

Clinical Evidence Requirements for CE certification under the *In-Vitro* Diagnostic Regulation in the European Union



Disclaimer & Copyright

This document was prepared by MedTech Europe based on information collected from both MedTech Europe members and the public domain. While MedTech Europe considers the information herein to be reliable it makes no warranty or representation as to its accuracy, completeness or correctness.

This document is intended for informational purpose only and should not be construed as regulatory, legal, medical or otherwise technical or expert advice for any individual circumstances or specific facts. MedTech Europe reserves the right to change or amend the overview at any time in order to keep the information up-to-date.

This document is owned by MedTech Europe, and all rights are reserved. This document can be downloaded, saved, hosted, added to online repositories, and transferred, all for non-commercial uses, provided the document is appropriately cited and acknowledged. For the avoidance of doubt, this document may be used as is, but it may not be sold or rented to any third party, and it may not be modified without express consent by MedTech Europe.

Clinical Evidence Requirements for CE certification under the *in-vitro* Diagnostic Regulation in the European Union

First Edition, May 2020

Many people contributed to this work, through group discussions, advices and reviews. We are acknowledging their input and engagement to develop the *in-vitro* diagnostics sector.

We would like to particularly mention following participants who were heavily involved in drafting this eBook:

Adams Neil	Masson Christine
Astola Annika	Mescalchin Alessandra
Bruinsma Anne Marie	Meyerovich Kira
Callaerts Geert	Nguyen Van Doan
Cheillan Frank	Ons Benny
Choudhary Mayank	Percivati Stefania
Ekholm Pettersson Frida	Petruschke Thorsten
Facheris Luisa	Rodrigues Cardoso Rute
Forssten Camilla	Rousseau Els
Franzen Volker	Saunders Richard
Gazin Muriel	Steenhuis Pieter
Giroud Claude	Sweeney Maranna
Homann Anke	Timonen Anne
Hughes Karin	Van den Eede Peter
Hughes Richard	Wettmarshausen Sascha
Lindroos Hanne	Wevelsiep Anja
Magana Laura	Zaugg Christian
Malcus Carine	

Table of Contents

Introduction	3
CHAPTER 1 - 'Intended Purpose/Use'	4
Appendix 1.1: Examples of intended purposes/uses	10
CHAPTER 2 - Analytical and clinical performance indicators.....	16
CHAPTER 3 – Scientific Validity, Clinical Benefit and Clinical Utility	22
Appendix 3.1: Examples of Clinical Benefit Assessments (according to the IVDR Article 2 (37) and Recital 64)	26
Appendix 3.2: Clinical benefit concept under the IVDR and its distinction from clinical utility	28
CHAPTER 4 – Clinical Evidence Levels	30
CHAPTER 5 - How to demonstrate evidence gained from 'published/documenting routine testing'	39
CHAPTER 6 – Equivalence and similarity concepts in the IVDR	41
CHAPTER 7 – Companion Diagnostics	48
CHAPTER 8 – Documentation of Performance Evaluation requirements	61
CHAPTER 9 - Summary of Safety and Performance	66
CHAPTER 10 – Post-Market Performance Follow-up	70
Appendix 10.1: Post-market Performance Follow-up Plan	76
Appendix 10.2: Post-market Performance Follow-up Report	80

*Note: Chapters: **Benefit-risk determination, State of the Art, Near-Patient Testing, Clinical Evidence Data outside EU** are in active development and will be released in the following Editions of this eBook.*

Introduction

A questions and answers guide to performance evaluation requirements of the new EU *In-Vitro* Diagnostic Medical Devices Regulation 2017/746 (IVDR)

Medical technologies are tightly regulated in the European Union. Before any medical technology can be legally placed on the EU market, a manufacturer must comply with the requirements of all applicable EU legislation and affix a CE mark to their product. Since the 1990s, in-vitro diagnostics (IVDs) have been regulated by an EU Directive (IVD Directive 98/79/EC). In May 2017, the In-Vitro Diagnostic Medical Devices Regulation (EU) 2017/746 (IVDR) was published. MedTech Europe, the European trade association representing the IVD industry, is working with our members and the authorities to support companies in complying with the new IVDR by the end of the transition period (2022).

The IVDR contains several provisions that are open to more than one interpretation. This brochure is designed to help stakeholders understand the new Regulation and the important changes it will bring. Where appropriate, information is presented in a Q&A format to make the text as accessible as possible. It reflects MedTech Europe's best efforts to interpret the IVDR. While we have invested considerable time and effort in developing this document, MedTech Europe does not assert that these opinions and advice are correct and accepts no legal responsibility for them. Specific legal advice should be sought before acting on any of the topics covered in this brochure. Readers should be reminded that it is ultimately for the courts to interpret legislation.

CHAPTER 1 - 'Intended Purpose/Use'

1) How is the term 'intended purpose' defined in the IVDR and how has it changed from the IVD Directive (IVDD)?

The IVDD defines 'intended purpose' as the use for which the device is intended, according to the data supplied by the manufacturer on the labelling, in the instructions for use and/or in promotional materials.

IVDD Article 1(2), (h)

The IVD Regulation defines 'intended purpose' as the use for which a device is intended according to the data supplied by the manufacturer on the label, in the instructions for use or in promotional **or sales materials or statements or as specified by the manufacturer in the performance evaluation.**

IVDR Article 2 (12)

The new element '*as specified by the manufacturer in the performance evaluation*' is the decisive difference between IVDD and IVDR.

2) Where can I find a detailed description of 'intended purpose' in the IVDR?

Descriptions of 'intended purpose' can be found in the instructions for use section in Annex I, as well as in the device description section in Annex II.

IVDR Annex I, Chapter III, section 20.4.1 (c)	
(i)	What is detected and/or measured;
(ii)	The device's function (e.g. screening, monitoring, diagnosis or aid to diagnosis, prognosis, prediction, companion diagnostic);
(iii)	The specific information that is intended to be provided in the context of: <ul style="list-style-type: none"> - a physiological or pathological state; - congenital physical or mental impairments; - the predisposition to a medical condition or a disease; - the determination of the safety and compatibility with potential recipients; - the prediction of treatment response or reactions; - the definition or monitoring of therapeutic measures;
(iv)	Whether it is automated or not;
(v)	Whether it is qualitative, semi-quantitative or quantitative;
(vi)	The type of specimen(s) required;
(vii)	Where applicable, the testing population;
(viii)	For companion diagnostics, the International Non-proprietary Name (INN) of the associated medicinal product for which it is a companion test.

Table 1: Components of device's intended purpose

Most of these elements are repeated in the 'device description' section of the technical documentation in Annex II. But it is notable that for the three specific elements, the wording is different, or the corresponding element can be found elsewhere in Annex I, Chapter III.

<p><i>IVDR Annex I, Chapter III, section 20.4.1 'The instruction for use shall' contain all of the following particulars'</i></p> <p><i>(c) the device's intended purpose</i></p>	<p><i>IVDR Annex II, 1.1 'Device description and specification'</i></p> <p><i>(c) 'the intended purpose of the device which may¹⁾ include information on'</i></p>
---	--

* According to the foreword to all ISO Standards (<https://www.iso.org/foreword-supplementary-information.html>)

- "shall" indicates a requirement
- "should" indicates a recommendation
- "may" is used to indicate that something is permitted

<p>(i) The specific information that is intended to be provided in the context of:</p> <ul style="list-style-type: none"> - a physiological or pathological state; - congenital physical or mental impairments; - the predisposition to a medical condition or a disease; - the determination of the safety and compatibility with potential recipients; - the prediction of treatment response or reactions; - the definition or monitoring of therapeutic measures; <p><i>IVDR Annex I, Chapter III, section 20.4.1 (c)</i></p>	<p>(iii) The specific disorder, condition or risk factor of interest that it is intended to detect, define or differentiate</p> <p><i>Annex II, 1.1 (c) ‘the intended purpose of the device which may include information on’</i></p>
<p>The intended user, as appropriate (e.g. self-testing, near patient and laboratory professional use, healthcare professionals);</p> <p><i>Annex I, Chapter III, 20.4.1 (e)</i></p>	<p>(viii) The intended user</p> <p><i>Annex II, 1.1 (c) ‘the intended purpose of the device which may include information on’</i></p>
<p>For companion diagnostics, the International Non-proprietary Name (INN) of the associated medicinal product for which it is a companion test.</p> <p><i>IVDR Annex I, Chapter III, section 20.4.1 (c)</i></p>	<p>(ix) For companion diagnostics, the relevant target population and the associated medicinal product(s)</p> <p><i>Annex II, 1.1 (c) ‘the intended purpose of the device which may include information on’</i></p>

Table 2: Comparative table between the ‘intended purpose’ requirements of Annex I and Annex II

3) The terms ‘intended purpose’ and ‘intended use’ are both used in the IVDR. Is there any difference in the meaning of the terms?

Unlike the term ‘intended purpose’, the term ‘intended use’ is not explicitly defined in the IVDR. However, the term ‘intended use’ is used several times throughout the Regulation.

This implies that it should not be understood differently from the term ‘intended purpose’

For example:

- Devices shall be designed, manufactured and packaged in such a way that their characteristics and performance during their intended use are not adversely affected during transport (...) *Annex I, Chapter I, section 7*
- The characteristics and performances of the device shall be specifically checked if they may be affected when the device is used for the intended use under normal conditions (...) *Annex I, Chapter I, section 9 (4)*

- The notified body's assessment of performance evaluations as referred to in Annex XIII shall cover 'the intended use specified by the manufacturer and claims for the device defined by it (...) Annex VII, section 4.5.4

Both intended purpose and intended use appear in the chapter on performance evaluation plans, stating that both should be specified:

As a rule, the performance evaluation plan shall include at least:

- a specification of the intended purpose of the device (...)
- a specification of the intended use of the device (Annex XIII 1.1)

4) What is the global view on the terms 'intended purpose' and 'intended use'? Are they used interchangeably? How does the global view of both terms impact the IVDR interpretations?

Analysis of the following international documents shows that 'intended use' is a synonym for 'intended purpose' and is used interchangeably. This has an important influence on the IVDR which explicitly emphasises in recital 5 that international guidance documents from GHTF/ IMDRF should be considered to promote global convergence.

For example:

- GHTF/SG1/N045:2008³ Principles of *In Vitro* Diagnostic (IVD) Medical Device Classification
'Intended use/purpose': the objective intent of the manufacturer; the use of a product, process or service as reflected in the specifications, instructions and information provided by the manufacturer (Chapter 4.0 Definitions)
- IMDRF Essential Principles of Safety and Performance of Medical Devices and IVD Medical Devices/January 2018 ⁴
'Intended Use/Intended Purpose': The objective intent of the manufacturer regarding the use of a product, process or services as reflected in the specifications, instructions and information provided by the manufacturer. (GHTF/SG/N77:2012) (Chapter 3.0 Definitions)
- EN ISO 18113-1:2011 *In vitro* diagnostic medical devices. Information supplied by the manufacturer (labelling)⁵. Part 1: Terms, definitions and general requirements (ISO 18113-1:2009, currently under revision)

3.31 **'intended use/intended purpose'**: objective intent of an IVD manufacturer regarding the use of a product, process or service as reflected in the specifications, instructions and information supplied by the IVD manufacturer.

5) How should the ‘intended purpose/use’ elements be presented in the instructions for use?

The instructions for use (IFU) section in Annex I does not specify a mandatory structure/layout. Therefore, how the applicable ‘intended purpose/use’ elements are presented in the IFU depends on the manufacturer’s concept of IFU. For example, these elements may be distributed over several sections or combined in one. If they are not combined, it may be helpful to describe where the applicable elements can be found, for audit purposes.

Annex I, Chapter III, section 20.4.

6) What is the relationship between a product’s ‘intended purpose/use’ and a ‘product claim’?

A device-specific intended purpose, as indicated in the IFU and labelling, serves as the basis for all product claims.

The manufacturer is prohibited from misleading the user or the patient through a product claim (e.g. text, names, pictures, figurative or other signs appearing on the label, in the IFU, or in promotional or sales materials) about the device’s ‘intended purpose/use’, safety and performance.

IVDR Article 7

7) How is the ‘intended purpose/use’ linked to the concept of clinical evidence?

The ‘intended purpose/use’ is fundamental to the building of the performance evaluation plan and includes information such as:

- What is detected and/or measured
- It’s function (see Table 1)
- The specific information set out in Table 1 and 2.

Therefore, the ‘intended purpose/use’ directly drives the level of performance evaluation, performance studies and post-market performance follow-up activities.

Annex I, Chapter III, section 20.4.1c; Annex II 1.1.c; Annex XIII Part A and B

It is the manufacturer’s solely responsibility to define a concept appropriate clinical evidence based on the ‘intended purpose/use’ and the environment where the product is used.

For more information about different levels of clinical evidence, see the [CHAPTER 4 – Clinical Evidence Levels](#). See below for a non-exhaustive list of examples ([Appendix 1.1: Examples of intended purposes/uses](#)).

References:

1. Directive 98/79/EC of the European parliament and of the council of October 27, 1998 on *in vitro* diagnostic medical devices
2. Regulation (EU) 2017/746 of the European parliament and of the council of April 5, 2017 on *in vitro* diagnostic medical devices
3. GHTF/SG1/N045:2008 Principles of *In Vitro* Diagnostic (IVD) Medical Device Classification
4. IMDRF Essential principles v 2017 GHTF/SG1/N77:2012 Principles of Medical Device Classification
5. EN ISO 18113-1:2011 *In vitro* diagnostic medical devices. Information supplied by the manufacturer (labelling). Part 1: Terms, definitions and general requirements (ISO 18113-1:2009)

Appendix 1.1: Examples of intended purposes/uses

Example 1: IVD device intended to detect magnesium

	Intended Purpose/ Intended Use	Scientific Validity	Analytical Performance	Clinical Performance Options
Physiological state	To detect and measure magnesium to assess electrolyte/magnesium homeostasis.	Mg ²⁺ is a cofactor of many enzyme systems, required by all ATP-dependent enzymatic reactions. It functions as an activator for various physiochemical processes, including phosphorylation, protein synthesis, and DNA metabolism. It is also involved in neuromuscular conduction and excitability of skeletal and cardiac muscle.	Quantitative determination of magnesium concentration in human serum, plasma, and urine with appropriate analytical sensitivity, specificity, precision, etc.	Agreement with other measures of magnesium (method comparison), standardised against atomic absorption spectrometry.
Clinical condition	To detect and measure magnesium to detect clinical conditions associated with abnormal magnesium levels in the body, hyper/hypomagnesemia.	<ul style="list-style-type: none"> - Increased serum magnesium concentrations occur in renal failure, acute diabetic acidosis, dehydration, or Addison's disease. - Hypomagnesemia may be observed in inherited disorders of isolated magnesium malabsorption, chronic alcoholism, malabsorption, severe diarrhoea, acute pancreatitis, diuretic therapy, hypertension, and kidney disorders such as glomerulonephritis and tubular reabsorption defects. 		Diagnostic/clinical sensitivity and specificity to detect specific clinical conditions e.g. kidney disorders, primary infantile hypomagnesemia, etc.
Clinical condition 'therapy monitoring'	To monitor drugs (e.g. proton pump inhibitors, diuretics, cytotoxic drugs), clinical interventions (e.g. dialysis) known to alter magnesium levels.	Composition of dialysis solution, and monitoring of blood pressure, along with measurement of magnesium concentration, are useful to monitor treatments/interventions known to alter magnesium levels. This supports dose adjustment and avoids adverse effects.		Appropriate diagnostic/clinical sensitivity and specificity to measure and monitor magnesium concentrations to adjust drug dosing and adjust treatment.

Example 2: IVD device intended to detect and measure C-reactive protein (CRP)

	Intended Purpose/ Intended Use	Scientific Validity	Analytical Performance	Clinical Performance Options
Physiological state	To detect and measure C-reactive protein to assess the inflammatory status of the body.	CRP is one of the strongest acute phase reactants and aids in non-specific host defence against infectious agents, rising after myocardial infarction, stress, trauma, infection, inflammation, surgery or neoplastic proliferation.	Quantitative determination of the CRP concentration in human serum, and plasma with appropriate analytical sensitivity, specificity, precision, etc.	Agreement with other assays standardised against reference preparation BCR470/CRM470 (method comparison) or erythrocytes sedimentation rate (ESR)
Clinical condition	To detect and measure C-reactive protein to detect systemic inflammatory processes due to an active disease.	Determination of CRP is clinically useful to screen for organic disease, to assess activity of inflammatory diseases such as rheumatoid arthritis, to detect intercurrent infection in systemic lupus erythematosus, in leukaemia or after surgery.		Diagnostic/clinical sensitivity and specificity to detect specific clinical condition.
Clinical condition 'therapy monitoring	To monitor efficacy of drugs which are known to suppress or prevent inflammatory processes (e.g. ISDs, anti-inflammatory drugs)	Serum CRP is clinically useful to monitor disease activity and detect renal allograft rejection. This supports dose adjustment and avoids adverse effects.		Appropriate diagnostic/clinical sensitivity and specificity to monitor kidney function to adjust drug dosing.

Example 3: IVD device intended to measure Troponin T

	Intended Purpose/ Intended Use	Scientific Validity	Analytical Performance	Clinical Performance Options
Physiological state	n/a		Quantitative determination of the troponin T concentration in human serum, and plasma with appropriate analytical sensitivity, specificity, precision, etc.	
Clinical condition	To determine cardiac troponin T levels in human serum and plasma to detect clinical conditions and risk associated with cardiomyocyte damage.	<p>Determination of troponin T in serum and plasma is useful in diagnosis of AMI/ACS due to the rapid increase of serum/plasma concentration after AMI.</p> <p>It is useful in risk stratification in patients presenting with ACS or cardiac risk in patients with renal disease.</p> <p>Determination of TnT aids in early diagnosis (PoC).</p> <p>Measurement of troponin T in serum and plasma aids in therapy selection in patients with elevated Troponin T levels.</p>		Diagnostic/clinical sensitivity and specificity to detect specific clinical condition, and hazard ratio to assess associated risk.
Clinical condition 'therapy monitoring'	To monitor troponin T levels in patients receiving drugs known to cause cardiac toxicity (such as anthracyclines, multikinase inhibitors, trastuzumab).	Currently, detection and monitoring of cardiac toxicity of cancer therapies are performed by assessment of LVEF using echocardiography, radionuclide ventriculography or MRI. Since a significant amount of myocardial damage is needed to result in a decrease of LVEF, the detection of cardiac toxicity can be delayed, leading to irreversible cardiac damage, late introduction of HF therapy, and suboptimal recovery. Early elevation of cardiac troponins after anthracycline is predictive of chronic cardiac toxicity, and the pattern of this elevation can add prognostic information.		Appropriate diagnostic/clinical sensitivity and specificity to monitor troponin T levels in order to adjust or induce appropriate treatment.

Example 4: IVD device intended to measure glucose in serum, plasma and urine (no PST/ CPS device)

	Intended Purpose/ Intended Use	Scientific Validity	Analytical Performance	Clinical Performance Options
Physiological state	To determine glucose levels in human serum, plasma and urine to assess glucose homeostasis.	Glucose is a breakdown product from carbohydrates and is used as an energy source in most organisms including humans. The concentration of glucose in the blood is regulated by the complex interplay of multiple pathways and is maintained within narrow limits.	Quantitative determination of the glucose concentration in human serum, and plasma with appropriate analytical sensitivity, specificity, precision, etc.	Agreement with other assays standardised against ID/MS (method comparison).
Clinical condition	To determine glucose levels in human serum, plasma and urine to detect clinical conditions associated with abnormal glucose concentrations such as diabetes mellitus.	Determination of glucose in serum, plasma and urine is useful in screening and diagnosis of diabetes. It is an aid in diagnosis of other diseases resulting in altered glucose levels such as insulinoma. Measurement of glucose in urine aids in diagnosis of renal tubular disorders such as Fanconi syndrome or familial renal glucosuria.		Diagnostic/clinical sensitivity and specificity to detect specific clinical condition.
Clinical condition therapy monitoring	To monitor glucose levels in patients receiving blood glucose lowering drugs (such as insulin, and other anti-diabetic drugs).	Measurement of glucose provides an index of short-term glycaemic control. This supports dose adjustment and avoids adverse effects.		Appropriate diagnostic/clinical sensitivity and specificity to monitor glucose homeostasis to adjust drug dosing.

Example 5: IVD device intended to detect oncology tumour marker – KRAS mutation test

	Intended Purpose/ Intended Use	Scientific Validity	Analytical Performance	Clinical Performance Options
Pathological state	To detect gene mutation to assess KRAS mutation status in samples from patients diagnosed with metastatic colorectal cancer.	Somatic mutation in the KRAS gene is an essential step in the development of colorectal cancer.	Qualitative detection of somatic mutations in the KRAS gene using extracted DNA from FFPE samples of CRC with appropriate analytical sensitivity, specificity, precision etc.	Appropriate clinical performance data. For KRAS codons 12 and 13 WHO reference panel NIBSC 16/250 available.
Companion diagnostic	To detect gene mutation to assess KRAS mutation status in samples from patients diagnosed with metastatic colorectal cancer. The test is intended to be used as an aid in the identification of metastatic colorectal cancer patients for whom treatment with drug (INN) may be indicated.	Somatic mutations in the KRAS gene are predictive biomarkers of resistance to human EGFR directed therapies.		Clinical trial to establish the safety and effectiveness of the therapeutic product in the appropriate population based on detection of the KRAS mutation status using the IVD test.

Example 6: IVD device intended as an oncology monitoring assay -BCR -ABL1

	Intended Purpose / Intended Use	Scientific Validity	Analytical Performance	Clinical Performance Options
Pathological state	To measure BCR-ABL1 mRNA p210 transcript levels in patients diagnosed with positive chronic myelogenous leukaemia during monitoring of treatment with Tyrosine Kinase Inhibitors.	The BCR-ABL1 transcript produced by the t (9;22) chromosomal translocation is associated with chronic myelogenous leukaemia. Therapy response in CML is associated with BCR-ABL1/ABL1 transcript levels.	Quantitative detection of BCR-ABL1 transcript using extracted RNA from whole blood with appropriate analytical dataset (sensitivity, specificity, precision etc.)	Appropriate clinical performance data. WHO International standard material for quantitation of BCR-ABL translocation available.
CDx	To measure BCR-ABL1 mRNA p210 transcript levels in patients diagnosed with t (9;22) positive chronic myelogenous leukaemia during monitoring of treatment with Tyrosine Kinase Inhibitors and to be used in the monitoring as an aid in identifying CML patients in the chronic phase being treated with drug (INN) who may be candidates for treatment discontinuation and for monitoring of treatment-free remission.	The BCR-ABL1 transcript produced by the t (9;22) chromosomal translocation is associated with chronic myelogenous leukaemia. Therapy response in CML is associated with BCR-ABL1/ABL1 transcript levels and treatment success is defined by specific transcript levels.		Clinical trial to establish the safety and effectiveness of the therapeutic product (incl. discontinuation of drug) in the appropriate population based on monitoring BCR-ABL1 transcript levels using the IVD test.

CHAPTER 2 - Analytical and clinical performance indicators

Analytical and Clinical Performance as Components of Clinical Evidence

IVDR Article 56 states:

‘(2) - The *clinical evidence* shall support the intended purpose of the device as stated by the manufacturer and be based on a continuous process of performance evaluation, following a performance evaluation plan.

(3) - A performance evaluation shall follow a defined and methodologically sound procedure for the demonstration of the following, in accordance with this Article and with Part A of Annex XIII:

- (a) scientific validity;
- (b) analytical performance;
- (c) clinical performance.

The data and conclusions drawn from the assessment of those elements shall constitute the *clinical evidence* for the device. The *clinical evidence* shall be such as to scientifically demonstrate, by reference to the state of the art in medicine, that the intended clinical benefit(s) will be achieved and that the device is safe. The *clinical evidence* derived from the performance evaluation shall provide scientifically valid assurance, that the relevant general safety and performance requirements set out in Annex I, are fulfilled, under normal conditions of use.’

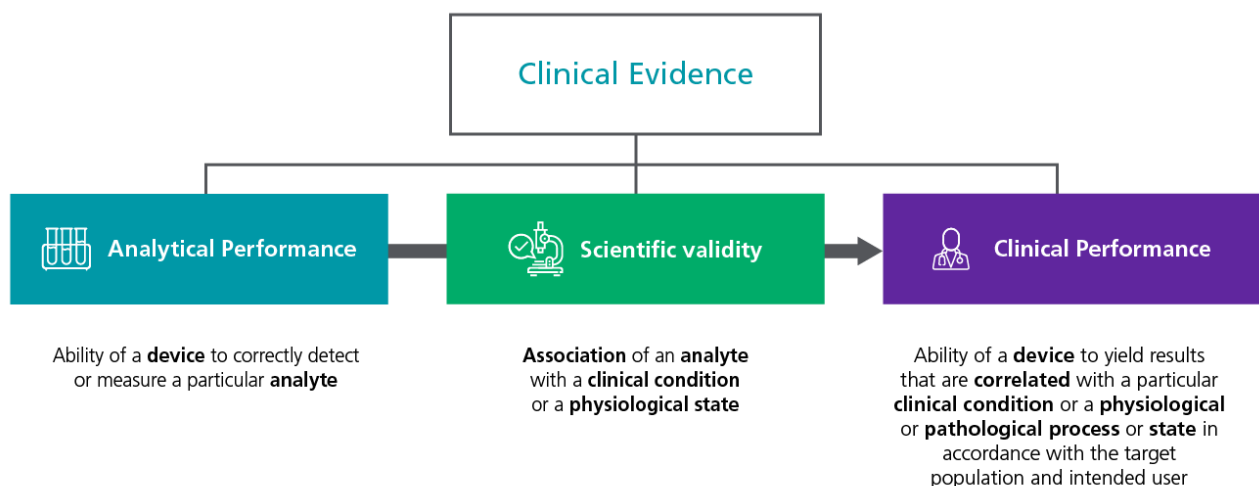


Figure 1. Components of clinical evidence according to IVDR 2017/746/EU

IVD devices shall achieve the performances stated by the manufacturer, and in particular, where applicable:

- (a) The analytical performance, such as, analytical sensitivity, analytical specificity, trueness (bias), precision (repeatability and reproducibility), accuracy (resulting from trueness and precision), limits of detection and quantitation, measurement range, linearity, cut-off, including determination of appropriate criteria for specimen collection and handling and control of known relevant endogenous and exogenous interference, cross-reactions.
- (b) The clinical performance, such as diagnostic sensitivity, diagnostic specificity, positive predictive value, negative predictive value, likelihood ratio, expected values in normal and affected populations.

Annex I, Chapter II, Section 9.1 and Annex II, Section 6.1.

1) What is the conceptual difference between analytical and clinical performance?

- Analytical performance and clinical performance studies have different objectives and endpoints.
- Analytical performance studies focus on the analyte, clinical performance studies focus on the patient.
- Analytical performance is the basis of the clinical performance of a device.
- Analytical performance data do not directly demonstrate the clinical performance of a device as they are assessing different performance characteristics. For example, a high analytical sensitivity does not guarantee acceptable diagnostic sensitivity ².

2) What are the typical indicators of analytical and clinical performance?

Indicators of analytical performance are typically similar or even identical across IVD devices. Guidance is provided by a set of Clinical & Laboratory Standards Institute (CLSI) documents. Conversely, indicators of clinical performance vary and depend strongly on the Intended Purpose. Specifically, the clinical function in the intended purpose/use defines the study endpoint or clinical performance data type, e.g. diagnostic sensitivity and specificity (also described as *clinical* sensitivity and specificity) for a test claiming a diagnostic intended purpose and a hazard ratio for a test claiming prognostic intended purpose (see Table 1 below).

The term “clinical study” by itself, without the specification of analytical or clinical performance study, can be confusing. Specifically, the term “clinical study” is sometimes applied to any study collecting or using of patients’ samples (sometimes called “clinical samples”), independent of the performance indicators. However, an analytical performance study utilising patient samples remains an analytical performance study and is not considered as a source of clinical performance data. The recommendation is, therefore, to use the specific and clearly defined terms such as “analytical performance study” and “clinical performance study”, as opposed to “clinical study”.

Typical Performance Indicators



 Analytical Performance	 Clinical Performance	
<ul style="list-style-type: none"> • Measuring Interval: LoQ as the lower limit and the upper limit of Linearity as the upper limit. • LoB (e.g. CLSI guideline EP17-A2) • LoD (=analytical sensitivity) (e.g. CLSI guideline EP17-A2) • LoQ (e.g. CLSI guideline EP17-A2) • Linearity (e.g. CLSI guideline EP06A) • Precision (repeatability) (e.g. CLSI guideline EP05-A3) • Intermediate Precision (e.g. CLSI guideline EP05-A3) • Reproducibility (e.g. CLSI guideline EP05-A3) • Carryover (e.g. CLSI guideline H26-A2) • Total Analytical Error (Accuracy) (e.g. CLSI guideline EP21-A) • Instrument Comparison (e.g. CLSI guideline EP09-A3) • Method Comparison (e.g. CLSI guideline EP09-A3) • Interfering Substances (=analytical specificity): Could be done by checking known and expected interferences, e.g. from vigilance cases and literature research. 	Intended Purpose	Performance Indicator
	Screening	Diagnostic Sensitivity & Specificity, AUC, or NPV, PPV
	Diagnosis	Diagnostic Sensitivity & Specificity, AUC, or NPV, PPV
	Classification	Agreement table, or Net Reclassification Index (NRI)
	Prognosis	Hazard or Odds Ratio, Kaplan-Meier curves, or C-index
	Disease monitoring	Diagnostic Sensitivity & Specificity, AUC, or NPV, PPV
	Therapy stratification Therapy selection	Outcome measure, e.g. response rate, survival, Hazard ratio, a.o.
	(Patho) physiological function / state	Agreement table
	For all Intended Purposes	Expected values in normal and affected populations

Table 1. Possible examples of analytical and clinical performance indicators based on the intended purpose as referred to in the complementary list of examples ³. For abbreviations please see below.

Box 1: Abbreviations

AUC: Area under the curve
 LoB: Limit of blank
 LoD: Limit of detection
 LoQ: Limit of quantification
 NPV: Negative predictive value
 NRI: Net reclassification index
 PPV: Positive predictive value

Please note that this table does not provide a comprehensive or prescriptive selection of performance indicators. It is the manufacturer's sole responsibility to define an appropriate concept of clinical evidence

Table 2. Examples of different intended purposes/use and how they drive the selection of clinical performance indicators, possible study populations, potential study designs, and IVD device examples.

Intended Purpose	Performance indicator	Study population	Study design	Examples
Screening (early detection of subclinical disease)	Diagnostic sensitivity & specificity (against the "gold standard"/ reference method), AUC , NPV, PPV	Subjects at risk (indicated for screening) Could be population level	Prospective or retrospective observational, longitudinal study (1-arm) or corresponding RWD	Bloodscreening for Infectious Diseases
Diagnosis	Diagnostic sensitivity & specificity (against the "gold standard"/ reference method), AUC , NPV, PPV	Subjects with signs and symptoms of disease	Prospective or retrospective observational cohort study or cross-sectional case-control study	Troponins for AMI
Classification / Grading	Agreement tables, NRI (Net Reclassification Index); if a gold standard available: also Sens/Spec	Subjects diagnosed with the disease of interest	Prospective or retrospective observational study, "case-control" study (cases with different grading)	Creatinine for kidney function / failure
Prognosis /Risk Stratification	Hazard ratio, Odds ratio, Kaplan-Meier curves, C-index, NRI, absolute survival estimates	Depending on IU, population level, or subjects with disease	Prospective or retrospective observational study (Less preferred: case-control study)	CRP, LDL
Disease monitoring	Diagnostic sensitivity & specificity, AUC (against gold standard), NPV, PPV	Diseased patients with or without treatment	Prospective or retrospective observational longitudinal study	Glucose, PSA
Therapy stratification (CDx)	Patient outcome measure and interaction analysis (CDx defined group for therapeutic efficacy and/or safety)	All-comers (all patients under treatment of the drug)	Clinical outcome studyprospective randomized controlled trial (RCT) or retrospective study Concordance (bridging) studies	HER2, BRAF, KRAS
Therapy selection (CDx)	Patient outcome measure and interaction analysis (CDx defined group for therapeutic efficacy and/or safety)	Biomarker-positive patients	Clinical outcome studyprospective RCT or retrospective study Concordance (bridging) studies	BRAF

Diagnostic sensitivity = Clinical sensitivity

Please note that this table does not provide a comprehensive or prescriptive selection of performance indicators, study populations, or study designs. It shows possible options of these clinical evidence concepts. It is the manufacturer's sole responsibility to define an appropriate clinical evidence concept. Furthermore, the demonstration of clinical utility is not a requirement according to the IVDR. A notable exception is the Intended Use of Therapy Prediction (Companion Diagnostic) where a clinical utility study involving the corresponding drug is typically required.

It should be noted that there are various analytical performance guidance and specifications approaches, e.g. standards from the Clinical and Laboratory Standards Institute (CLSI), the Milan performance specifications¹¹, and others. These are established guidelines that could be considered, but it is beyond the scope of this brochure to provide a comprehensive overview.

3) Where should cut-offs be documented?

- IVDR mentions cut-offs under analytical performance. Therefore, cut-offs should be documented in the analytical performance report, unless justified.
- IVDR, Annex II, Section 6.1.2.6. Definition of assay cut-off:
 This Section shall provide a summary of analytical data with a description of the study design including methods for determining the assay cut-off, such as:
 - (a) the population(s) studied: demographics, selection, inclusion and exclusion criteria, number of individuals included;

- (b) method or mode of characterisation of specimens; and
- (c) statistical methods such as Receiver Operator Characteristic (ROC) to generate results and if applicable, define grey zone/equivocal zone.

4) What are the requirements if analytical and/or clinical performance studies are performed externally instead of internally?

- External studies have the same objectives and endpoints as their internal counterparts.
- The level of required documentation is higher for performance evaluation studies, if conducted externally.
- For external studies, manufacturers need to consider a number of additional factors and activities, e.g. number of study sites, site initiation, monitoring, sponsorship, contract an investigator. Depending on the type of study, ethics approval may be needed. For clinical performance studies, see also ISO 20916 'In vitro diagnostic medical devices — Clinical performance studies using specimens from human subjects – Good study practices'
- If testing in an end-user setting (external study) is omitted by the manufacturer, it has to be justified that the internal conditions of use cover the normal conditions of use mentioned in Annex I.
- IVDR Annex I, Section 9.4. 'The characteristics and performances of the device shall be specifically checked in the event that they may be affected when the device is used for the intended use under normal conditions:
 - (a) For devices for self-testing, performances obtained by laypersons;
 - (b) For devices for near-patient testing, performances obtained in relevant environments (for example, patient home, emergency units, ambulances).'
- IVDR Annex XIII, 2.3.1. 'Clinical performance study design type: Clinical performance studies shall be designed in such a way as to maximize the relevance of the data while minimising potential bias.'
- IVDR Article 57. 2. 'Where appropriate, performance studies shall be performed in circumstances similar to the normal conditions of use of the device.'

References:

- 1) Regulation (EU) 2017/ 746 of the European parliament and of the council of April 5, 2017 on *in vitro* diagnostic medical devices
- 2) Saah A J, Hoover D R. "Sensitivity" and "specificity" reconsidered: the meaning of these terms in analytical and diagnostic settings. *Ann Intern Med.* 1997;126:91–94
- 3) [Appendix 1.1: Examples of intended purposes/uses.](#)
- 4) CLSI document EP09-A3: Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline - Second Edition (Interim Revision). Wayne, PA : Clinical and Laboratory Standards Institute, 2013.
- 5) CLSI document EP06A: Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline. Wayne : NCCLS, 2012.
- 6) CLSI document H26-A2: Validation, Verification, and Quality Assurance of Automated Hematology Analysers; Approved Standard - Second Edition. Wayne, PA : Clinical and Laboratory Standards Institute, 2010.
- 7) CLSI document EP17-A2: Protocols for Determination of Limits of Detection and Limits of Quantification; Approved Guideline. Wayne, PA : s.n., 2012.
- 8) CLSI document EP21-A Estimation of Total Analytical Error for Clinical Laboratory Methods. CLSI. 20, s.l. : CLSI, 2012, Vol. 23.
- 9) CLSI document EP24-A2: Assessment of the diagnostic accuracy of laboratory tests using receiver operating curves; approved guideline - second edition, 2011.
- 10) CLSI document EP07: Interference testing in Clinical Chemistry, third edition, 2018

The CLSI numbers and version is valid at the time of publication/revision of this document.

- 11) Sandberg S et al. "Defining analytical performance specifications: Consensus Statement from the 1st Strategic Conference of the European Federation of Clinical Chemistry and Laboratory Medicine" *Clin Chem Lab Med* 2015; 53(6): 833-835

CHAPTER 3 – Scientific Validity, Clinical Benefit and Clinical Utility

Background

Scientific validity is a new term and requirement that has been introduced in the IVD Regulation. It is important to clarify the concept of scientific validity, and its relationship to clinical utility, in order to understand what the responsibilities of the manufacturer are under the Regulation.

1) What is the concept of ‘scientific validity’ and the relationship between ‘scientific validity’ and ‘clinical utility’?

- The IVDR Article 2 (38) defines ‘**scientific validity** of an analyte’ as the association of an analyte with a clinical condition or a physiological state.
- The IVDR does not mention or define **clinical utility**.
- The IMDRF document GHTF/SG5/N6:2012 explains that **scientific validity** is often identified in academic research and is supported by studies evaluating the analyte (measurand) for potential clinical applications. Literature review and, where applicable, feasibility and/or scientific validity studies, will help to establish the potential scientific validity. For many analytes (measurands) the scientific validity is well established; e.g. the scientific validity for calcium (measurand) is well established as being linked to parathyroid disease, a variety of bone diseases, chronic renal disease and tetany. However, some IVD medical devices are developed when the scientific validity of the analyte is still emerging. An example would be a newly characterised biomarker that is potentially useful in monitoring recurrence or progressive disease in patients with cancer.
- In the same IMDRF document, a definition of **clinical utility** is given as: ‘The usefulness of the results obtained from testing with the IVD medical device and the value of the information to the individual being tested and/or the broader population.’
- The IMDRF provides a link between **clinical utility** and **scientific validity** by the following explanation:

Clinical utility of an IVD medical device supports clinical decisions for patient management such as effective treatment or preventive strategies. Clinical utility has been described as including many elements such as acceptability, appropriateness, availability of treatments/interventions, and health economics. Scientific validity and clinical performance are the only elements of clinical utility considered in this document (see APPENDIX I).

In line with the IVDR, a manufacturer is expected to demonstrate clinical evidence, which includes scientific validity, analytical performance and clinical performance, for all IVD medical devices unless any omission can be justified as not applicable. Aside from scientific validity and clinical performance, a manufacturer is not

required to demonstrate any other elements of clinical utility for premarket conformity CE marking assessment purposes. See *APPENDIX II* for a *diagrammatic representation of the manufacturers' responsibilities*.

2) What are the examples of scientific validity and clinical utility?

Scientific validity

A self-testing blood glucose product which measures the amount of glucose in the blood has a **scientific validity** in that glucose levels are associated with diabetes.

Clinical utility

The **clinical utility** of testing the blood for glucose is that if the patient monitors their glucose levels regularly to ensure it remains within the normal range and, as needed, adjusts their insulin levels to keep their blood glucose levels normal, this will have longer-term effects on patient outcomes. It can reduce the potential for damage to the large blood vessels of the heart, brain and legs (called macrovascular complications) and damage to the small blood vessels (microvascular complications) causing problems in the eyes, kidneys, feet and nerves. These complications will cause hospitalisation and further cost to the health service.

3) What are the responsibilities of the manufacturer under the IVD Regulations to provide information on scientific validity to enable a product to be CE marked?

- a. The manufacturer is responsible for demonstrating **scientific validity** as defined in Annex XIII Part A (1.2. (1)) 'Performance evaluation and Performance Studies'.
 - i. As a general methodological principle, the manufacturer shall:
 - I. identify through a systematic scientific literature review the available data relevant to the device and its intended purpose and identify any remaining unaddressed issues or gaps in the data;
 - II. appraise all relevant data by evaluating their suitability for establishing the safety and performance of the device;
 - III. generate any new or additional data necessary to address outstanding issues.
 - ii. The manufacturer shall demonstrate scientific validity based on one or a combination of the following sources:
 - I. relevant information on the scientific validity of devices measuring the same analyte or marker;
 - II. scientific (peer-reviewed) literature;
 - III. consensus expert opinions/positions from relevant professional associations;
 - IV. results from proof of concept studies;
 - V. results from clinical performance studies.

As stated in Article 56 (5) – 'The scientific validity data, their assessment and the clinical evidence derived therefrom shall be documented in the performance evaluation report referred to in Section

1.3.2 of Part A of Annex XIII. The performance evaluation report shall be part of the technical documentation, referred to in Annex II, relating to the device concerned.'

- b. Aside from scientific validity and clinical performance, a manufacturer is not required to demonstrate any other elements of clinical utility for premarket conformity CE marking assessment purposes. See Appendix 1 for a diagrammatic representation of the manufacturers' responsibilities.

4) What is meant by clinical benefit for an IVD device?

- **The IVDR Article 2 (37) defines clinical benefit** as 'the positive impact of a device related to its function, such as that of screening, monitoring, diagnosis of patients, or a positive impact on patient management or public health'.
- In addition, **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 device 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 patients is dependent on further diagnostic and/or therapeutic options which could be available.'
- Thus, the clinical benefit focuses on the 'accurate medical information' output of an IVD device, in context of the intended purpose as defined by the manufacturer and in conjunction with other medical information. The clinical benefit and the corresponding clinical evidence do not include the potential benefits as a result of patient management (i.e. clinical utility; see figure in APPENDIX II).

5) How do manufacturers assess the clinical benefit of their device?

- Annex XIII (1.3.1) of the IVDR states: 'The manufacturer shall assess all relevant scientific validity, analytical and clinical performance data to verify the conformity of its device with the general safety and performance requirements as referred to in Annex I. The amount and quality of that data shall allow the manufacturer to make a qualified assessment whether the device will achieve the intended clinical benefit or benefits and safety, when used as intended by the manufacturer.'
- Hence, mindful of the Regulation and its definitions above, manufacturers first describe the intended clinical benefit (based on the intended purpose) and then perform a qualified assessment of the acceptability of benefit-risk of a device and the corresponding clinical evidence as to whether the clinical benefit is achieved. It should be noted that this can be a qualitative assessment based on the judgement of a qualified person taking into consideration other diagnostic information on a

patient as provided by the state of the art in medicine. As outlined in the chapter on Plans and Reports for Performance Evaluation, the intended clinical benefit needs to be described in the Performance Evaluation Plan. The assessment of benefit-risk and clinical evidence towards the achievement of the clinical benefit must be documented in the performance evaluation report.

Appendix 3.1: Examples of Clinical Benefit Assessments (according to the IVDR Article 2 (37) and Recital 64)

The following clinical benefit assessment examples describe the medical information on patients (e.g. screening, monitoring, diagnosis). Although clinical utility is beyond the IVDR requirements, the following examples should aim at illustrating the differences between the concepts of clinical benefit and clinical utility (see also IVDR Annex II).

Clinical Benefit Assessment of a Cyclosporine IVD Device

Based on its analytical performance and scientific validity, this IVD device achieves the clinical benefit of accurately measuring concentrations of cyclosporine in blood. Based on clinical guidelines and textbooks, and when used in conjunction with other diagnostic technologies and options, this medical information is useful in the context of the narrow therapeutic range of cyclosporine, whereby underdosing is associated with an increased risk for transplant rejection, and overdosing is associated with toxicity and an increased risk for nephropathy. This clinical benefit supports physicians in establishing and maintaining efficacious therapeutic drug concentrations and ultimately (the clinical utility of) graft tolerance, while minimising potentially toxic effects of overdosing.

Clinical Benefit Assessment of a Magnesium IVD Device

Based on the clinical evidence, this IVD device achieves the clinical benefit of accurately measuring magnesium in plasma or serum. Based on clinical guidelines and textbooks, and when used in conjunction with other diagnostic technologies and options, this medical information is useful for diagnosing and monitoring magnesium imbalance, including hypomagnesemia (magnesium deficiency) and hypermagnesemia (magnesium excess), both of which can be associated with (or observed during) a number of underlying disease states or pathological conditions. This clinical benefit allows physicians to consider (the clinical utility of) timely clinical interventions or exclusion of magnesium dysregulation.

Clinical Benefit Assessment of a Troponin T/I IVD Device

Based on the analytical and clinical performance (high NPV and PPV), this IVD device achieves the clinical benefit of accurately measuring Troponin T/I in plasma or serum and providing medical information about myocyte (heart cell) injury that can, in conjunction with other diagnostic technologies and options (e.g. chest pain and electrocardiogram) and per clinical guidelines, be used as an aid in the diagnosis of myocardial infarction in patients presenting with chest pain. This clinical benefit allows physicians to consider (the clinical utility of) timely therapeutic interventions or exclusion of myocardial infarction.

Clinical Benefit Assessment of a CD45 2D1 IVD Device

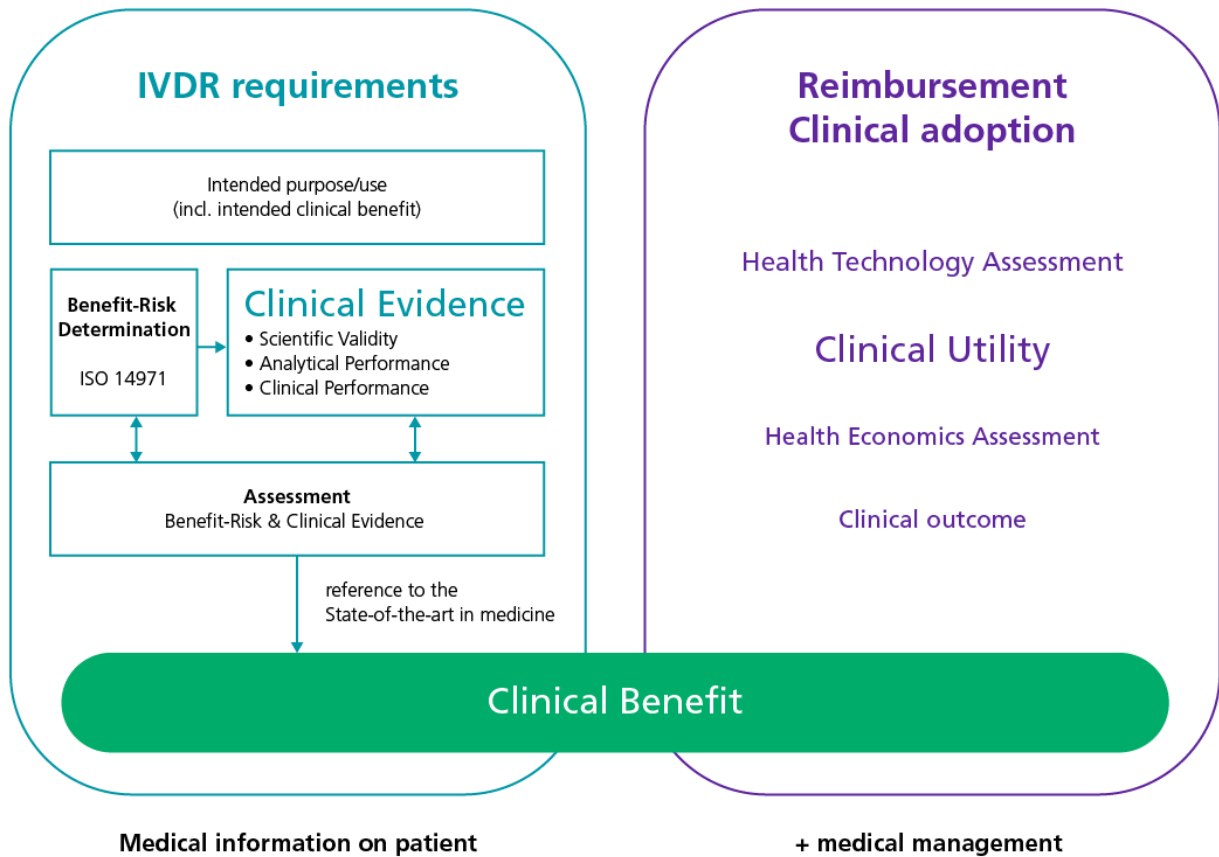
Based on the analytical performance, this IVD device achieves the clinical benefit of accurate identification of haematopoietic cells expressing the CD45 antigen. Based on clinical guidelines for the immunophenotyping of haematopoietic cells, and when used in conjunction with further diagnostic tests or procedures, this medical

information is useful for the assessment of immune status. This clinical benefit allows physicians to consider timely diagnostic or therapeutic options for disorders of the immune system.

Clinical Benefit Assessment of a TBNK (T cells, B cells, Natural Killer cells) IVD Device

Based on the analytical and clinical performance, this IVD device achieves the clinical benefit of accurate identification and measurement of T, B and Natural Killer (NK) lymphocyte subsets, including percentages and absolute counts. Based on clinical guidelines for the identification and enumeration of lymphocyte subsets, and when used in conjunction with further diagnostic tests or procedures, this medical information is useful for the assessment of individuals that have (or are at risk of having) autoimmune diseases or immune deficiencies. This clinical benefit allows physicians to consider timely diagnostic or therapeutic options for autoimmune diseases or immune deficiencies.

Appendix 3.2: Clinical benefit concept under the IVDR and its distinction from clinical utility



References:

- 1) Regulation (EU) 2017/ 746 of the European Parliament and of the Council of April 5, 2017, on *in vitro* diagnostic medical devices
- 2) ISO 14971:2019 - Medical devices — Application of risk management to medical devices

CHAPTER 4 – Clinical Evidence Levels

1) How is clinical evidence defined in the IVDR?

The IVDR introduces a new clinical evidence concept which is defined as follows:

Article 2(36) '*clinical evidence*' means clinical data and performance evaluation 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;

Article 56 (2) - The *clinical evidence* shall support the intended purpose of the device as stated by the manufacturer and be based on a continuous process of performance evaluation, following a performance evaluation plan.

(3) A performance evaluation shall follow a defined and methodologically sound procedure for the demonstration of the following, in accordance with this Article and with Part A of Annex XIII:

- (a) scientific validity (as defined in Art. 2 (39));
- (b) analytical performance (as defined in Art. 2 (40));
- (c) clinical performance (as defined in Art. 2 (41)).

The data and conclusions drawn from the assessment of those elements shall constitute the *clinical evidence* for the device. The *clinical evidence* shall be such as to scientifically demonstrate, by reference to the state of the art in medicine, that the intended clinical benefit(s) will be achieved and that the device is safe. The *clinical evidence* derived from the performance evaluation shall provide scientifically valid assurance that the relevant general safety and performance requirements, set out in Annex I, are fulfilled under normal conditions of use.

2) What is the justification for clinical evidence levels?

'The manufacturer shall specify and justify the level of the 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.' (IVDR, Article 56 (1))

The IVDR does not define how much clinical evidence is required. It is the responsibility of the manufacturer to decide what is appropriate for their device, based on the intended use and risk class.

According to the principles of evidence-based medicine², the term *evidence levels* refers to strength, robustness, and/or quality of the evidence. These levels reflect the source of the evidence, statistical validity, clinical relevance, and peer-review acceptance. The concepts outlined below are specific to IVD medical devices and are based on general principles of evidence-based medicine.

3) What is the general guidance on clinical evidence levels?

The necessity and levels of clinical evidence may vary among IVD devices and classes.

‘Where specific devices have no analytical or clinical performance or specific performance requirements are not applicable, it is appropriate to justify in the performance evaluation plan, and related reports, omissions relating to such requirements’ (IVDR, Preamble 65). Devices without analytical performance include pipets or specimen receptacles, while devices without clinical performance include DNA extraction kits or therapeutic drug monitoring (TDM). As a consequence, performance evaluation reports do not need to include corresponding performance data (Annex XIII Part A (1.3.2)). Due to the applicability of clinical evidence components, the following chapters focus on class B, C and D devices.

If applicable, evidence levels for *analytical performance* and *scientific validity* can be similar for IVD devices regardless of the risk class. Because the IVDR classes are largely based on risks to individuals and/or to public health), the robustness and strength of the evidence should primarily relate to clinical performance. Consequently, evidence levels for *clinical performance* follow a risk-based approach. Thus, the strength and robustness of the clinical performance evidence should follow the following pattern: class B < class C < class D devices (see Figure 2 below).

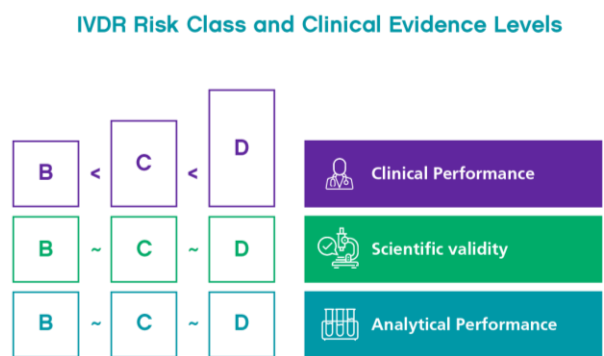


Figure 2. Risk-based evidence levels for analytical performance, scientific validity, and clinical performance

4) How much data is sufficient to demonstrate scientific validity?

Evidence is always needed to prove scientific validity. However, depending on how well established the analyte is, the level and source of required evidence for demonstration of scientific validity may vary. For instance, if the device is well established and in routine clinical use, and if the association of the analyte to a clinical condition or physiological state is well established, evidence from the literature is enough to prove scientific validity. For novel devices, and in the absence of literature, scientific validity should be proven via clinical performance studies or proof of concept studies (GHTF/SG5/N7:2012, Section 6.0) ³.

5) What are the sources for demonstrating clinical performance?

Demonstration of the clinical performance of a device shall be based on one or a combination of the following:

- Clinical performance studies
- Scientific peer-reviewed literature
- Published experience gained by routine diagnostic testing

IVDR Article 56 (4) states that clinical performance studies in accordance with Section 2 of Part A of Annex XIII shall be carried out unless it is duly justified to rely on other sources of clinical performance data.

6) What are the options for clinical performance data?

As per the definition in the IVDR Article 2 (41), clinical performance means '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 use'.

Based on this definition, there are three options for clinical performance:

1. Clinical performance defined as correlation with clinical condition/disease: For devices measuring specific analytes that are associated with a clinical condition/disease and have medical decision points (cut-offs), clinical performance data and a corresponding clinical performance report are required;
2. Clinical performance defined as correlation with a physiological or pathophysiological process or state: For devices measuring analytes without clear medical decision points (cut-offs) or for devices measuring analytes that are not (yet) associated with a clinical condition, clinical performance may be defined as correlation with physiological or pathophysiological process or state, or a justification for omission of clinical performance data may be considered; or
3. No clinical performance data based on a justification, e.g. for devices without analytical or clinical performance or specific performance requirements or a device that does not yield results correlating with a clinical condition or a physiological or pathological process or state.

Justification of omission of any clinical performance data is based on the following IVDR sections:

- Article 2 (39) 'performance of a device' means the ability of a device to achieve its intended purpose as claimed by the manufacturer. It consists of the analytical and, where applicable, the clinical performance supporting that intended purpose.
- Annex XIII Part A (1.2.3) Demonstration of the clinical performance: The manufacturer shall demonstrate the clinical performance of the device in relation to all the parameters described in point (b) of Section 9.1 of Annex I, unless any omission can be justified as not applicable.

In such cases, a clinical performance report is not applicable, but a performance evaluation report including the other clinical evidence components would still be required.

Options for clinical performance	IVD Device	Function / Intended Purpose / Intended Use	Clinical Performance
Correlation with clinical condition/ disease	Troponin T / I test	Diagnosis of acute myocardial infarction	Diagnostic sensitivity and specificity, AUC, NPV, PPV
Correlation with physiological process or state	Creatinine test	Assessment of kidney function	Agreement with other method measuring kidney function
No correlation with a clinical condition or a physiological or pathological process or state	Cyclosporine test	Therapeutic drug monitoring †	Not applicable, reference ranges (if applicable). Omission to be justified in the respective Clinical Performance section of Performance Evaluation Plan and Report

Table 1. Examples of IVD devices along with intended purpose and possible clinical performance. Please note that this table does not provide a comprehensive or prescriptive selection of intended purpose and clinical performance options.

† A Therapeutic Drug Monitoring (TDM) device is a device without medical decision points. Clinical performance data cannot be generated for many TDM devices and the clinical benefit lies in the accurate information about the drug concentration for which different subtherapeutic and toxic drug levels may exist, depending on indications and population.

Rationale for TDM: According to IVDR Article 2 (41), 'clinical performance' means 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. For products for Therapeutic Drug Monitoring (TDM), the assays measure the level of the administered drug and/or its metabolites in bodily fluids, e.g. blood, urine. These levels can show tremendous intra- and inter-patient variability, depending on a variety of factors, including time after treatment, concomitant medication, organ function, drug toxicity and others. Since the drug is usually administered to treat an underlying clinical condition and measurement of the concentration of the drug is used to determine whether the levels are within the therapeutic window for that specific patient, there is no direct connection of the device to a clinical condition or physiological process or state. Therefore, none of the clinical performance parameters referenced in IVDR Annex I, 9.1(b), e.g. diagnostic sensitivity, diagnostic specificity, positive or negative predictive value, likelihood ratio, expected values, is applicable.

Determination of the therapeutic window, toxic or sub-therapeutic levels for each drug is the responsibility of the drug manufacturers and demonstration of clinical performance of an IVD device for TDM does not imply that IVD manufacturers determine sensitivity or specificity of finding such levels. Also, it has been demonstrated that the establishment of generalized reference (or therapeutic) ranges for most therapeutic drugs that require monitoring is extremely difficult, due to a wide variety of influencing factors. E.g. for cyclosporine therapeutic ranges in solid organ (kidney, liver, heart) transplant settings are not absolutely defined, as they can be widely variable, dependent on a clinical protocol, organ transplanted, time after transplant, risk of rejection, concomitant immunosuppressive drugs, organ function and cyclosporine toxicity.

As a result, the analytical performance data (including method comparisons to a reference method or device) are sufficient to demonstrate that such a product is able to accurately and precisely measure the concentration of the drug and/or its metabolites, and, in consequence, is capable of monitoring the drug accordingly. If the data presented in the Analytical Performance Report show that the analyte is measured with sufficient accuracy and precision in human specimens, within the measuring range which covers the therapeutic range and potentially toxic concentrations (as established by the drug manufacturer), in accordance with IVDR Recital (65), Article 2 (39), Article 56 (1-3), product-specific clinical performance data can be judged to be unnecessary, and performance claims are addressed sufficiently by the analytical performance.

7) How much clinical performance data is sufficient to demonstrate ‘clinical evidence’?

Clinical performance data and evidence levels

As outlined in Annex XIII Part A (1.2.3) of the IVDR, clinical performance data can be demonstrated based on one or a combination of clinical performance studies, scientific peer-reviewed literature, and/or published experience gained by routine diagnostic testing (see also the chapter of this brochure on ‘published experience gained by routine diagnostic testing’). In any case, the strength and robustness of clinical performance evidence will ultimately depend on study design and biostatistical considerations.

In principle, demonstration of clinical performance can be direct or indirect or a combination thereof. Direct demonstration of clinical performance indicates that the data are based on the particular device produced by the IVD manufacturer and are obtained from studies using prospectively collected specimens or biobank/leftover specimens. Indirect demonstration indicates that the data are based on literature search or a comparison with a reference device (e.g. method comparison). Direct demonstration yields stronger evidence levels of clinical performance data than indirect demonstration and should accordingly be applied to higher risk class and / or novel devices. It should be noted that these principles relate to an individual clinical performance data set of a particular IVD device and not to the available pool of evidence of a reference IVD device. For example, a method comparison study may provide appropriate evidence for a particular IVD showing equivalence with a selected reference device that has a published and accepted strong body of clinical evidence.

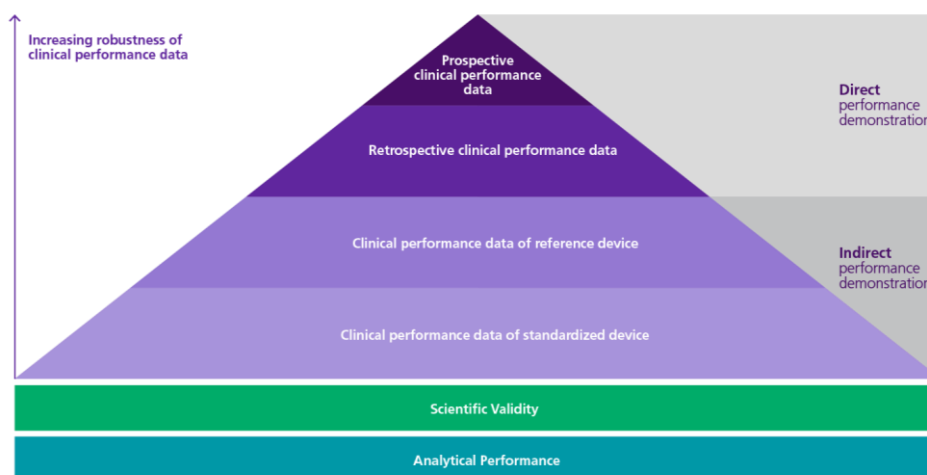


Figure 3. Clinical evidence levels for IVD classes B, C, and D

It should be noted that multiple general evidence grading systems exist (e.g. GRADE⁵), QUADAS-2⁶, Hayes⁷) and they have been reviewed and considered under the proposed framework above.

Drivers of the evidence level of clinical performance data include:

- I) Intended purpose/use
- II) Groups according to the Global Harmonization Task Force (GHTF)³
 - a) established, standardised device
 - b) established, non-standardised device
 - c) novel device
- III) IVDR class

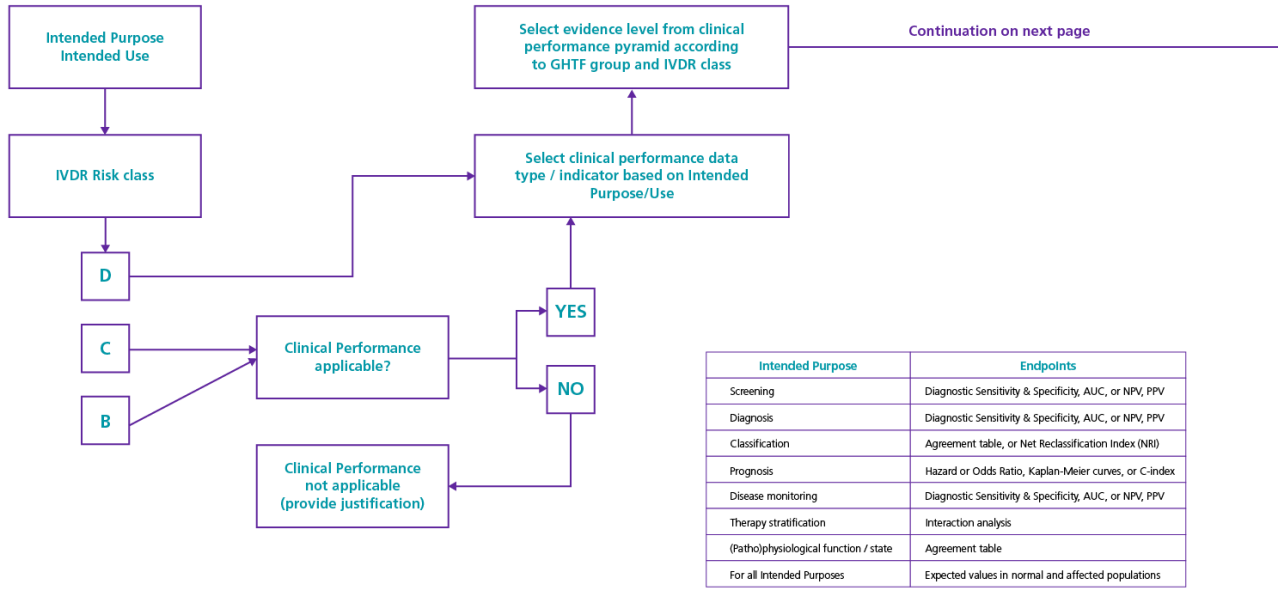
Determining clinical performance indicators and study endpoints

A clear definition of the intended purpose/use is the first and essential step to determine the clinical performance indicator(s) and corresponding study endpoint(s) or data type(s) (see [CHAPTER 1 - 'Intended Purpose/Use'](#) and [CHAPTER 2 - Analytical and clinical performance indicators](#)). Specifically, the clinical function in the intended purpose defines the clinical performance indicator(s)/data type(s) and the study endpoint(s) the study endpoint(s), e.g. diagnostic sensitivity and specificity for a test claiming a diagnostic intended purpose/use and a hazard ratio for a test claiming prognostic intended purpose. A device's intended purpose and target population also define the IVD risk class.

The strongest clinical performance data are derived from adequately statistically powered prospective clinical performance studies. The vast majority of these studies are typically observational, thus non-interventional in design. This may be an option for novel devices, if no biobank or leftover samples are available. Wherever available or applicable, the generation of clinical performance data should follow the EU Common Specifications (CS) or international technical specifications (e.g. WHO, ISO 15197 'Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus' and ISO 17593 'Requirements for in vitro monitoring systems for self-testing of oral anticoagulant therapy').

Retrospective studies typically use biobank or leftover samples representing the intended purpose/use population along with the necessary clinical data to determine clinical performance. Like prospective studies, they need to be adequately powered to yield robust clinical performance data. Retrospective studies may lead to more bias than prospective studies (selection bias, changes in medical practice, etc.). Therefore, retrospective clinical performance studies may be an option for novel and established devices depending on the quality of the samples.

Indirect demonstration of clinical performance can be shown using a method comparison study against a reference device, provided that the clinical performance of the reference device is known and published. This may be an option for established devices, but not standardised devices. Finally, an option for established and standardised devices may be indirect demonstration of clinical performance via published data from reference devices, provided the analytical performance determination is performed using standardised device and reference material.



Flowchart for Clinical Performance

**Please note that it is the manufacturer's sole responsibility to choose an appropriate and applicable performance indicator and that not all mentioned performance indicators are applicable to all devices*

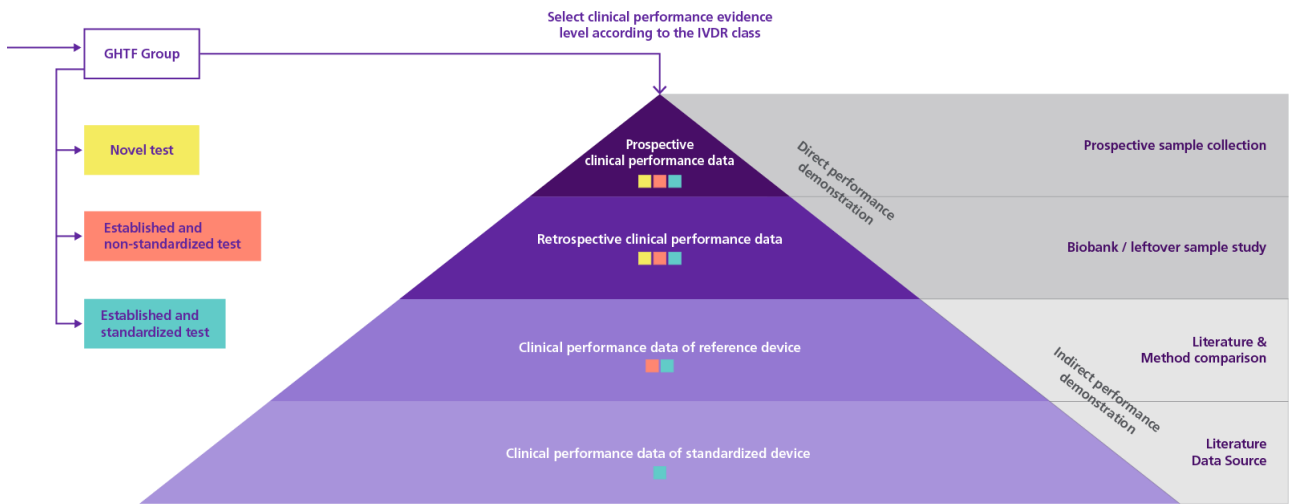


Figure 4. Flowchart for options of clinical performance data types and evidence levels.

8) How can post-market data be used to satisfy the clinical evidence requirements of established products?

Post-market data may allow manufacturers to comply with clinical evidence requirements in the technical files of established products. Annex XIII of the IVDR requires that manufacturers demonstrate clinical performance of their products (unless duly justified to omit it), which will be documented in the Clinical Performance Report (CPR) (IVDR, Annex XIII, Section 1.2.3). The demonstration of clinical performance of a device can be based on one or a combination of clinical performance studies, scientific peer-reviewed literature or published experience gained by routine diagnostic testing. See [CHAPTER 5 - How to demonstrate evidence gained from 'published/documentated routine testing'](#) and [CHAPTER 8 – Documentation of Performance Evaluation requirements](#)

The use of post-market data to address clinical evidence requirements should be subject to the appropriate risk analysis. This should consider how critical it is for the safety and performance of the device in question.

Definitions of Novel, Established and Standardised Devices ^{3, 4}

Novel Device

- a device which incorporates technology (the analyte, technology or test platform) not previously used in diagnostics and not continuously available on the European Community market during the previous three years, or;
- an existing device which is being used for a new intended purpose for the first time.

Established Status

- Established tests have clinical guidelines and/or consensus for the use of the test and/or are medically accepted as gold standard

Standardisation

- An international standard or accepted reference materials (e.g. WHO) of the analyte exists, and
- More than one commercial test is available, and
- Standardised devices/tests produce equivalent results for the analyte regardless of the method/manufacturer. Equivalence will depend on the device, intended purpose/use, risk class, and authority view.

References:

- 1) Regulation (EU) 2017/ 746 of the European parliament and of the council of April 5, 2017 on in-vitro diagnostic medical devices
- 2) U.S. Preventive Services Task Force (August 1989). Guide to clinical preventive services: report of the U.S. Preventive Services Task Force. DIANE Publishing. pp. 24–. ISBN 978-1-56806-297-6.
- 3) GHTF/ GHTF/SG5/N7:2012 Clinical Evidence for IVD medical devices – Scientific Validity Determination and Performance Evaluation
- 4) Definitions from MDEG New and Emerging Technologies Task Force
- 5) Whiting PF, Rutjes AW, Westwood ME, et al, the QUADAS-2 Group. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med. 2011;155:529–536.
- 6) Whiting PF, Rutjes AW, Westwood ME, et al, the QUADAS-2 Group. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. Ann Intern Med. 2011;155:529–536.
- 7) The Hayes Rating, <https://www.hayesinc.com/hayes/about/hayes-rating/>

CHAPTER 5 - How to demonstrate evidence gained from 'published/documentated routine testing'

According to the IVDR, demonstration of the clinical performance of a device shall be based on one or a combination of clinical performance studies, scientific peer-reviewed literature and/or published experience gained by routine diagnostic testing.

Under the IVDD, clinical performance studies are already a source of data for the demonstration of clinical performance. Scientific peer-reviewed literature includes articles from journals, posters from conferences, guidance or documents from official websites (i.e. MedTech Europe, IMDRF, WHO, local authorities, European Medicines Agency etc.) and/or guidelines and textbooks, provided that the data is peer-reviewed. However, the third possible source (published experience gained by routine diagnostic testing) is open to more interpretation. This brochure aims to help manufacturers meet the expectations implied by the IVDR.

If a manufacturer chooses to use experience data from routine diagnostic testing, it is important that any reports or collations of data contain sufficient information. This information must allow the undertaking of a rational and objective assessment and ultimately support the conclusion of its significance with respect to the performance of the IVD medical device in question. Reports of such experience that are not adequately supported by data, such as anecdotal reports or opinion, should not be used. For established products, routine diagnostic testing (including Post Market Performance Follow-up (PMPF) data) is expected to be immediately available and can be used as clinical evidence, in addition to existing performance evaluations and scientific literature.

1) As literature is 'published', does published experience refer to literature?

No, it is a supplementary item in the Regulation, separate from literature, since literature is already covered in the second indent of Annex XIII, Part A, 1.2.3

2) What do we mean by published?

The definition[‡] is broad and includes:

- Information that is issued (printed or otherwise reproduced textual material etc.) for sale or distribution to the public
- Information that is issued publicly[§]
- Information that is submitted (content) online, (e.g. laboratory/hospital intranet)
- Information that is announced formally or officially; proclaimed; promulgated
- Information that can be accessed upon request (e.g. internal document)

[‡] Modified from *Dictionary.com*

[§] Might be free of charge (e.g. website from clinical labs)

Any published item should be authored (identifiable source) and cover the intended purpose.

3) What does published experience refer to?

Any document or set(s) of data coming from the use of the device and are published (according to the above definition).

4) Can we use PMPF data as part of published experience gained by routine diagnostic testing?

Yes, post-market surveillance data generated by the manufacturer (e.g. customer testing results) can be used. PMPF data can be complemented if required, by literature, other routine diagnostic testing or further studies.

5) What other kind of data are included in published experience gained by routine diagnostic testing?

After having considered the quality and robustness of data (case by case analysis), we propose including any of the following:

- data from evaluation or re-evaluation by competent authorities (e.g. [ANSM in France](#))
- data from accreditation (laboratory validation data)
- proficiency data report/external quality assurance data (e.g. independent medical and/or laboratory associations such as WHO or IFCC)
- data from post-launch studies (after CE marking)
- data from investigator-initiated studies
- data from real-world evidence, e.g. registries
- data from Health Economics and Outcome Research (HEOR) studies

CHAPTER 6 – Equivalence and similarity concepts in the IVDR

1) Where and how are the terms ‘equivalence’ and ‘similar’ used in the IVDR? And how are they defined?

The IVDR does not include a definition of ‘equivalence’ or ‘similar’ even though both terms are used either alone or in combination in relation to performance evaluation and post-market surveillance.

The IVDR uses the terms ‘equivalence’ or ‘equivalent’ or ‘similar’ or ‘equivalent and/or similar’ in the following ways:	
Preamble	(53) For class D devices for which no Common Specifications (CS) exist it is appropriate to provide that where it is the first certification for that specific type of device and there is no similar device on the market having the same intended purpose and based on similar technology, notified bodies should (...) be obliged to request expert panels to scrutinise their performance evaluation assessment reports
Article 2: Definitions Generic device group	(8) Means a set of devices having the same or similar intended purposes or a commonality of technology allowing them to be classified in a generic manner not reflecting specific characteristics
Annex I: General Safety and Performance Requirements; Chapter II	11.6 The labelling of the device shall distinguish between identical or similar devices placed on the market in both a sterile and a non-sterile condition additional to the symbol used to indicate that devices are sterile
Annex II: Technical Documentation 1.2	(b) an overview of identified similar devices available on the Union or international markets, where such devices exist
Annex III: Technical Documentation on Post-Market Surveillance	<p>1. The post-market surveillance plan drawn up in accordance with Article 79</p> <p>(a) post-market surveillance plan shall address the collection and utilization of available information, in particular</p> <ul style="list-style-type: none"> - publicly available information about similar medical devices. <p>(b) The post-market surveillance plan shall cover at least:</p> <ul style="list-style-type: none"> - a proactive and systematic process to collect any information referred to in point (a). The process shall allow a correct characterization of the

	performance of the devices and shall also allow a comparison to be made between the device and similar products available on the market
Annex VII: Requirements to be met by Notified Bodies	Section 4.5.4 Performance Evaluation Assessment The notified body's assessment of the performance evaluation as referred to Annex XIII shall cover: <ul style="list-style-type: none"> - Validity of equivalence claimed in relation to other devices, the demonstration of equivalence, the suitability and conclusions data from equivalent and similar devices
Annex VII: Requirements to be met by NB	4.10 Surveillance activities and post-certification monitoring The NB shall, if listed as part of the conditions for certification: conduct an in-depth review of the performance evaluation as most recently updated by the manufacturer based on the manufacturer's post-market surveillance, on its PMPF and on clinical literature relevant to the condition being treated with the device or on clinical literature relevant to similar devices
Annex IX: Conformity Assessment based on a Quality Management System and on assessment of Technical Documentation	Chapter I Quality Management System (c) the procedures and techniques for monitoring, verifying, validating and controlling the design of the devices, and the corresponding documentation as well as the data and records arising from those procedures and techniques. Those procedures and techniques shall specifically cover <ul style="list-style-type: none"> - the strategy for regulatory compliance, including processes for identification of relevant legal requirements, qualification, classification, handling of equivalence, choice of, and compliance with, conformity assessment procedures Chapter II Assessment of the Technical Documentation 4.5 The notified body shall, in circumstances in which the clinical evidence is based partly or totally on data from devices which are claimed to be equivalent to the device under assessment, assess the suitability of using such data, taking into account factors such as new indications and innovation. The notified body shall clearly document its conclusions on the claimed equivalence , and on the relevance and adequacy of the data for demonstrating conformity.
Annex X: Conformity Assessment based on Type-Examination	3. Assessment (d) in circumstances in which the clinical evidence is partly or totally based on data from devices which are claimed to be similar or equivalent to the device under assessment, assess the suitability of using such data, taking into account factors such as new indications and innovation. The notified body shall clearly document its conclusions

	on the claimed equivalence , and on the relevance and adequacy of the data for demonstrating conformity;
Annex XIII: Post-Market Performance follow up	5.2 The PMPF plan shall include at least: (f) an evaluation of the performance data relating to equivalent or similar devices, and the current state of the art
Annex XIV: Interventional clinical performance studies and certain other performance studies	2. Investigator´s brochure 2.1 Identification and description of the device, including information on the intended purpose, the risk classification and applicable classification rule pursuant to Annex VIII, design and manufacturing of the device and reference to previous and similar generations of the device. 2.4 Existing 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; — 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.

Table 1: Compilation of references of terms ‘equivalence’, ‘equivalent’, ‘similar’ throughout the IVDR

2) Do the terms ‘equivalence’ and ‘similar’ have different meanings?

The IVDR does not suggest different meanings for ‘equivalent’ and ‘similar’ as both terms are associated with product characteristics which can be assessed by comparison. Nevertheless, the results of such comparison can be interpreted differently.

- **‘Similar’** can be interpreted as a broader and softer term. Devices can be considered as similar based on a review of publicly available product data, including e.g. instruction for use, product composition, design, features, intended purpose and/or the performance of another, comparator device. No in-depth analysis or systematic method comparison study is required.
- **‘Equivalent’** can be considered as a narrower and stronger term. Objectively, a device is considered as equivalent when, based on a review of publicly available product data, the device in question is

either almost identical to the comparator device or identical to the comparator device regarding the product composition, design, features, or intended purpose. In order to demonstrate equivalent performance, a systematic method comparison is required, where performance should correspond to the performance of a comparator device within the pre-defined limits (e.g. CLSI guidelines for method comparison).

- Hence, a device can be considered as **similar** if there are no meaningful differences in safety as well as analytical and/or clinical performance of the device. A device can be considered as **equivalent** if there are no meaningful differences in the critical characteristics.

3) How can similarity or equivalence of a device in question be assessed?

Table 2 aims at providing guidance on how to assess similarity or equivalence of an IVD device based on the IVD-relevant characteristics, such as technical, analytical, biological and clinical features. The goal of this comparison is to identify any meaningful difference in the safety as well as the analytical and/or clinical performance of a device under evaluation. In order to perform such an assessment, manufacturers are required to be able to access the relevant data of a comparator device to which they claim equivalence or similarity.

Device characteristics	Device 1 (device under evaluation)	Device 2 (device to which IVD similarity and/or equivalence is claimed)	Differences Device 1 vs Device 2	Applied standards and/or other guidelines	Justification for claiming IVD similarity and/or equivalence
Medical device nomenclature code					
Technology (e.g. ELISA, Western Blot, PCR, Flow Cytometry)					
Device Design (e.g. sample volume, processing and incubation time, critical reaction component(s), read-out technology (e.g. chemi-luminescence))					

Automated or manual system, operating conditions					
Analytical performance characteristics (Annex I, Chapter II, 9.1 and Annex II, Section 6.1)					
Specimen type(s)					
Biological controls (metrological traceability)					
Antibodies (polyclonal/monoclonal)					
Intended purpose					
Target population					
Intended user (professional use, near patient test, self-testing)					
Test limitations					
Scientific validity					
Clinical performance Annex I, Chapter II, 9.1 (b)					
Clinical benefit					

Table 2: Assessment of similarity and/or equivalence of IVD devices. Please note that this table does not provide a comprehensive or prescriptive selection of meaningful characteristics. It is the manufacturer's sole responsibility to define an appropriate concept.

4) How to use this table?

The table lists possible technical, analytical, biological and clinical characteristics of an IVD device. It is a non-exhaustive and non-prescriptive compilation of different parameters; therefore, the chosen comparison criteria shall be relevant to a device under evaluation. Based on the proposed definitions for similarity and/or

equivalence, each feature (technical, analytical, biological and clinical) will be rated as either similar or equivalent, followed by a qualitative evaluation.

References:

- 1) Regulation (EU) 2017/ 746 of the European parliament and of the council of April 5, 2017 on in vitro diagnostic medical devices

CHAPTER 7 – Companion Diagnostics

1) How are Companion Diagnostics described in the IVDR?

Recitals 10 to 12 and Article 2 (f) of the IVDR introduce a new companion diagnostics concept.

Recital 10 (...) tests that provide information to predict treatment response or reactions, such as companion diagnostics, are in vitro diagnostic medical devices

Recital 11 Companion diagnostics are essential for

- defining patients' eligibility for specific treatment with a medicinal product through the quantitative or qualitative determination of specific markers identifying subjects at a higher risk of developing an adverse reaction to the medicinal product in question or
- identifying patients in the population for whom the therapeutic product has been adequately studied and found safe and effective. Such biomarker or biomarkers can be present in healthy subjects and/or in patients.

Article 2(f) Companion diagnostic means a device which is essential for the safe and effective use of a corresponding medicinal product to:

- identify, before and/or during treatment, patients who are most likely to benefit from the corresponding medicinal product; or
- identify, before and/or during treatment, patients likely to be at increased risk of serious adverse reactions as a result of treatment with the corresponding medicinal product;

2) What are NOT companion diagnostics **?

A) The IVDR Recital 12 clarifies that “Devices that are used with a view to monitor treatment with a medicinal product in order to ensure that the concentration of relevant substances in the human body is within the therapeutic window are not considered to be companion diagnostics”.

Examples include:

– **Cyclosporine as a Therapeutic Drug Monitoring Devices (TDM)**

The introduction of cyclosporine into clinical practice improved transplant outcome. A narrow therapeutic index coupled with variable absorption and unpredictable pharmacokinetics has resulted

** **Complementary Diagnostic** Assays are neither defined nor described in the IVDR but are generally understood as recommended but not required for the safe and effective use of a medicinal product. They may for instance aid physicians in identifying patients who may be relatively more likely to derive benefit from treatment with a particular medicinal product ³

in the need to measure cyclosporine blood concentrations to enable the dose of the drug to be individualized to the patient. When this is done correctly therapeutic efficacy can be maximized while toxicity is kept to a minimum².

Such a device intended to monitor levels of medicinal products, substances or biological component is classified IVDR Annex XIII, rule 3 (j). Please see for further information the MTE [CHAPTER 4 – Clinical Evidence Levels](#), section ‘Clinical Performance of IVD Devices for Therapeutic Drug Monitoring (TDM)’

– **Blood glucose monitoring devices**

These devices are intended for the quantitative measurement of blood glucose levels in freshly collected capillary blood samples. Such monitors provide immediate information to the user on whether the blood sugar is too high (**hyperglycaemia**) or too low (hypoglycaemia). In cases of hyperglycaemia, the test result is then used to calculate an adequate insulin dosage to be administered to the patient. Such devices, intended to monitor by determination of the blood glucose levels whether results are within the acceptable range, do not follow the definition of CDx in Article 2 (f) as described in question 1.

B) If a study test result does not lead to any treatment decision or is used in the context of enrichment and/or exploratory studies, such devices are *not* companion diagnostics with the meaning of the CDx definition in Article 2 (f) as described in question 1.

– Enrichment is the prospective use of any patient characteristic to select a study population in which detection of a drug effect (if one is in fact present) is more likely than it would be in an unselected population. Enrichment strategies are intended to increase the efficiency of drug development and support precision medicine, i.e., tailoring treatments to those patients who will benefit based on clinical laboratory, genomic, and proteomic factors⁴.

– Exploratory investigational new drug (IND) study is intended to describe a clinical trial that

- is conducted early in phase 1
- involves very limited human exposure
- has no therapeutic or diagnostic intent (e.g., screening studies, micro-dose studies)⁵

3) What are the requirements for companion diagnostics performance studies?

A CDx performance study is:

– A certain performance study as described in Article 58 (2) as follows: ‘performance studies involving companion diagnostics shall be subject to the same requirements as the performance studies listed in Article 58 paragraph (1)’

– covered by the term ‘interventional clinical performance study’ as defined in the IVDR §2 (46): ‘interventional clinical performance study is a clinical performance study where the test results may

influence patient management decisions and/or may be used to guide treatment or where the conduct of the study involves additional invasive procedures or other risks for the subjects of the studies’

It follows that performance studies involving companion diagnostics must meet the

- General requirements set out in Article 57 and Annex XIII
- Additional requirements set out in Art 58 to 77 and Annex XIV.

In the special situation where only leftover or archived samples^{††} are used, the IVDR emphasizes that most of the additional requirements do not apply to performance studies involving companion diagnostics / Article 58(2). Such studies must, however, be notified to the competent authority.

A study concept with leftover or archived samples may play a role in bridging studies e.g. bridging clinical trial assay (CTA) with final CDx with samples taken at time of the CTA or adaption of an established CDx test on a new instrument platform by linking the existing clinical data set to the new combination.

CDx studies should be conducted based on an adequate analytical performance and scientific validity data set. If the scientific validity for the Companion Diagnostic is not established, manufacturers must provide the scientific rationale for the use of the biomarker.

†† How are leftover & archived specimens defined?

- Retrospective samples may include leftover, banked, archived or residual specimens.
- The IVDR text does not define any of these terms.
- The ISO standard contains no definition for banked or residual samples but refers to tissue banks or biobanks.
- The ISO 20916 defined these terms as follows⁶:
‘Leftover specimen = leftover sample as unadulterated remnants of human derived specimens collected as part of routine clinical practice and after all standard analysis has been performed
Note 1 to entry: Such specimens/samples would be otherwise discarded as there is no remaining clinical need for them. Note 2 to entry: This can include specimens collected for research or other purposes not connected to the clinical performance study in question’.
- The GHTF/SG5/N8: 2012 defined archived samples as follows⁷
Archived specimen = archived sample specimen or *sample* (3.42) that was collected in the past and is obtained from repositories (e.g. tissue banks, commercial vendor collections)

An overview of the IVDR general and additional requirements in relation to CDx performance studies is shown in the Figure 1 below.

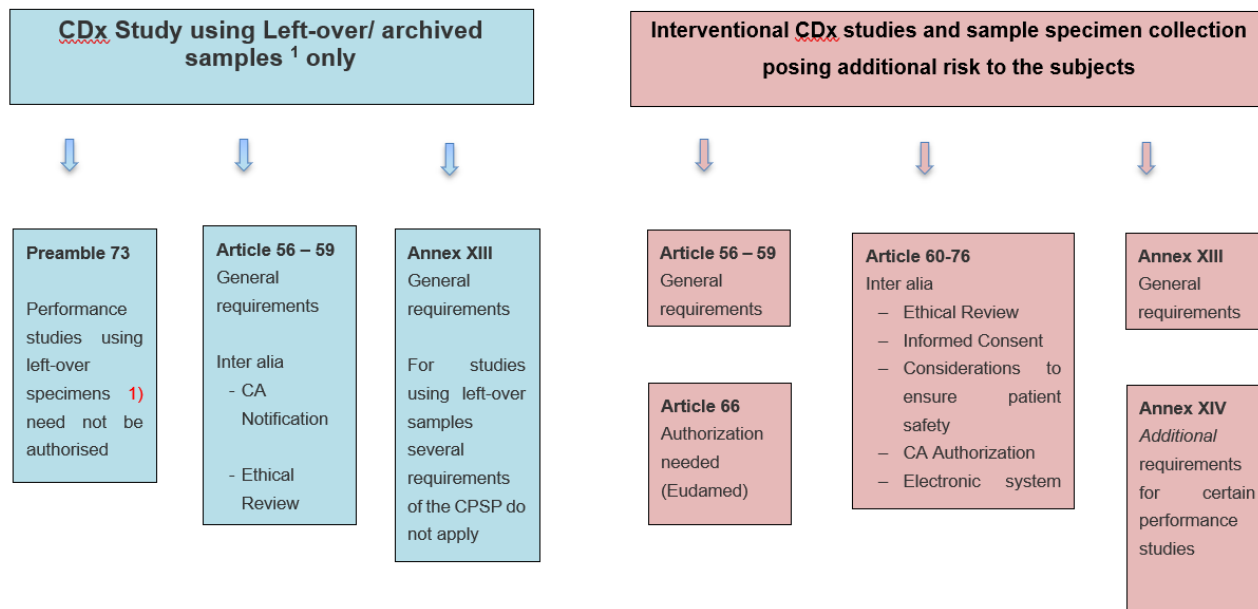


Figure 1. IVDR Requirements for CDx study using leftover/archived samples compared to interventional CDx study and specimen collection with additional risk to the subjects

4) When can a CDx interventional clinical performance be initiated?

In addition to the ethics review and other local requirements, an interventional clinical performance study needs to be authorised by the Member State(s) in which the study is to be conducted (Article 58 (5) a) according to the procedure described in Article 66.

The application for the interventional study includes in principle the unique single identification number for the study, the opinion of the ethics committee, Informed consent from the study subjects and the application dossier in accordance with section 2 and 3 of Annex XIII and Chapter I of Annex XIV.

Based on Article 66 the Notified Body is not involved into the application process. However, with regard to the documents to be submitted to the authorities, further developments need to be tracked. Submission takes place via the clinical module of the EUDAMED system (Article 69).

The Member States notify the sponsor of the authorization. If the study is conducted in more than one Member State, the so called 'coordinating Members State' (Article 74) will inform the sponsor. It must be noted that the 'Coordinated assessment procedure for performance studies' under Article 74 is not yet introduced.

The process flow about the application for an interventional CDx performance study based on the articles 66,67 and 71 is displayed in the Figure 2 below.

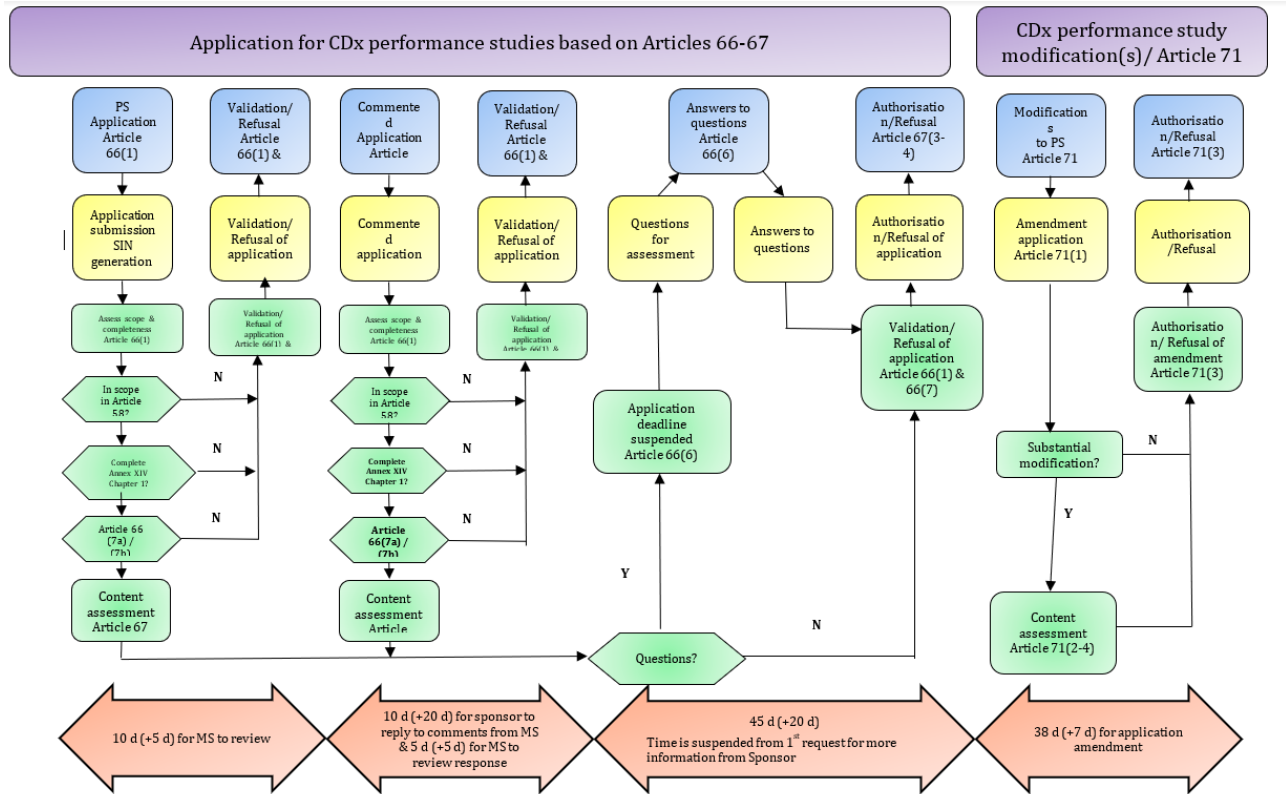


Figure 2. Process flow about the application for an interventional CDx study and related timelines based on Articles 66, 67 and 71.

5) When can a CDX study with leftover or archived samples be initiated?

This type of study must be notified to the competent authorities(s) (Article 58(2)) from the Member State(s) where the study is conducted. Prerequisite for the notification is no objection from an ethics committee from the Member State where the study is conducted.

Unlike the authorization, it is unclear if this notification is planned as a national notification or if it will be done over the clinical module from EUDAMED (Article 69). In principle the sponsor can start the study after the notification. However, national laws should be considered.

6) What are the specific labelling requirements of devices used in interventional performance studies?

CDx devices, used in an interventional or performance study using leftover samples only, should indicate on the product label that this is a 'device for performance study' (Annex I, 20.2 (e)). Such a product label cannot

bear the CE -mark because only devices, *other* than devices for performance studies, considered to be in conformity with the requirements of the Regulation shall bear the CE marking of conformity (Article 18.1).

7) What are the components of Clinical Evidence relevant for CDx?

The clinical evidence aspects for CDx devices are similar to other IVD devices as discussed previously in this brochure. Specifically, clinical evidence for CDx IVD devices includes the demonstration of scientific validity, analytical performance, and clinical performance in accordance with IVDR Article 56 and with Part A of Annex XIII and Article 58 with Annex XIV.

8) What are the typical indicators of analytical and clinical performance?

Indicators of *analytical performance* are typically similar or even identical across IVD devices, including CDx devices (see Q&A on Analytical vs Clinical Performance). Conversely, indicators of *clinical performance* vary and depend strongly on the Intended Purpose/Use. Specifically, the clinical function in the Intended Purpose/Use defines clinical performance indicator (see Table 1 below).

In the case of CDx devices, the two typical clinical functions in the Intended Purpose/Use are:

- ‘therapy stratification’ (also known as ‘therapy response prediction’, or ‘predictive CDx Intended Use’ in other references), or less frequently
- ‘therapy selection’ (also known as ‘selective CDx Intended Use’ similar to therapy stratification, but applied when a “marker positive only” study design is used).

No other Intended Purpose/Use than ‘therapy stratification’ or ‘therapy selection’ is considered in this Q&A document (e.g. ‘complementary diagnostics’ or ‘precision dosing’ diagnostics are not CDx and are therefore out of scope as described under 2).

This CDx-specific Intended Purpose/Use requires evidence to describe the IVD device performance in the context of the corresponding therapy with regards to the efficacy and safety of the therapeutic. Thus, the medical treatment of the patient needs to be taken into consideration in order to generate appropriate clinical evidence for a CDx device to stratify or select a specific therapy. This is possible during co-development of IVD CDx and therapeutic or after development of the therapeutic.

In the latter case, a clinical trial assay (CTA) instead of the final CDx can be used for patient management in the clinical trial. In this case, a concordance study (or bridging study) including appropriate statistical analysis is required to assess the agreement between CDx and CTA in order to bridge the clinical data (e.g. overall survival) from CTA to CDx and to evaluate the therapeutic efficacy in CDx intended use population⁸.

Another example of CDx development after launch of a therapeutic is a follow-on CDx device, when concordance to a previously developed comparator CDx to a therapeutic can already be shown⁹.

In any case, a corresponding study and analysis needs to show that the proposed CDx device is able to stratify or select the patients into likely responders or on-responders (see Table 1), and subsequently also show that the group of patients that was characterized as likely responders were also the ones that benefitted the most from the treatment and/or show favourable safety¹⁰. Accordingly, clinical performance indicator(s) and thus, the endpoints of the corresponding studies, are typically driven by the intended benefit of the therapeutic. Moreover, such a study may consist of a retrospective analysis of biobank samples and corresponding clinical data (typically from drug development trials using a similar IVD device) and/or a prospective study, i.e. a randomized controlled interventional clinical outcome study that is typically the pivotal drug trial. The selected study design may depend on the development phase of the therapeutic, the scientific validity of the test (including similarity of molecular diagnostic and therapeutic targets), the benefit risk ratio of the therapeutic, and other factors.

Typical Performance Indicators



 Analytical Performance	 Clinical Performance	
<ul style="list-style-type: none"> • Measuring Interval: LoQ as the lower limit and the upper limit of Linearity as the upper limit. • LoB (e.g. CLSI guideline EP17-A2) • LoD (=analytical sensitivity) (e.g. CLSI guideline EP17-A2) • LoQ (e.g. CLSI guideline EP17-A2) • Linearity (e.g. CLSI guideline EP06A) • Precision (repeatability) (e.g. CLSI guideline EP05-A3) • Intermediate Precision (e.g. CLSI guideline EP05-A3) • Reproducibility (e.g. CLSI guideline EP05-A3) • Carryover (e.g. CLSI guideline H26-A2) • Total Analytical Error (Accuracy) (e.g. CLSI guideline EP21-A) • Instrument Comparison (e.g. CLSI guideline EP09-A3) • Method Comparison (e.g. CLSI guideline EP09-A3) • Interfering Substances (=analytical specificity): Could be done by checking known and expected interferences, e.g. from vigilance cases and literature research. 	Intended Purpose	Performance Indicator
	Screening	Diagnostic Sensitivity & Specificity, AUC, or NPV, PPV
	Diagnosis	Diagnostic Sensitivity & Specificity, AUC, or NPV, PPV
	Classification	Agreement table, or Net Reclassification Index (NRI)
	Prognosis	Hazard or Odds Ratio, Kaplan-Meier curves, or C-index
	Disease monitoring	Diagnostic Sensitivity & Specificity, AUC, or NPV, PPV
	Therapy stratification Therapy selection	Outcome measure, e.g. response rate, survival, Hazard ratio, a.o.
	(Patho) physiological function / state	Agreement table
	For all Intended Purposes	Expected values in normal and affected populations

Table 1: Possible examples of analytical and clinical performance indicators based on the intended purpose. Therapy stratification or therapy selection is the typical intended purpose/use of CDx devices.

Please note that this table does not provide a comprehensive or prescriptive selection of performance indicators. It is the manufacturer's sole responsibility to define an appropriate clinical evidence concept.

Box 1: Abbreviations

AUC: Area under the curve
 LoB: Limit of blank
 LoD: Limit of detection
 LoQ: Limit of quantification
 NPV: Negative predictive value
 NRI: Net reclassification index
 PPV: Positive predictive value

Intended Purpose	Performance indicator	Study population	Study design	Examples
Screening (early detection of subclinical disease)	Diagnostic sensitivity & specificity (against the "gold standard"/ reference method), AUC , NPV, PPV	Subjects at risk (indicated for screening) Could be population level	Prospective or retrospective observational, longitudinal study (1-arm) or corresponding RWD	Bloodscreening for Infectious Diseases
Diagnosis	Diagnostic sensitivity & specificity (against the "gold standard"/ reference method), AUC , NPV, PPV	Subjects with signs and symptoms of disease	Prospective or retrospective observational cohort study or cross-sectional case-control study	Troponins for AMI
Classification / Grading	Agreement tables, NRI (Net Reclassification Index); if a gold standard available: also Sens/Spec	Subjects diagnosed with the disease of interest	Prospective or retrospective observational study, "case-control" study (cases with different grading)	Creatinine for kidney function / failure
Prognosis /Risk Stratification	Hazard ratio, Odds ratio, Kaplan-Meier curves, C-index, NRI, absolute survival estimates	Depending on IU, population level, or subjects with disease	Prospective or retrospective observational study (Less preferred: case-control study)	CRP, LDL
Disease monitoring	Diagnostic sensitivity & specificity, AUC (against gold standard), NPV, PPV	Diseased patients with or without treatment	Prospective or retrospective observational longitudinal study	Glucose, PSA
Therapy stratification (CDx)	Patient outcome measure and interaction analysis (CDx defined group for therapeutic efficacy and/ or safety)	All-comers (all patients under treatment of the drug)	Clinical outcome study/prospective randomized controlled trial (RCT) or retrospective study Concordance (bridging) studies	HER2, BRAF, KRAS
Therapy selection (CDx)	Patient outcome measure and interaction analysis (CDx defined group for therapeutic efficacy and/ or safety)	Biomarker-positive patients	Clinical outcome study/prospective RCT or retrospective study Concordance (bridging) studies	BRAF

Diagnostic sensitivity = Clinical sensitivity

Table 2: Examples of different Intended Purposes/Uses and how they drive the selection of clinical performance indicators, possible study populations, potential study designs, and IVD device examples.

Please note that this table does not provide a comprehensive or prescriptive selection of performance indicators, study populations, or study designs. It shows possible options of these clinical evidence concepts. It is the manufacturer's sole responsibility to define an appropriate clinical evidence concept. Furthermore, the demonstration of clinical utility is not a requirement according to 2017/746/EU. For the CDx Intended Use of Therapy Stratification or Therapy Selection, a clinical outcome study may be involved defining the clinical performance of the CDx in terms of the corresponding therapeutic.

9) Where should the manufacturer document the cut-offs/medical decision points?

As mentioned in the Q&A on Analytical vs Clinical Performance, IVDR mentions cut-off under analytical performance. Therefore, cut-offs should be documented in the analytical performance report, unless justified. The selection of a cut-off of a CDx device may require a) clinical (or surrogate) outcome data arising from prospective or retrospective trial data involving the therapeutic to be stratified or a comparator CDx device in case of a follow-on CDx.

10) What is the Clinical Benefit of a CDx device?

For the vast majority of (standalone) IVD devices, the clinical benefit focuses on the 'accurate medical information' output of an IVD device, in context of the Intended Purpose/Use as defined by the manufacturer and in conjunction with other medical information (see Q&A on Intended Purpose/Use). In contrast to standalone IVD devices, the clinical benefit and the corresponding clinical evidence of CDx IVD devices include the potential benefits as a result of treatment with the corresponding therapeutic product (i.e. clinical outcome; see also Figure 3 below).

Accordingly, recital (11) states "Companion diagnostics are essential for defining patients' eligibility for specific treatment with a medicinal product through the quantitative or qualitative determination of specific markers identifying subjects at a higher risk of developing an adverse reaction to the medicinal product in question or identifying patients in the population for whom the therapeutic product has been adequately studied and found safe and effective. Such biomarker or biomarkers can be present in healthy subjects and/or in patients."

Determination of safety and effectiveness is covered by the corresponding drug law.

11) What are typical examples of a CDx Clinical Benefit Assessment (according to IVDR 2017/746/EU Article 2 (37) and Recital 64)

The following clinical benefit assessment examples relate to the potential clinical benefit of a CDx-specific intended purpose/use of therapy stratification and/or therapy selection.

Clinical Benefit Assessment of a HER2 CDx Device (therapy stratification)

Based on the analytical and clinical performance, this IVD device achieves the clinical benefit of accurately detecting HER2 antigen in normal and neoplastic breast and gastric tissue and providing medical information about breast and gastric cancer patients for whom Anti-HER2 therapy is considered. In conjunction with histological examination, relevant clinical information, and proper controls, this information allows physicians to consider therapeutic interventions using anti-HER2 therapies per individual drug labels and/or clinical guidelines.

Clinical Benefit Assessment of a KRAS CDx Device (therapy stratification)

Based on the analytical and clinical performance, this IVD device achieves the clinical benefit of identifying CRC patients for whom treatment with cetuximab or with panitumumab may be indicated based on a no mutation detected result. In conjunction with relevant clinical information, this information allows physicians to consider therapeutic interventions per individual drug labels and/or clinical guidelines.

Clinical Benefit Assessment of a BRAF CDx Device (therapy stratification or selection)

Based on the analytical and clinical performance, this IVD device achieves the clinical benefit of selecting melanoma patients whose tumours carry the BRAF V600E or V600K mutation for treatment with trametinib. In conjunction with relevant clinical information, this information allows physicians to consider therapeutic interventions per individual drug labels and/or clinical guidelines.

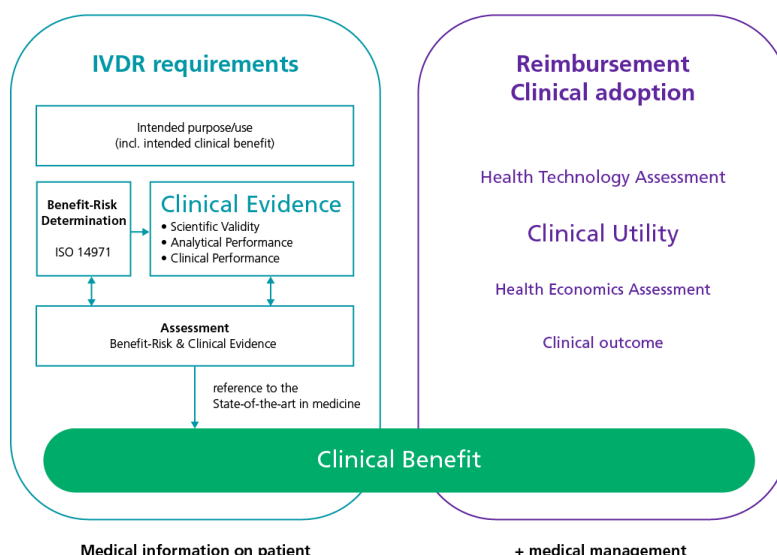


Figure 3. Clinical benefit and clinical utility concepts under the IVDR for CDx devices

The CDx-specific Intended Purpose/Use may require studying the IVD device together with the corresponding therapeutic with regards to the efficacy and safety of the therapeutic. Thus, the medical treatment and outcome of the patient need to be taken into consideration. Though clinical utility is not required for all IVDR, in this case the clinical utility of the therapeutic product (clinical outcome) is required for CDx because of their Intended Purpose. However, as for all IVDs, Health Technology Assessments or Health Economic Studies are not a requirement under IVDR. They are required for the therapeutic product.

12) What the Clinical evidence level considerations for CDx devices?

As for other IVD devices, evidence levels for *analytical performance* and *scientific validity* can be similar for various CDx devices. Similar to standalone IVD devices, the robustness and strength of the evidence should primarily relate to *clinical performance* and follow a risk-based approach. However, as all CDx devices are expected to be in class ‘C’, the strength and robustness of the clinical performance evidence for CDx is expected to be similar. Moreover, levels of available clinical evidence of CDx devices may depend of the

related therapeutic, the scientific validity of the test (including similarity of molecular diagnostic and therapeutic targets), the availability of similar or equivalent CDx devices, and the benefit risk ratio of the therapeutic product, and other factors influencing the risk of patients.

13) How much data is sufficient to demonstrate scientific validity?

As stated in Q&A on Scientific Validity, evidence is always needed to prove scientific validity. In the specific case of a CDx device, the evidence for the scientific validity of the product should include expression of the associated therapeutic product's clinical performance in the CDx-stratified or selected patient population, such as positive results of an interaction analysis of outcome measures that demonstrate the ability of the CDx device to stratify or select the therapeutic product.

14) What are the sources for clinical performance data?

Based on the Intended Purpose/Use of therapy stratification, CDx devices always require clinical performance data (omission cannot be justified). Specifically, they require evidence demonstrating that the CDx can successfully stratify or select the patients into responders or likely non-responders to the therapy in question. Demonstration of the clinical performance of a CDx device (i.e. the ability to select or stratify a therapeutic in support of the Intended Use) can be based on the following:

- Clinical performance studies that may include clinical outcomes (expression of therapeutic benefit and/or safety in IVD stratified or selected group)
- Concordance analysis between CDx and a comparative/predicate device, supported with statistical analysis of the therapeutic effect in the population defined by the CDx
- Real-world evidence generated using the CDx

As stated earlier, the Intended Purpose/Use of the IVD devices drives the clinical performance indicator. Some examples for CDx devices are shown in the Table 3 below.

IVD CDx Device	Function / Intended Purpose / Intended Use	Clinical Performance
HER2	Therapy stratification: aid in the assessment of breast and gastric cancer patients for whom Anti-HER2 therapy is considered.	Interaction analysis demonstrating that the CDx can successfully stratify the patients into responders or likely non-responders to Anti-HER2 therapy.
KRAS	Therapy stratification: aid in the identification of patients with colorectal cancer for treatment with cetuximab or	Interaction analysis demonstrating that the CDx can successfully stratify the patients into responders or likely non-

	panitumumab based on a no mutation detected test result.	responders to cetuximab or panitumumab therapy.
BRAF	Therapy selection: aid in selecting melanoma patients whose tumours carry the BRAF V600E or V600K mutation for treatment with trametinib Therapy selection	Expression of the drug performance in the population defined by the CDx.

Table 3. Examples of CDx IVD devices along with Intended Purpose and possible clinical performance.

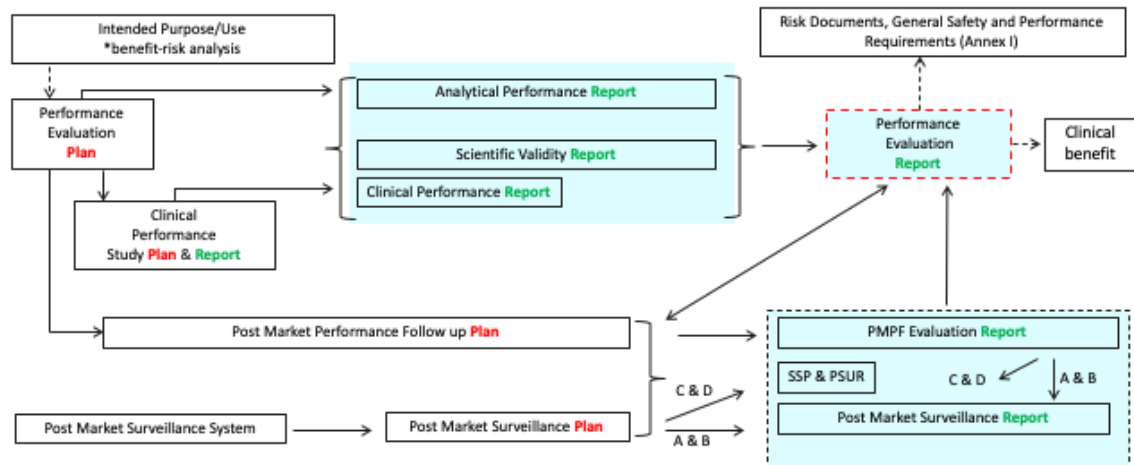
Please note that this table does not provide a comprehensive or prescriptive selection of Intended Purpose and clinical performance options.

References:

- 1) Regulation (EU) 2017/ 746 of the European parliament and of the council of April 5, 2017 on in-vitro diagnostic medical devices
- 2) Jorga A, Holt DW, Johnston A. Therapeutic drug monitoring of cyclosporine
- 3) Status of Companion and Complementary Diagnostics: Strategic Considerations for Development and Launch <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5355969/>
- 4) Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products. Center for Drug Evaluation and Research (CDER) and Center for Biologics Evaluation and Research (CBER), March 2019, Guidance for Industry
- 5) Exploratory IND Studies; Guidance for Industry, Investigators, and Reviewers. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), January 2006 Pharmacology/ Toxicology
- 6) ISO 20916:2019, In vitro diagnostic medical devices — Clinical performance studies using specimens from human subjects — Good study practice
- 7) GHTF/ GHTF/SG5/N8:2012 Clinical Performance Studies for In Vitro Diagnostic Medical Devices
- 8) Meijuan Li, Statistical consideration and challenges in bridging study of personalized medicine. J Biopharma Stat. 2015;25(3):397-407.
- 9) Meijuan Li, Statistical Methods for Clinical Validation of Follow-On Companion Diagnostic Devices via an External Concordance Study, Statistics in Biopharmaceutical Research, 2016 8:3, 355-363
- 10) JT Jørgensen and M. Hersom (Ann Transl Med. 2016 Dec; 4(24): 482
- 11) CLSI document EP09-A3: Method Comparison and Bias Estimation Using Patient Samples; Approved Guideline - Second Edition (Interim Revision). Wayne, PA: Clinical and Laboratory Standards Institute, 2013.
- 12) CLSI document EP06A: Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach; Approved Guideline. Wayne: NCCLS, 2012.
- 13) CLSI document H26-A2: Validation, Verification, and Quality Assurance of Automated Hematology Analysers; Approved Standard - Second Edition. Wayne, PA : Clinical and Laboratory Standards Institute, 2010.
- 14) CLSI document EP17-A2: Protocols for Determination of Limits of Detection and Limits of Quantification; Approved Guideline. Wayne, PA: s.n., 2012

CHAPTER 8 – Documentation of Performance Evaluation requirements

Annex XIII of the IVDR sets out the respective requirements for the plans and reports on Performance Evaluation and Post-Market Performance Follow up (PMPF). This document describes the flow of plans and reports (Figure 1), the required frequency for updating the reports, and seeks to clarify elements of the wording.



*A risk analysis should be performed before a performance evaluation is started so that requirements and acceptable limits can be established, this drives the Performance Evaluation Plan content.

Figure 1. Flow of Plans and Reports for Performance Evaluation.

The flowchart describes the relevant information that is required in the design control process. How this is documented and indexed will depend on the individual company documentation system.

Although the IVDR does not explicitly mention analytical study documentation, Annex XIII, Section 3 refers to studies other than clinical performance studies which shall be documented in the same way. Analytical performance study documentation is included in the performance evaluation plan and is therefore addressed in a similar manner as the clinical performance study plan and report.

The performance evaluation and its documentation shall be updated throughout the lifecycle of the device concerned with data obtained from the manufacturer's PMPF plan in accordance with Part B of Annex XIII and the post-market surveillance plan referred to in Article 79.

Table 1 below provides an overview of required frequency of different documents depending on the device class.

Device Class	Document	Required frequency of update	Article
All	Performance evaluation and associated documentation	Throughout the lifecycle of the device. From implementation of the manufacturer's PMPF plan in accordance with Part B of Annex XIII and the post-market surveillance plan referred to in Article 79	Article 56, section 6
A & B	Post Market Surveillance Report	When necessary and made available to the notified body and the competent authority upon request	Article 80
C & D	Periodic Safety Update Report (PSUR)	At least annually	Article 81, section 1
	Performance Evaluation Report	As necessary and at least annually	Article 56, section 6
	Summary of Safety and Performance (SSP)	As soon as possible, where necessary	Article 56, section 6

Table 1. Required frequency of updates of reports

1) What level of performance evaluation documentation will Notified Bodies expect for established products?

The same information will be required for established products as other products. For established products it is reasonable to refer to existing documents instead of generating a new performance evaluation plan.

2) Annex XIII, section 1.1 states 'As a general rule, the performance evaluation plan shall include at least'. What is meant by 'As a general rule'?

The text states 'As a general rule', indicating that some points may be excluded as long as a justification is given.

3) Annex XIII, section 1.1, 10th indent: Why should a benefit-risk analysis be performed before a performance evaluation is started (required to be referenced as part of the plan)?

The benefit-risk analysis according to ISO 14971² is intended to determine if the medical/clinical benefits of the intended use outweigh the overall residual risk.

4) Annex XIII, section 2, Clinical Performance Studies: Where can I find additional information on how to conduct clinical performance studies?

See the new ISO 20916³ for additional information.

5) Annex XIII, section 2.1. What are the criteria that determine whether a clinical performance study is needed?

When clinical performance is applicable, then in the absence of sufficient clinical performance data, a clinical performance study shall be performed to supplement the available clinical performance data from other sources, such as literature and experience from routine diagnostic testing.

6) Annex XIII, section 2.3.2(a), single identification number of the clinical performance study: Does this requirement apply to all studies?

No, this requirement only applies to Annex XIV studies as these cover interventional performance studies and certain other performance studies as referred to in Article 58 (1) and (2).

7) Annex XIII, section 2.3.2(h): Where should the benefit-risk analysis be documented?

The benefit-risk analysis will be a part of the risk management report and should be referred to in the Performance Evaluation Plan (PEP) and Performance Evaluation Report (PER). PEP/R can refer to the risk management report according to ISO 14971.

8) Annex XIII, section 2.3.2 (o), monitoring plan: Does this refer to data integrity and/or the monitoring of patients?

This refers to the monitoring of study conduct (e.g. follow the CPSP, integrity of data, adequate qualification of personnel conducting the study). For additional information, please consult ISO 20916.

9) Annex XIII, section 2.3.2 (p), data management: What does this refer to?

This is referring to the process of how the data will be captured and managed. Where relevant, it would be appropriate to state how the requirements of the GDPR⁴ are being met within the data management process. For additional information, please consult ISO 20916.

10) Annex XIII, section 2.3.3: Where can additional guidance be found on the structure and content of the clinical performance study report?

ISO 20916 can provide additional guidance on the conduct of a clinical performance study.

11) Annex XIII, section 3, Other Performance Studies: Is this referring to analytical performance studies? If 2.3.2 structure is used for analytical performance study plans, can all listed items be applicable?

There is no clear indication of additionally required performance studies in the Regulation. Clinical and analytical performance studies require individual reports using similar headings and structure. The level of detail may vary between analytical and clinical performance study reports. Therefore, depending on the analytical performance study, it would be reasonable to state which parts are relevant rather than listing all parts that are not relevant.

12) Do analytical and clinical performance study reports need to be signed?

Yes, both reports need to be signed by competent/authorised persons and are part of the Design Control Management System.

References:

- 1) Regulation (EU) 2017/ 746 of the European Parliament and of the Council of April 5, 2017 on in-vitro diagnostic medical devices
- 2) ISO 14971 Medical Devices – Application of risk management to medical devices
- 3) ISO 20916 In vitro diagnostic medical devices – Clinical performance studies using specimens from human subjects – Good study practice
- 4) Regulation EU 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons regarding the processing of personal data and on the free movement of such data (General Data Protection Regulation)

CHAPTER 9 - Summary of Safety and Performance

The Summary of Safety and Performance (SSP) is one of the requirements of the new Regulation, specific for class C and D devices, to enhance transparency and adequate access to information. It intends to provide public access to summarised data on the safety and performance of class C and class D IVD devices to all intended users – professionals and lay persons.

The present document aims at guiding manufacturers where relevant information for the different SSP requirements of Article 29 can be found in the manufacturer's documentation. The template below offers possible sources for the SSP. It by no means replaces the EUDAMED template or mandates the format of the SSP. It is the manufacturer's sole responsibility to document the SSP in an appropriate manner, fulfilling the requirements of Article 29 of the IVDR.

Requirements based on IVDR Article 29	Potential regulatory sources
Device identification and general information	
Name or trade name including any model number or version	<Excerpt from IFU or declaration of conformity>
Manufacturer (name and address)	<Excerpt from label or declaration of conformity >
Manufacturers single registration number (SRN), if available	<Excerpt from declaration of conformity >
Basic UDI-DI	< Excerpt from declaration of conformity >
Intended purpose of the device	
Intended purpose and indications	<Excerpt from IFU >
Target populations	<Excerpt from IFU > A clear specification of indications. Description of target groups shall be specified, e.g. age, gender, specific medical conditions, etc.
Contraindications (limitations)	<Excerpt from IFU or clinical evidence report > Annex I, 20.1 (g) Limitations (medical and technical)
Device description	
Device description	< Excerpt from IFU and summary of Technical Documentation, Annex II, 1.1, as appropriate > To include e.g. operating principles
Reference to previous generation(s) or variants of the device (as applicable) and a description of the differences	<Excerpt from IFU, technical documentation (Annex II, 1.2a) >

	Should include e.g. differences of the operating principles (e.g. manual vs automated); any novel features
Description of accessories intended to be used in combination with the device (as applicable)	<Excerpt from IFU, if exists, of the accessory, technical documentation (Annex II, 1.1.m) >
Description of other devices and products intended to be used in combination with the device (as applicable)	<Excerpt from IFU, if exists, of the other devices; technical documentation (Annex II, 1.1.m) >
Standards Reference	
Harmonised standards and Common Specifications (CS) applied	Provide a list of applicable CS and harmonised standards. If CS exists for the device in question, provide a reference to the CS that can be found in the Declaration of Conformity. The SSP can also include the monographs of the EU Pharmacopoeia adopted in accordance with the Convention on the Elaboration of the European Pharmacopoeia, if references to those monographs have been published in the OJEU.
Summary of the performance evaluation and Post-Market Performance Follow-Up	
<Excerpt from Performance evaluation report including PMPF section (Annex XIII, 1.3.2.)> This shall be an objective, balanced summary from the performance evaluation report that is written in a comprehensive and traceable manner, including relevant aspects of safety and performance, conclusions from benefit-risk analysis, and a statement regarding whether equivalence was used in the assessment of the conformity of the device. The summary shall be provided in appropriate terminology understandable to the respective intended user(s) of the device.	
Metrological traceability	
Metrological traceability of assigned values	<excerpt from IFU, Annex I, 20.4.1 (u) > Where performance of a device depends on calibrators and/or control materials
Users	
User Profile	<Excerpt from IFU, annex I, 20.4.1 (e) >
User Training	<Excerpt from IFU, Annex I, 19.1; 20. 4.1(p) >
Device Risks Information	
Residual risks and undesirable effects	<Excerpt from IFU (Annex I, 20.4.1 (n)) and from Risk analysis documentation >

Warnings and precautions	<Excerpt from IFU (Annex I, 20.4.1 (n) and Safety Data sheet >
--------------------------	--

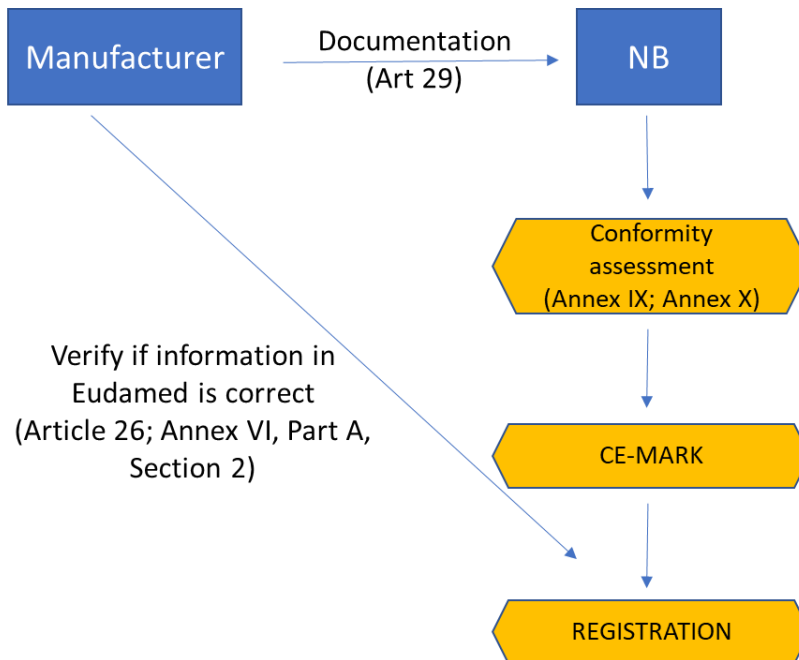
Table 1: Possible sources for the SSP

1) Which sources can be used for the SSP?

The content in this document shall be sources from the technical documentation (Annex I and Annex II), the EU Declaration of Conformity and may be identical to some parts of the instruction for use (IFU). However, this document is not intended to substitute the IFU. The present document is a *summary* of safety and performance, therefore all entries, especially the part on performance evaluation, shall be provided in a concise and summarised form rather than include detailed reports.

2) Who should upload the SSP?

The manufacturer should submit a draft SSP, as part of the application documents, to the Notified Body (NB) involved in the conformity assessment (Annex IX and X). After issuing the certificate, the NB will upload the validated SSP in EUDAMED. Before uploading the SSP, the NB will verify that all required elements are covered in the SSP and that the information provided in the draft SSP conforms with the technical documentation assessed under the conformity assessment process. Upon receiving the CE- certification and before the device can be placed on the market, the manufacturer shall verify in EUDAMED the information related to the device, including the SSP (Article 26; Annex VI, Part A, Section 2.11).



3) What is the frequency of update?

Article 56 (6): 'The Summary of Safety and Performance shall be updated as soon as possible, where necessary', suggesting that it should be updated only if the manufacturer's post-market surveillance (including PMPF) identifies any issues that will lead to a change in the technical documentation rendering the information in the SSP outdated. However, if no changes have been found, the SSP shall remain unchanged regardless of the frequency of updates to any reports that may constitute the SSP.

CHAPTER 10 – Post-Market Performance Follow-up

Post-Market Performance Follow-Up (PMPF) is a continuous process that updates the performance evaluation referred to in Article 56 and Part A of Annex XIII and shall be addressed specifically in the manufacturer's post-market surveillance plan. When conducting PMPF, the manufacturer shall proactively collect and evaluate performance and relevant scientific data from the use of a device which bears the CE marking and is placed on the market or put into service within its intended purpose as referred to in the relevant conformity assessment procedure. The PMPF aims to confirm the safety, performance and scientific validity throughout the expected lifetime of the device, to ensure the continued acceptability of the benefit-risk ratio and to detect emerging risks on the basis of factual evidence. Figure 1 describes how PMPF relates to other elements of the IVDR.

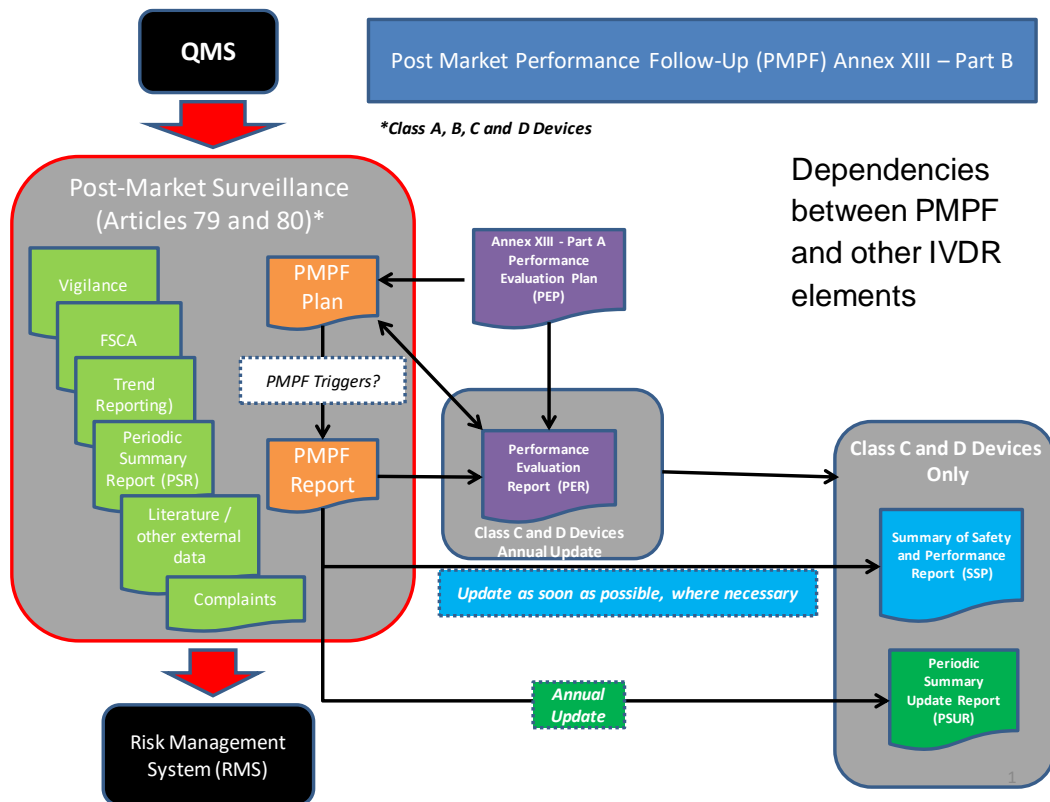


Figure 1. Dependencies between PMPF and other IVDR elements

1) What should be included in the PMPF and where can this information be found?

Annex XIII, part B describes the requirements for PMPF. The PMPF shall be planned and performed as deemed required by the manufacturer and as documented in the manufacturer's PMPF plan. Table 1 gives examples of what should be included as the general methods and procedures. The PMPF plan shall describe the specific methods and procedures, rationale for method and procedure appropriateness, and the objective and frequency/timeline. Post-market studies may be included as a specific method and procedure in the PMPF plan. References to relevant Common Specifications harmonised standards consulted and relevant PMPF guidance shall also be listed, as well as a reference to the relevant parts of the performance evaluation report referred to in the IVDR Section 1.3 of Annex XIII and to the risk management referred to in Section 3 of Annex I.

Elements potentially overlapping with the periodic safety update report (PSUR) or post-market surveillance report, such as scientific literature evaluation or complaint data, may be available through these reports.

Overall objectives of the PMPF are to:

- Confirm the safety, performance and scientific validity of the device throughout the expected lifetime
- Identify systematic misuse^{††};
- Identify new safety issues;
- Analyse benefit/risk ratio;
- Identify new risks;
- Identify limits to performance and, if applicable, contra-indications; and
- If applicable, review the performance data relating to equivalent or similar devices, and the current state of the art.

In addition, any product specific objectives (e.g. sourcing of rare samples) will be included in the PMPF plan.

Note: Misuse should not be confused with "Use Error", which is defined in MEDDEV 2. 12-1 (Guidelines on a medical devices vigilance system) as "*Act or omission of an act, that has a different result to that intended by the manufacturer or expected by the operator of the medical device*"³. Use Error would be handled through the normal Post-Market Surveillance vigilance system of the manufacturer.

^{††} IVDR includes provisions for manufacturers around systematic misuse and reasonably foreseeable misuse. Modification of a device that is subject to the requirements of the exemption including appropriate performance study does not constitute foreseeable or systematic misuse. The modification and use of the device should be verified against the original device when used as intended by the manufacturer to demonstrate and document whether the function, performance or purpose has been altered. Modification could include using an existing device for a purpose not intended by the manufacturer, modifying a device for a new purpose, use of sample types, accessories or components or combining devices not specified by the manufacturer. Therefore, off-label use may also be a modification or manufacture and the exemption requirements would apply ². An example of misuse is using HIV monitoring assays for screening of blood bags. Systematic misuse is different to use error, as described in MEDDEV guidance³.

Table 1. PMPF plan template example – general elements and examples.

Please note that this table does not provide a comprehensive or prescriptive section of elements and methods. It is the manufacturer's sole responsibility to define an appropriate concept.

General methods and Procedures	Specific methods and procedures	Rationale for method and procedure appropriateness	Objectives	Frequency / timeline
Scientific literature evaluation ^	Conduct literature search according to specified methodology. Evaluate new guidelines (e.g. technical or medical guidelines)	This method will provide the relevant scientific information on the biomarker and test. This method will also provide information on similar devices/state of the art	If applicable, review the performance data relating to equivalent or similar devices, and the current state of the art Verify that product claims are met Identify systematic misuse Identify safety issues Identify new limitations and contra-indications	Product class-dependent. TBD by the manufacturer
Feedback from users	Evaluate customer complaint data ^ Evaluate published data on user perspectives. Information from sales and training (e.g. surveys)	These methods will raise potential issues experienced by product users	Verify that product claims are met Identify systematic misuse Identify of new risks Identify new limitations and contra-indications	Product class-dependent. TBD by the manufacturer

Gathering of clinical experience gained	Post-market study data generation	Post-market studies will allow further collection of safety and performance data, including large-scale data where applicable	Verify that product claims are met Identify safety issues Analyse the benefit/risk ratio Identify new risks Identify new limitations and contra-indications	Product class-dependent. TBD by the manufacturer
	<ul style="list-style-type: none"> - Conduct company-sponsored or investigator-initiated post-market studies - Evaluation of patient registers, where applicable 	Applicable, for example, in circumstances when additional clinical data is required to support claims of pre-launch data, such as for rare samples or where only retrospective samples have been available for pre-market studies		
	Evaluation of published experience gained by routine diagnostic testing Evaluation of specific results, such as patient mean results	These methods will allow further collection of safety and performance data	Verify that product claims are met Identify safety issues Analyse the benefit/risk ratio Identify new risks Identify new limitations and contra-indications	Product class-dependent. TBD by the manufacturer
	External / Internal Quality Assessment data generation Conduct external quality assessments at selected laboratories/customer sites, e.g. ring trials	This method will allow further collection of analytical performance data	Verify that product claims are met	Product class-dependent. TBD by the manufacturer

^ Examples where PSUR data or post-market surveillance report data can be utilised, where available.

2) What are appropriate timelines for PMPF report updates?

The PMPF plan and/or triggers will determine the frequency/timeline of the PMPF update for a device. Accordingly, PMPF can be performed based on pre-planned dates and/ or based on certain triggers, which will be defined in the PMPF plan, see question 3. The frequency of PMPF shall be determined by the manufacturer and the rationale for this shall be described in the PMPF plan. For class C and D products, the PMPF report shall be updated annually⁴ with important updates and the PMPF main findings will be included in the periodic safety update report (PSUR). If no action has been required according to the PMPF plan, for example, in instances where no triggers have occurred, nothing further is required, and this will be stated in the PMPF report update. If the manufacturer concludes no PMPF is required for a device, a justification for this shall be provided and documented within the performance evaluation report.

3) What elements can be pre-specified triggers for PMPF?

Pre-specified results can trigger additional tasks and activities. Pre-specified triggers for PMPF activities are based on their impact on product claims and benefit-risk and can include customer complaints, emergence of data from e.g. publications, external quality assessment programs.

For example, the emergence of new mutations or interference from medicinal products will likely trigger PMPF. The IVDR states that relevant new information should trigger a reassessment of the clinical evidence of the device thus ensuring safety and performance through a continuous process of performance evaluation⁵. Relevant data and information gathered through post-market surveillance, as well as lessons learned from any implemented preventive and/or corrective actions, should be used to update any relevant part of technical documentation, such as those relating to risk assessment and performance evaluation, and should also serve the purposes of transparency⁶.

4) What IVDR elements are linked to PMPF and what are the dependencies between these?

The PMPF plan is part of the Performance Evaluation Plan (PEP), and the PMPF evaluation report forms part of the performance evaluation report (PER). PMPF is included in post-market surveillance (PMS), and the PMPF shall be specifically addressed in the manufacturer's PMS plan. Relevant information on the PMPF shall be included in the Summary of Safety and Performance (SSP), which shall be updated as soon as possible, where necessary. The Periodic Safety Update Report (PSUR) shall also contain the main findings of the PMPF and shall be part of the technical documentation. The dependencies between PMPF and other IVDR elements are illustrated in Figure 1 and Table 2 in this Q&A document. The Q&A on Documentation further describes the flow of plans and reports.

Table 2. PMPF and PMS requirements

	A	B	C	D
POST-MARKET				
Post-Market Surveillance Plan	X	X	X	X
Post-Market Surveillance Report	X	X		
Periodic Safety Update Report			X	X
PMPF Plan	X	X	X	X
PMPF Report	X	X	X	X
Performance Evaluation Report	X	X	X	X
Summary of Safety and Performance	X	X	X	X
VIGILANCE				
Manufacturer Incident Report	X	X	X	X
Periodic Summary Report	X	X	X	X
Trend Report	X	X	X	X
Field Safety Corrective Action	X	X	X	X
Field Safety Notice	X	X	X	X

Submitted to EUDAMED

PMPF confirms safety and performance of the device throughout its expected lifecycle

- Previously unknown risks or limits to performance and contraindications
- Emergent risks on basis of factual evidence
- Continued applicability of the clinical evidence and of the benefit-risk ratio
- Possible systematic misuse

Periodic Safety Update Report (PSUR)

- Conclusions of the benefit-risk determination
- Main findings of the PMPF
- Volume of sales of device and an estimate of the size and other characteristic of the population using the device
- Usage frequency of the device if practicable

PMPF Plan and PMPF Report are used to update the Performance Evaluation Report

- Justification of approach taken to gather clinical evidence
- Literature search methodology and protocol
- Technology on which the device is based, intended purpose of the device and performance and safety claims
- Nature and extent of scientific validity and analytical and clinical performance data that has been evaluated
- Clinical evidence as the acceptable performance against the state of art in medicine

5) In what instances is PMPF not deemed appropriate?

Post-market surveillance is a requirement of the regulation, whereas PMPF activities may not be required where other PMS activities do not identify any triggers, such as for products where foreseeable or actual changes are less likely to negatively impact the benefit-risk ratio. If PMPF is deemed not appropriate, a justification shall be provided in PER (IVDR, Annex XIII, Part B (8)).

- Class A - IVD Instrument – stand-alone:
 - o Justification: Performance is typically related to reagents running on the instruments; other PMS activities (see Figure 1) should be sufficient to monitor performance
- Class A - Washing solution – separate, not included in IVD test/kit:
 - o Justification: Performance is typically related to the IVD test/kit. PMS activities of the IVD test/kit should be sufficient to monitor performance
- Class B and C – Established and Standardized tests on the market:
 - o Justification: Sufficient data from other devices available to mitigate the risk so that other PMS activities should be sufficient to monitor performance

Appendix 10.1: Post-market Performance Follow-up Plan

Example 1

Date and Version	13 August 2019 / Version 001
Name of the Device	HIV Ab-Ag combo Assay
Class	D
Intended Use	Semi-quantitative enzyme immunoassay kit for the detection of HIV-1 p24 antigen and antibodies to HIV-1 (groups M and O) and HIV-2 in human serum or plasma. This kit can be used for both HIV Ag and HIV Ab screening of blood donations and as an aid in the diagnosis of HIV infection.

Aim:

- Verify Clinical Safety and Performance over expected lifetime
- Identify previously unknown risks or limits to performances and contra-indications
- Identify and analyze emergent risks on the basis of factual evidence
- Ensure continuous acceptability of the clinical evidence and the benefit risk ratio
- Identify possible systematic misuse

Benefit /risk ratio: Refer to “Product” Risk Management Plan document

Clinical Evidence, Performance: Refer to “Product” PER document

Performance of equivalent or similar devices and the current State of the Art: Refer to “Product” State of the Art Report document

References:

- CTS 2009/886/EC CS:
- Standards:

PMPF Time Schedule

The data will be reviewed each year and gathered in a report according to table 3 (PMPF plan example 1)

Table 3. PMPF plan example 1

Examples - General methods and procedures	Specific methods and Procedures	Rationale for method and procedure appropriateness	Objectives	Frequency / timeline
Clinical experience gained	Collecting additional data from internal/external studies	To collect new performance information on the product	Appraise the sensitivity and specificity results	If new sample panels (seroconversion, sensitivity panels) are identified and available Or new standard (ex WHO standard)
Clinical experience gained	Collecting additional data from internal/external studies	To collect new performance information on the product	Appraise the specificity and results	If complaints linked to specificity performance
Clinical experience gained	Conducting a post-market clinical study according Annex XIII IVDR /ISO 20/916	To collect new performance information on the product	Appraise the specificity or sensitivity results in other countries (with different prevalence, and different subtypes)	If new variants identified and available
Scientific literature search *	SOP on literature search	To collect new scientific information on the targeted marker	Look at new variants, subtypes	Regular literature survey
	SOP on literature search	To collect new performance information on the product, on similar competitor products	Appraise the specificity or sensitivity results	Regular literature survey
Feedback from users ^	Investigate the data linked to the event	Complaint linked to performance	Improve sensitivity or specificity performances	Depending of occurrence of the event

^ This information may be extracted from the PSUR report data or post-market surveillance report data can be utilised, where available

Example 2

Date and Version	13 August 2019 / Version 001
Name of the Device	Influenza A & B rapid diagnostic test
Class	C
Intended Use	Immunochromatographic assay for the qualitative detection of influenza A and B nucleoprotein antigens in nasopharyngeal (NP) swab and nasal swab specimens.

Aim:

- Verify Clinical Safety and Performance over expected lifetime
- Identify previously unknown risks or limits to performances and contra-indications
- Identify and analyze emergent risks on the basis of factual evidence
- Ensure continuous acceptability of the clinical evidence and the benefit risk ratio
- Identify possible systematic misuse

Risk management: Refer to “Product” Risk Management Plan document

Clinical Evidence, Performance: Refer to “Product” PER document

Performance of equivalent or similar devices and the current State of the Art: Refer to “Product” State of the Art Report document

References:

- Standards:

PMPF Time Schedule

The data will be reviewed each year and gathered in a report according to table 3 (PMPF plan example 2)

Table 4. PMPF plan example 2

Examples General Methods and Procedures	- Specific methods Procedures	and	Rationale method procedure appropriateness	for Objectives and	Frequency / timeline
---	-------------------------------------	-----	---	-----------------------	----------------------

Clinical Experience gained	Internal studies and / or post-market external clinical studies	Internal and / or external studies may be conducted to validate that the product continues to meet the product claims.	Verify that product claims are met	If product complaints emerge, or if information becomes available regarding new mutants or cross-reactants that have not previously been validated with the test
Scientific literature search*	To collect new scientific information that is relevant for test performance, such as new mutants. To collect information on similar competitor products	SOP on literature search	Verify that product claims are met Identify safety issues Analyse the benefit/risk ratio Identify new risks Identify new limitations	Regular literature survey
Feedback from users*	Evaluate customer complaint data	This method will raise issues with products in the field	Verify that product claims are met Identify safety issues Analyze the benefit/risk ratio Identify new risks Identify new limitations	Customer complaint data will be monitored continuously through PMS activities

* This information may be extracted from the PSUR report data or post-market surveillance report data can be utilised, where available.

Appendix 10.2: Post-market Performance Follow-up Report

Date and Version
State the PMPF plan date and version
State the PMPF report date and version

Device identification
Name:
Classification:
Intended use:

Results
State the results (for key elements see PMPF plan)

Conclusion(s)
State the conclusion(s) and if needed action items, such as CAPA

References:

- 1) Regulation (EU) 2017/ 746 of the European Parliament and of the Council of April 5, 2017 on in-vitro diagnostic medical devices
- 2) MHRA Draft guidance on the health institution exemption (HIE) – IVDR and MDR, draft v. 0.2, December 2017
- 3) MEDDEV 2 12-1 Rev 8, January 2013
- 4) Article 56 paragraph 6, Regulation EU 2017/ 746 of the European Parliament and of the Council of April 5, 2017 on in-vitro diagnostic medical devices
- 5) Recital 63, Regulation EU 2017/ 746 of the European Parliament and of the Council of April 5, 2017 on in-vitro diagnostic medical devices
- 6) Recital 75, Regulation EU 2017/ 746 of the European Parliament and of the Council of April 5, 2017 on in-vitro diagnostic medical devices

Other useful reference documents:

- 1) Regulation (EU) 2017/ 746 of the European Parliament and of the Council of April 5, 2017 on in-vitro diagnostic medical devices
- 2) ISO 20916 In vitro diagnostic medical devices – Clinical performance studies using specimens from human subjects – Good study practice
- 3) ISO/TC 210/WG 6 (Working Group 6): Application of post market surveillance systems to medical devices
- 4) GHTF/SG5/N7:2012 Clinical Evidence for IVD medical devices – Scientific Validity Determination and Performance Evaluation

About MedTech Europe

MedTech Europe is the European trade association for the medical technology industry including diagnostics, medical devices and digital health. Our members are national, European and multinational companies as well as a network of national medical technology associations who research, develop, manufacture, distribute and supply health-related technologies, services and solutions.

For more information, visit www.medtecheurope.org.

For further information on the content of this publication, please contact:

Oliver Bisazza

Director Regulations and Industrial Policy

MedTech Europe

o.bisazza@medtecheurope.org

MedTech Europe Clinical Evidence Working Group



The *In-Vitro* Diagnostic Medical Devices Regulation contains several provisions that are capable of being given more than one interpretation. In the preparation of this series of Questions and Answers, MedTech Europe has used its best efforts to ensure that the opinions and advice expressed are sound. However, the Association makes no assertion that those opinions and advice are correct and it accepts no legal responsibility for them. Specific legal advice should be sought before acting on any of the topics covered. MedTech Europe reserves the right to change or amend this document at any time without notice in order to keep the information up to date.

Members are reminded that, while competent authorities and notified bodies may be helpful in providing views as to the meaning of the 2017/746/EU Regulation, it is ultimately for the courts to interpret legislation.