies, to confirmatory investigations, such as pivotal clinical investigations and a PMCF with an indication of milestones and a description of potential acceptance criteria. Source: EU 2017/745 (MDR), Annex XIV, part A
CLINICAL EVALUATION (MDR)	A systematic and planned process to continuously generate, collect, analyse and assess the clinical data pertaining to a device in order to verify the safety and performance, including CLINICAL BENEFITS, of the device when used as intended by the manufacturer. Source: EU 2017/745 (MDR), Article 2 (44)
CLINICAL EVIDENCE	Clinical data and CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) results pertaining to a device of a sufficient amount and quality to allow a qualified assessment of whether the device is safe and achieves the intended CLINICAL BENEFIT(S), when used as intended by the manufacturer. Source: EU 2017/745 (MDR), Article 2 (51)); EU 2017/746 (IVDR), Article 2 (36)
CLINICAL INVESTIGATION (MDR)	Any systematic investigation involving one or more human subjects, undertaken to assess the safety or performance of a device. Source: EU 2017/745 (MDR), Article 2 (45)
CLINICAL PERFORMANCE	Article 2 (52) MDR defines clinical performance as the ability of a device, resulting from any direct or indirect medical effects which stem from its technical or functional characteristics, including diagnostic characteristics, to achieve its intended purpose as claimed by the manufacturer, thereby leading to a CLINICAL BENEFIT for patients, when used as intended by the manufacturer; whereas Article 2 (41) IVDR defines clinical performance as the ability of a device to yield results that are correlated with a particular clinical condition or a physiological or pathological process or state in accordance with the target population and intended user. Source: EU 2017/745 (MDR), Article 2 (52); EU 2017/746 (IVDR), Article 2 (41)
CURATED DATABASE/ CURATED REGISTRY	For the purpose of this document, a curated database/curated registry is any kind of structured repository such as a traditional database, an ontology or an XML file, that is created and updated with a great deal of human effort through the consultation, verification, and aggregation of existing sources, and the interpretation of new (often experimentally obtained) raw data.
GENERALISABILITY	Generalisability refers to the ability of a MDSW to extend the intended performance tested on a specified set of data to the broader intended population.

HUMAN FACTORS ENGINEERING	Human factors engineering refers to the application of knowledge about human behaviour, abilities, limitations, and other characteristics to the design of and interactions with a MDSW to achieve adequate USABILITY.
PERFORMANCE EVALUATION (IVDR)	An assessment and analysis of data to establish or verify the SCIENTIFIC VALIDITY, the ANALYTICAL and, where applicable, the CLINICAL PERFORMANCE of a device. Source: EU 2017/746 (IVDR), Article 2 (44)
PERFORMANCE STUDY (IVDR)	A study undertaken to establish or confirm the analytical or CLINICAL PERFORMANCE of a device. Source: EU 2017/746 (IVDR), Article 2 (42)
REAL-WORLD PERFORMANCE	Information on real-world device use and performance from a wider patient population than a controlled study. Source: Definition derived from IMDRF/SaMD WG/N41FINAL:2017
STATE-OF-THE-ART	Developed stage of current technical capability and/or accepted clinical practice in regard to products, processes and patient management, based on the relevant consolidated findings of science, technology and experience. *Note: The STATE-OF-THE-ART embodies what is currently and generally accepted as good practice in technology and medicine. The state-of-the-art does not necessarily imply the most technologically advanced solution. The STATE-OF-THE-ART described here is sometimes referred to as the "generally acknowledged STATE-OF-THE-ART"* Source: Modified from IMDRF/GRRP WG/N47 FINAL:2018
TECHNICAL PERFORMANCE (MDR) /ANALYTICAL (IVDR)) PERFORMANCE	Capability of a MDSW to accurately and reliably generate the intended technical/analytical output from the input data. Source: IMDRF/SaMD WG/N41FINAL:2017 Source: EU 2017/746 (IVDR) Article 2 (40)
USABILITY	For the purpose of this document, usability refers to the characteristic of the user interface that establishes effectiveness, efficiency and ease of user learning and user satisfaction.
VALID CLINICAL ASSOCIATION (MDR) / SCIENTIFIC VALIDITY (IVDR)	Means the association of an MDSW output with a clinical condition or physiological state. Source: Derived from IMDRF/SaMD WG/N41FINAL:2017 Source: EU 2017/746 (IVDR), Article 2 (38)

4. General principles of the MDSW CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) process

4.1. Introduction

CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) is an ongoing process, conducted throughout the life cycle of a MDSW. It is a structured, transparent, iterative and continuous process which is part of the quality management system for a device. Software that qualifies as a MD or an IVD is subject to the same general CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) principles, laid down in the applicable guidelines and regulatory documents, as other MDs/ IVDs, such as:

- Establishing and maintaining a CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) plan and criteria applied to generate the necessary CLINICAL EVIDENCE based on the characteristics of the device;
- Identification of the relevant data pertaining to performance and/ or safety of the device and any remaining unaddressed issues or gaps in the data;
- Appraisal of the relevant data in terms of quality and its contribution to the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR);
- Analysis of the available data and its relevance with regard to demonstrating conformity with the relevant General Safety and Performance Requirements (GSPRs);
- Documenting the relevant data, their assessment and the CLINICAL EVIDENCE derived therefrom, in the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) report;
- Updating the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) and its documentation throughout the life cycle of the MDSW concerned with data obtained from implementation of the manufacturer's Post Market Clinical Follow-up / Post Market Performance Follow-up (PMCF /PMPF) plan.

These methodological principles are depicted in Figure 1.

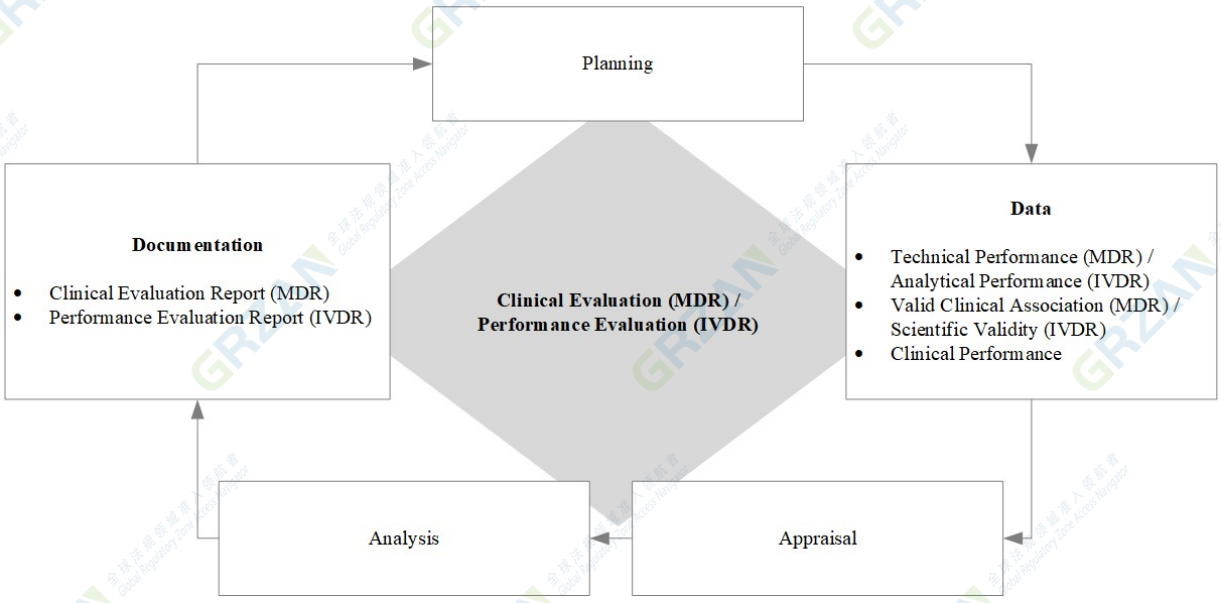


Figure 1 Overview of the stages of the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR)

The requirements for CLINICAL EVALUATION and PERFORMANCE EVALUATION are outlined in Article 61 of the MDR (including Annex XIV) and Article 56 of the IVDR (including Annex XIII), respectively.

While the definition of CLINICAL EVALUATION in the MDR and PERFORMANCE EVALUATION in the IVDR are not identical (see section 0), there is a shared expectation for providing sufficient CLINICAL EVIDENCE to demonstrate conformity with relevant GSPRs under the normal conditions of the device's intended use. CLINICAL EVIDENCE should be sufficient and appropriate in view of the characteristics of the device, clinical risks and its intended purpose. The level of CLINICAL EVIDENCE necessary should be specified and justified by the manufacturer.

Three key components should be taken into account when compiling CLINICAL EVIDENCE for every MDSW (Figure 1), and each is described below in further detail.

VALID CLINICAL ASSOCIATION / SCIENTIFIC VALIDITY is understood as the extent to which, the MDSW's output (e.g. concept, conclusion, calculations) based on the inputs and algorithms selected, is associated with the targeted physiological state or clinical condition. This association should be well founded **or** clinically accepted (e.g. existence of a scientific framework or sufficient level of evidence as further elaborated in section 4.2 of this document). The VALID CLINICAL ASSOCIATION / SCIENTIFIC VALIDITY of a MDSW should demonstrate that it corresponds to the clinical situation, condition, indication or parameter defined in the intended purpose of the MDSW.

NOTE: The VALID CLINICAL ASSOCIATION / SCIENTIFIC VALIDITY seeks to establish that there are sound scientific principles underpinning the use of the MDSW in question. The information provided for the establishment of the VALID CLINICAL ASSOCIATION / SCIENTIFIC VALIDITY should put forward the case that the MDSW has an association with a clinical condition or physiological state. This association may not always be readily established. Thus, the CLINICAL PERFORMANCE can serve as an additional input to the VALID CLINICAL ASSOCIATION/ SCIENTIFIC VALIDITY from a clinical perspective for the specific intended purpose (see Annex I).

Example: MDSW that detects heart arrhythmia by analysing auscultation sound obtained by a digital stethoscope requires demonstrating VALID CLINICAL ASSOCIATION *of the association between abnormal cardiac sounds and heart arrhythmia.*

Evidence supporting VALID CLINICAL ASSOCIATION / SCIENTIFIC VALIDITY can be generated e.g. through literature research, professional guidelines, proof of concept studies, or manufacturer's own clinical investigations/clinical performance studies.

Validation of the **TECHNICAL PERFORMANCE / ANALYTICAL PERFORMANCE** is the demonstration of the MDSW's ability to accurately, reliably and precisely generate the intended output, from the input data.

Evidence supporting TECHNICAL PERFORMANCE / ANALYTICAL PERFORMANCE can be generated through verification and validation activities, e.g. unit-level, integration, and system testing or by generating new evidence through use of curated databases, curated registries, reference databases or use of previously collected patient data.

Validation of the **CLINICAL PERFORMANCE** is the demonstration of a MDSW's ability to yield clinically relevant output in accordance with the intended purpose. The clinical relevance of a MDSW's output is a positive impact

- on the health of an individual expressed in terms of measurable, patient-relevant clinical outcome(s), including outcome(s) related to diagnosis, prediction of risk, prediction of treatment response(s), or

- related to its function, such as that of screening, monitoring, diagnosis or aid to diagnosis of patients, or

- on patient management or public health.

Evidence supporting CLINICAL PERFORMANCE can be generated by testing the MDSW under evaluation, or an equivalent device, in the target population and for the intended use. The applied methodology should be appropriate in light of the device characteristics and intended purpose and may include pre-clinical testing, a clinical investigation or a clinical performance study.

Specifically, for MDSW not claiming CLINICAL BENEFITS that can be specified through measurable, patient-relevant clinical outcome(s), clinically relevant outputs are achieved through demonstrated predictable and reliable use and USABILITY (please refer to section 4.2 of this document).

In addition, CLINICAL EVALUATION or PERFORMANCE EVALUATION of MDSW must consider the benefit-risk ratio in light of the STATE-OF-THE-ART related to practice of medicine for diagnosis, treatment or patient management. It is further expected that the assessment of MDSW considers all components of the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) (see Figure 1 and Annex 0).

The three components described above do not represent a distinct stepwise approach but rather portray a methodological principle for the generation of CLINICAL EVIDENCE.

To determine and justify the level of CLINICAL EVIDENCE, both amount and quality of supporting data should be evaluated. This assessment may be guided by the following non-exhaustive questions:

Sufficient amount

- Does the data support the intended use, indications, target groups, clinical claims and contraindications?
- Have the clinical risks and analytical performance/ clinical performance been investigated?
- Have relevant MDSW's characteristics, such as the data input and output, the applied algorithms or type of interconnection been considered when generating the data to support the performance of the device?
- What is the grade of innovation/ history on the market (how big is the body of scientific evidence)?
- Other, as applicable.

Sufficient quality

- Were the type and the design of the study/ test appropriate to meet the research objectives?
- Was the data set appropriate and actual (state of the art)?
- Was the statistical approach appropriate to reach a valid conclusion?
- Were all ethical, legal and regulatory considerations/ requirements taken into account?
- Is there any conflict of interest?

- Other, as applicable.

4.2. Determination of the valid clinical association / scientific validity

In the first step, the manufacturer should verify the association between the output of the MDSW (based on the inputs and algorithms selected) and the targeted physiological/ clinical condition, clinical situation or clinical parameter, as defined in the intended purpose of the MDSW. MDSW may include a multitude of clinical features governed by its intended purpose which require individual assessment.

This association should be clinically accepted or well founded, which means accepted by the broad medical community and/or described in scientific (peer-reviewed) literature.

VALID CLINICAL ASSOCIATION/ SCIENTIFIC VALIDITY can be demonstrated through the use of existing CLINICAL PERFORMANCE DATA while taking into account the generally acknowledged STATE-OF-THE-ART.

VALID CLINICAL ASSOCIATION / SCIENTIFIC VALIDITY may further be demonstrated by the creation of new CLINICAL PERFORMANCE DATA in the cases where existing data is not sufficient. For example, as a result of a gap analysis, the manufacturer could conclude that additional data may be required.

Examples of existing data (in no particular order)

- Technical standards
- Professional medical society guidelines
- Systematic scientific literature review
- CLINICAL INVESTIGATIONS/ CLINICAL PERFORMANCE STUDIES
- Published CLINICAL DATA (e.g. Summary of Safety and Clinical Performance (SSCP) / Summary of Safety and Performance (SSP), Registries and databases from authorities)

Examples of generating new evidence (in no particular order)

- Secondary data analysis (Analysis of real-world data)
- Perform CLINICAL INVESTIGATION / CLINICAL PERFORMANCE STUDY

4.3. Technical Performance /Analytical Performance

The manufacturer should verify that the MDSW reliably, accurately and consistently meets the intended purpose in real-world usage.

The relevant performance characteristics, as part of the GSPRs and linked to the analytical and / or clinical features, should be supported by evidence generated during verification and validation activities as part of good manufacturing practices for software, or by generating new evidence through the use of curated databases, curated registries, reference databases or use of previously collected patient data.

TECHNICAL PERFORMANCE / ANALYTICAL PERFORMANCE is confirmed by the examination and provision of objective evidence that the MDSW specifications conform to user needs and intended uses, and that the requirements implemented can be consistently fulfilled.[7]

For example, performance verification and validation in the intended computing[8] and use environments[9,10] can be characterised by the demonstration of

- availability,
- confidentiality,
- integrity,
- reliability,
- accuracy (resulting from trueness and precision),
- analytical sensitivity,
- limit of detection,
- limit of quantitation,
- analytical specificity,
- linearity,
- cut-off value(s),
- measuring interval (range),
- GENERALISABILITY,
- expected data rate or quality,
- absence of inacceptable cybersecurity vulnerabilities,
- HUMAN FACTORS ENGINEERING.

Identification of gaps during the validation of the TECHNICAL PERFORMANCE /ANALYTICAL PERFORMANCE could require generation of new evidence, for example, to demonstrate generalisability with real-life datasets or to extend the usability evaluation to omitted user groups.

4.4. Clinical Performance

For the validation of a MDSW's CLINICAL PERFORMANCE, the manufacturer should demonstrate that the MDSW has been tested for the intended use(s), target population(s), use condition(s), operating- and use environment(s) and with all intended user group(s). Section 4.1 of this document further provides context that validation of CLINICAL PERFORMANCE includes the assessment of clinical safety, effectiveness, performance and can support the demonstration of CLINICAL BENEFIT. Validation of the CLINICAL PERFORMANCE should be considered at each change of the software to a new release. If no validation is performed, a justification should be stated in the technical documentation.

With a validation of CLINICAL PERFORMANCE, it is demonstrated that users can achieve clinically relevant outputs through predictable and reliable use of the MDSW.

[7] Derived from Source: GHTF/SG3/N18:2010.
[8] Computing environment: e.g., hardware, memory size, processing unit, time zone, network infrastructure) under which the software is to perform.
[9] Use environment: actual conditions and setting in which users interact with the medical device.
[10] Example on operating environments with distinct requirements are cloud or remote networks.

The manufacturer should consider the intended use(s), indication(s), desired clinical output(s) expressed as claims, leading to expected CLINICAL BENEFITs as part of the CLINICAL PERFORMANCE validation.

A MDSW may have multiple features with only some features claiming a specific CLINICAL BENEFIT. CLINICAL PERFORMANCE is only applicable to those features. Since MDSW can be modular in nature, validation of the CLINICAL PERFORMANCE is also permissible on module level when the functionality of the modules is independent of the other modules. This would allow the confirmation of a continuous benefit / risk acceptability only for the MDSW modules that have changed. In cases where the final combination of modules changes product indications and intended purposes, the performance of that final product configuration should also be evaluated. Validation of the CLINICAL PERFORMANCE can be characterised by the demonstration of applicable CLINICAL DATA to the MDSW in question, such as (non-exhaustive):

- clinical/ diagnostic sensitivity,
- clinical/ diagnostic specificity,
- positive predictive value,
- negative predictive value,
- number needed to treat (average number of patients that need to be diagnosed/ treated in order to have an impact on one person),
- number needed to harm (number of patients that need to be diagnosed/ treated in order have an adverse effect on one patient),
- positive likelihood ratio,
- negative likelihood ratio,
- odds ratio,
- USABILITY/ user interface,
- confidence interval(s).

CLINICAL DATA can be obtained by one or multiple methods such as those referred to in GHTF/SG5/N7:2012 and IMDRF/SaMD WG/N41FINAL:2017.

In addition to the considerations above, CLINICAL EVALUATION of class III and implantable devices (MDR), shall include data from a CLINICAL INVESTIGATION unless the conditions of Article 61(4), (5) or (6) of the MDR have been fulfilled.

For MDSW falling under the IVDR, the evaluation of clinical performance requires the carrying out of clinical performance studies regardless of the classification of the device, unless due justification is provided for relying on other sources of clinical performance data.

Relevant common specifications should be taken into account.

4.4.1. Clinical investigations and clinical performance studies

The practical and achievable benefits of a CLINICAL INVESTIGATION / CLINICAL PERFORMANCE STUDY should be considered as part of determining what data are needed for demonstrating the safety and performance of a new or modified MDSW. The investigation or study should account for potential risks, should follow appropriate ethical requirements, and should be compliant with all relevant legal and regulatory requirements.

MDSW has specific characteristics that should be considered when setting up a clinical investigation or clinical performance study. If the MDSW is used for the determination of a patient's future state (e.g. predisposition, prognosis, prediction) or if the output of the MDSW impacts clinical outcomes (e.g. treatment efficacy) or patient management decisions, then a prospective study may be required as part of the device's CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR). In other situations, retrospective analysis may be more appropriate to generate the necessary data to support compliance with the GSPRs, as there is no impact on patient management and the research does not introduce any risks to the patients. Such an approach is only possible under condition that there is an adequate access to data sets of sufficient amount and quality and obtained from the target population.

Formal requirements of MDR Articles 62 (1), 74 and 82 need to be met as far as appropriate for pre-market retrospective studies of MDSW falling under the MDR.

4.4.2. Where demonstration of conformity based on clinical data is not deemed appropriate

In line with the provisions of MDR Article 61 (1) and IVDR Article 56(1), the level of CLINICAL EVIDENCE required should be appropriate in view of the device claims and characteristics. For medical devices, where the demonstration of conformity with GSPRs based on clinical data is not deemed appropriate (MDR Article 61 (10)), the manufacturer shall duly substantiate in the technical documentation why it is adequate to demonstrate conformity based on the results of non-clinical testing methods alone, including PERFORMANCE EVALUATION, bench testing and preclinical evaluation, and USABILITY assessment.

The justification must be based on the output of the risk management process. This should include an evaluation of clinical STATE-OF-THE-ART, including alternative diagnostic and treatment options, including those identified from literature, and an appraisal of their relevance to the device under evaluation. The device / body interaction, the CLINICAL PERFORMANCEs intended, and the claims of the manufacturer should be specifically considered.

A CLINICAL EVALUATION (MDR) is still required, and the above information and evidence-based justification should be presented in the clinical evaluation report.

Similarly for IVDs, where due to specific device characteristics, demonstration of conformity with GSPRs based on clinical data is not deemed appropriate, a PERFORMANCE EVALUATION (IVDR) is still required and a justification shall be provided and documented in the Performance Evaluation Plan and the corresponding Performance Evaluation Report.

4.5. Final analysis and conclusion of the clinical evaluation (MDR) / performance evaluation (IVDR)

The manufacturer should compile evidence, perform the benefit-risk analysis and document the CLINICAL or PERFORMANCE EVALUATION and its output in the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) report.

4.6. Continuous update of the clinical evaluation (MDR) / performance evaluation (IVDR)

The safety, effectiveness and performance of the MDSW should be actively and continuously monitored by the manufacturer.

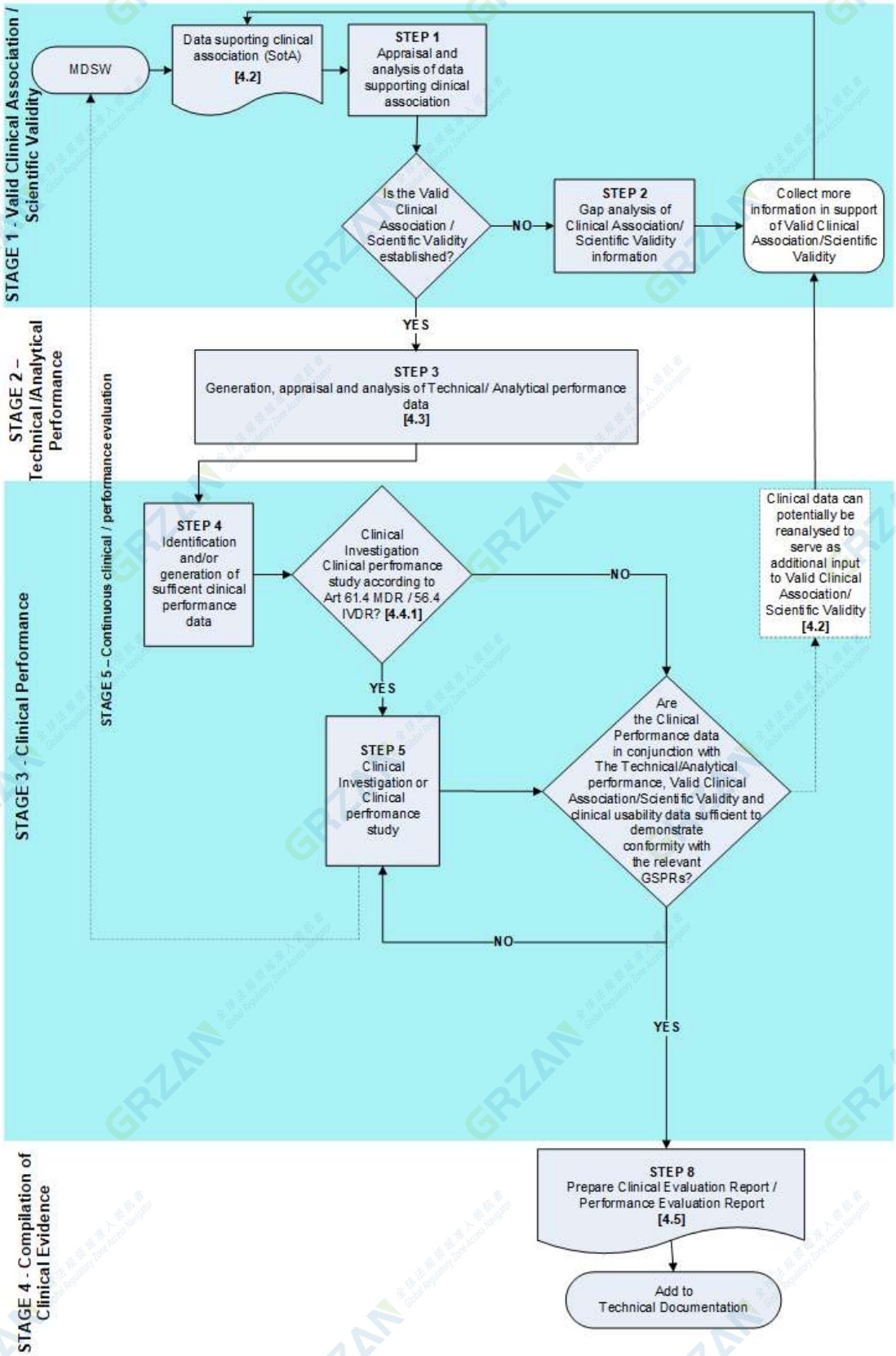
Such data may include, but is not limited to post-market information such as complaints, PMCF/ PMPF data, REAL-WORLD PERFORMANCE data, direct end-user feedback or newly published research / guidelines and should be subject to the CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) principles depicted in Figure 1.

The unique level of connectivity of MDSW facilitates access to REAL-WORLD PERFORMANCE data, which can be used for multiple purposes, including, but not limited to

- timely detection and correction of malfunctions;
- detection of systematic misuse;
- understanding user interactions;
- to conduct ongoing monitoring of CLINICAL PERFORMANCE;
- to improve effectiveness;
- develop the claims in the CLINICAL DEVELOPMENT PLAN (MDR) or future releases.

MDSW can be released for CE marking with initially claimed and validated CLINICAL BENEFITS. Monitoring of REAL-WORLD PERFORMANCE data can help formulate hypotheses about future MDSW functionalities and intended use(s).

Annex I – Methodological principle for generation of CLINICAL EVIDENCE



Annex II – Examples of CLINICAL EVALUATION (MDR) / PERFORMANCE EVALUATION (IVDR) strategies

The high-level examples provided here are for guidance purposes only and aim to provide general indications on how to develop a CLINICAL EVALUATION / PERFORMANCE EVALUATION strategy. The strategy presented in each example is **not a confirmation** of the pathway for a CLINICAL EVALUATION / PERFORMANCE EVALUATION of the device, as other factors need to be considered.

Moreover, the proposed pathway reflects the specific intended purpose, or the healthcare context or situation, in which the device is used as described in the example itself. Any change to the intended purpose or the healthcare context / situation in which that same device is used might result in a different approach.

Data source	**Examples**
Peer-reviewed, relevant scientific literature	- Existing data from studies conducted with the subject device or equivalent device
CLINICAL INVESTIGATION/ CLINICAL PERFORMANCE STUDIES	- Prospective or retrospective studies - Existing manufacturer data - Data from equivalent devices - Data from curated databases/registries/reference databases - Data from outside the EU with justification on applicability
Published experience gained by routine diagnostic testing	- REAL-WORLD PERFORMANCE DATA - Data obtained from PMPF/ PMCF

a) MDSW intended to analyse sleep quality data

An independent MDSW intended to take into account accelerometer and microphone data to determine quality of sleep and to estimate the expected success rate of CPAP (continuous positive airway pressure) treatment for sleep apnoea.

The Manufacturer claims that the MDSW

- determines the quality of sleep that impacts the general well-being.
- monitors quality of sleep in patients with sleep disorders such as sleep apnoea (using phone sensors/wearable devices)
- estimates the expected success rate of CPAP therapy.

Valid Clinical Association

To establish VALID CLINICAL ASSOCIATION, review literature.

- Objective quality of sleep is measured by sleep duration, efficiency and fragmentation. It is further well-established that quality of sleep impacts general well-being such as concentration, risk-factors for cardiovascular disease, mood, cognitive abilities, etc.

- It is not well-established that the success of CPAP therapy can be predicted by monitoring the quality of sleep.
- Address the association of accelerometer and microphone data to established quality of sleep parameters (e.g. sleep duration, efficiency and fragmentation).

The VALID CLINICAL ASSOCIATION has been not established without gaps for prediction of success of CPAP therapy, which requires generation of missing clinical data.

Technical Performance

- Confirm with verification and validation tests that the app can reliably and reproducibly calculate sleep quality scoring.
- Confirm compatibility between the MDSW and the device equipped with the sensors to ensure data can be utilised in the intended way.

Clinical Performance

- In addition to the USABILITY assessment, the manufacturer would perform a retrospective study on previously obtained data to confirm that success of CPAP therapy can be predicted based on the quality of sleep.

b) MDSW intended for image segmentation

An independent MDSW intended to allow automatic detection of organs and anatomical structures (such as the aorta) in CT scans with the accuracy of a radiologist.

The Manufacturer claims that the MDSW:

- detects abdominal aortic aneurisms on abdominal CT scans,
- detects compression fractures on vertebrae,
- detects liver cysts.

Valid Clinical Association

To establish VALID CLINICAL ASSOCIATION, review literature.

- The normal shape and size of anatomy is well established.
- Segmentation techniques on cross-sectional images correlates well with the actual size and shape.

The VALID CLINICAL ASSOCIATION has been established without gaps identified.

Technical Performance

- Confirm with verification and validation tests the basic technical performance such as display, modification, window levelling of images, measurements including confirmation of accuracy, sensitivity and reliability of the MDSW as per the expected performance.

Clinical Performance

- USABILITY assessment including the intended user groups in conjunction with the VALID CLINICAL ASSOCIATION and validation of TECHNICAL PERFORMANCE results has been determined as sufficient to demonstrate conformity with relevant GSPRs.
- In cases where data is available, a retrospective analysis can be performed. In cases where data does not represent the variability of input parameters, for the CLINICAL PERFORMANCE of the segmentation algorithm, the missing data could be generated in a prospective CLINICAL INVESTIGATION.

c) MDSW intended to detect inflammatory bowel diseases (IBD)

Self-testing independent MDSW intended for the semi-quantitative detection of calprotectin from a faecal sample. Reagents are added to the sample resulting in a colour change. The sample is then photographed on a smartphone, and the image is evaluated by an MDSW application (app) running on the phone. The MDSW app detects the colour change in the sample and interprets the concentration of calprotectin. The test is intended as an aid in monitoring and staging of patients with inflammatory bowel disease (IBD).

Manufacturer's claims that the MDSW app

- aids in monitoring and staging the disease level of patients with inflammatory bowel diseases (IBD).
- aids in differentiation between IBD and functional bowel disorders.
- helps patients avoid unnecessary clinical visits.

Scientific Validity

To establish SCIENTIFIC VALIDITY, review literature.

- The SCIENTIFIC VALIDITY could address how the calprotectin level corresponds to the IBD level and stages. Furthermore, it should address, whether calprotectin levels are suitable to differentiate between IBD and functional bowel disorders.
- It is well-established that calprotectin concentration in faecal matter can be reliably measured in test strips by change of colour.
- The colour intensity is directly representative of the concentration of calprotectin.

Analytical Performance

- Confirm the MDSW app can detect reliably and accurately the colour of the test strip compared to human observation, taking into account environmental factors.

Clinical Performance

- The manufacturer should assess the initial performance and feasibility by creating CLINICAL PERFORMANCE metrics, taking into account sensitivity, specificity and confidence intervals.
- Any claims regarding CLINICAL BENEFIT should be supported by sufficient clinical performance data.
- USABILITY should be confirmed by the manufacturer.

d) Active devices containing MDSW to enable their intended purpose

Active devices, such as diagnostic or therapeutic devices, that include MDSW which drives the device in a way that, without the software it would not be able to fulfil its intended purpose. This software does not perform a medical purpose on its own.

The CLINICAL EVALUATION of the MDSW should not be performed independently but should be performed together with the driven device.

e) MDSW which provides an additional user-interface to control an insulin pump

A MDSW intended to virtualise controls of an insulin pump additionally on a smartphone app by connecting to it.

As the software is driving the insulin pump, it is not performing a medical purpose on its own, nor is it creating information on its own for medical purposes.

The CLINICAL EVALUATION of the MDSW app should not be performed independently but should be performed together with the driven insulin pump.

f) MDSW intended to analyse exhaled CO2 in a life-sustaining device in order to control ventilator settings

The MDSW uses physiological data of the patient (e.g. exhaled CO_2, blood oxygen saturation) to control a ventilation device (e.g. frequency, volume and pressure).

The MDSW allows the device to maintain the pre-set value at a desired target (defined by the clinician) without periodic user adjustments needed. This MDSW is part of a closed-loop system.

The CLINICAL EVALUATION should not be limited to the MDSW and should include pre-clinical and clinical investigations, encompassing the entire closed-loop system.