MedTech Europe
from diagnosis to cure

Clinical Evidence Requirements under the EU *In Vitro* Diagnostics Regulation (IVDR)

AUGUST 2025

MedTech Europe Regulatory eBook
Clinical Evidence Requirements under the EU *In Vitro* Diagnostics Regulation (IVDR)

Welcome to the Fourth Edition!

This eBook is a collection of questions and answers designed to support manufacturers in navigating their performance evaluation obligations under the *In Vitro* Diagnostic Medical Devices Regulation (IVDR) (EU) 2017/746.

The content reflects the collective expertise of numerous regulatory and clinical professionals, including both members of MedTech Europe and a range of external experts and contributors. It has been developed as a practical resource, grounded in the real-world experience of the IVD industry and enriched by continuous feedback from Notified Bodies.

We are pleased to present the **Fourth Edition** (August 2025), which builds on the foundation of previous versions and incorporates the latest insights and developments in the field. This edition includes the following updates:

- Chapter 7 on Companion Diagnostics has been revised, now featuring new questions that reflect the latest trends and stakeholder feedback.
- Chapter 14 on Clinical Evidence for Next Generation Sequencing (NGS)-based IVD Assays, with a particular focus on oncology applications.
- Chapter 15 on Real-World Evidence, acknowledging its growing importance in IVD development and assessment.

This edition was coordinated by MedTech Europe, Iana Slobodeaniuc and Alice Bova, with valuable support from MedTech Europe's Clinical Evidence WG co-chairs, Christian Zaugg and Megha Iyer. It was shaped through recent discussions with experts from a leading IVD-designated Notified Body. Additionally, references and standards have been revised to ensure alignment with current best practices.

We hope this updated resource continues to serve as a practical and insightful guide for IVD manufacturers, regulators, Notified Bodies, and other stakeholders committed to advancing innovation and ensuring patient access to safe and effective diagnostics.

As with previous editions, MedTech Europe is pleased to offer the Fourth Edition of Clinical Evidence Requirements for CE Certification as a free download on its website: Clinical Evidence Requirements for CE certification - MedTech Europe.

We welcome your feedback and hope you find this edition valuable. For any inquiries or suggestions, please contact Iana Slobodeaniuc at regulatory@medtecheurope.org

For EU legislation please consult the latest consolidated version.

Short historical context to the ebook

The **First Edition** was published in 2020, when there was still limited hands-on experience with the application of the IVDR. As regulatory understanding and implementation have matured over time, this document has evolved accordingly.

The **Second Edition**, released in November 2021, achieved great success, with over 6,000 downloads within its first year. Its widespread adoption underscored the eBook's growing influence within the IVD community. The Second Edition marked a turning point, solidifying the eBook as a highly influential resource and setting new trends in the industry. It quickly became clear that the eBook had established itself as an essential tool for stakeholders across the board, demonstrating its power through broad usage and significant impact.

The **Third Edition** of the eBook (February 2023) marked a comprehensive revamp, introducing additional examples, enhanced clarity, improved flow, and updated references and diagrams. To achieve these updates, MedTech Europe commissioned an independent expert review of the chapters from the **Second Edition** of *IVDR Clinical Evidence Requirements*, as well as new chapters under development. This review was led by **Steve Lee**, an independent expert, with significant contributions from MedTech Europe's **Iana Slobodeaniuc**, and Clinical Evidence Working Group co-chairs, **Volker Franzen**, and **Christian Zaugg**.

Authors list

Astola Annika
Adam Neil
Baker Amanda
Batchen Ashleigh (representing BIVDA)
Belonogova Marina
Bhatia Ramya
Bova Alice
Bruinsma Anne Marie
Callaerts Geert
Cardoso Rute Rodrigues
Chaube Amita
Cheillan Frank
Choudhary Mayank
Connell Barbara
Deviprasad Iyer Megha
Ekholm Pettersson Frida
Facheris Luisa
Forssten Camilla
Gannon Allison
Gazin Muriel
Giroud Claude
Goossens Dirk
Homann Anke
Hughes Karin
Hughes Richard
Javey Mana
Kasturi Roshni
Katarzyna A Koscielska
Li Lily
Lindroos Hanne
Louw Gail
Love Joanna
Magana Laura
Malcus Carine
Mamadou Norbert
Masson Christine
McBride David
Mechthild Merz
Mescalchin Alessandra
Meyerovich Kira
Miazga Natalie
Neumann Vanessa
Ons Benny
Partheniou Faidra
Percivati Stefania
Petruschke Thorsten
Plenert Karli
Plumridge Neil
Punwani Divya
Rousseau Els
Rutter Andrew
Saunders Richards
Slobodeaniuc Iana
Smartt Sherrie
Steenhuis Pieter
Sweeny Maranna
Therese de la Torre
Thottakam Bensita
Timonen Anne
Torbjörn Johansson
Van den Eede Peter
Van doan Nguyen
Wettmarshausen Sascha
Wevelsiep Anja
Wierzba Mike
Yates Alton
Ylianttila Mervi
Young Emma
Zaugg Christian
Ziegler Saskia
Zoellner Petra

TABLE OF CONTENTS

Introduction

This document is a Q&A guide to performance evaluation requirements of the *In Vitro* Diagnostic Medical Devices Regulation (EU) 2017/746.

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 add a CE mark to their product. Since the 1990s, *in vitro* diagnostic medical devices (IVDs) have been regulated by an EC Directive (IVD Directive (EC) 98/79). Since May 2022, the *In Vitro* Diagnostic Medical Devices Regulation (EU) 2017/746 (IVDR) fully applies. Most IVDs are able to benefit from a three to five years period of extended transition to the IVD Regulation. During this time, all IVDs will gradually transition to the IVD Regulation. MedTech Europe, the European trade association representing the IVD industry, works with our members and the authorities to support companies in complying with the IVDR.

The IVDR contains several provisions that are open to more than one interpretation. This brochure is designed to help stakeholders understand the IVD Regulation. 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.

Disclaimer

This document represents the understanding of MedTech Europe about the covered topics at the time of publication, and while we have invested considerable time and effort in developing this document, MedTech Europe does not assert that these opinions and interpretations 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 3, section 20.4.1 The Instructions for use shall contain all of the following particulars: The following table refers exclusively to **(c) the device's intended purpose** as the basis for the clinical evidence. Further requirements are not considered in the Instructions for use.	
(i)	What is detected and / or measured;
(ii)	The device's function (e.g. screening, monitoring, diagnosis or aid to diagnosis, prognosis, prediction, companion diagnostics);
(iii)	The specific information that is intended to be provided in the context of: – a physiological or pathological state; – congenital physical or mental impairments; – the predisposition to a medical condition or a disease; – the determination of 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 reiterated 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 3. It should also be noted that the intended user is not formally required to be part of the intended purpose under the instructions for use. However, the intended user shall be provided with the intended purpose under technical documentation.

IVDR Annex I, Chapter 3, section 20.4.1 'The instructions for use shall[1] contain all of the following particulars' ***(c) the device´s intended purpose***	***IVDR Annex II, 1.1 'Device description and specification'*** ***(c) 'the intended purpose of the device which may[1)] include information on'***
(i) The specific information that is intended to be provided in the context of: – 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; *IVDR Annex I, Chapter 3, section 20.4.1 (c)*	(iii) The specific disorder, condition or risk factor of interest that it is intended to detect, define or differentiate *Annex II, 1.1 (c) 'the intended purpose of the device which may include information on'*
The intended user, as appropriate (e.g. self-testing, near patient and laboratory professional use, healthcare professionals); *Annex I, Chapter 3, 20.4.1 (e)*	(viii) The intended user *Annex II, 1.1 (c) 'the intended purpose of the device which may include information on'*
For companion diagnostics, the International Non-proprietary Name (INN) of the associated medicinal product for which it is a companion test. *IVDR Annex I, Chapter 3, section 20.4.1 (c)*	(ix) For companion diagnostics, the relevant target population and the associated medicinal product(s) *Annex II, 1.1 (c) 'the intended purpose of the device which may include information on'*

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.

[1] 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

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 1, 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 1, 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/N045:2008[3] 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 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 Definitions)

- ISO 18113-1:2022 *In vitro* diagnostic medical devices. Information supplied by the manufacturer (labelling)[5]. Part 1: Terms, definitions and general requirements

3.1.37 '**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 section in Annex I does not specify a mandatory structure / layout. Therefore, how the applicable 'intended purpose / use' elements are presented in the instructions for use depends on the manufacturer's concept of instructions for use. For example, these elements may be distributed over several sections or combined in one (depending on discussions with your Notified Body). If they are not combined, it may be helpful to describe where the applicable elements can be found, for audit purposes. In case of new products, it is recommended to present this in one section.
Annex I, Chapter 3, section 20.4.

For more guidance on the intended purpose, MedTech Europe's members may consult the internal guidance (available exclusively for MedTech Europe via this link).

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 instructions for use 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

Ambiguous or misleading claims about the intended purpose of the device may lead to a higher classification and should be avoided. Any limitations to the intended purpose of the product (that is, what the device is NOT intended for) should be clearly stated (link to classification guidance "MDCG 2020-16 rev.4 Guidance on Classification Rules for *in vitro* Diagnostic Medical Devices under Regulation (EU) 2017/746").

The performance characteristics of the device should be suitable for the intended purpose taking account of the generally acknowledged state of the art. Performance characteristics may have been established in Harmonised Standards or Common Specifications, *IVDR Annex I para 9:* or adapt solutions that ensure a level of safety and performance that is at least equivalent thereto.

Further, manufacturers may wish to establish performance characteristics through e.g., Target Product Profiles. WHO and FIND offer descriptions of Target Product Profiles at https://www.who.int/ and https://www.finddx.org/

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
- Its function (see Table 1)
- The specific information set out in Tables 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 3, section 20.4.1c; Annex II 1.1.c; Annex XIII Part A and B

It is the manufacturer's sole responsibility to define an appropriate clinical evidence concept based on the 'intended purpose / use' and the environment where the product is used.
*For more information about different levels of clinical evidence, see Chapter 4. See below for a non-exhaustive list of examples (*Appendix 1.1: Examples of intended purposes/uses*)*

Appendix 1.1: Examples of intended purposes/uses.

The following examples only refer to what is detected or measured (part i), the function (part ii) and the specific clinical evidence (part iii) (annex I 20.4.1c). These examples do not provide the full description of the intended purpose. ICD codes[2] may be helpful in expressing the specific medical purpose of the device.

The examples here represent different products in principles. For each example of intended use, concepts of clinical evidence have been suggested (scientific validity, analytical performance, clinical performance).

Example 1: IVD intended to detect and measure magnesium

Products different in principles	**Intended Purpose / Intended Use (function and specific medical purpose only)**	**Scientific Validity**	**Analytical Performance**	**Clinical Performance**
A) Physiological state	To detect and measure magnesium to aid in the diagnosis and / or monitoring disorders of magnesium metabolism[3]	Mg2+ 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,	Agreement with other measures of magnesium (method comparison), standardised against atomic absorption spectrometry. Reference ranges appropriate to the clinical condition and target population could be included either from literature or a new study.

[2] ICD codes - International Statistical Classification of Diseases and Related Health Problems https://www.who.int/standards/classifications/classification-of-diseases
[3] ICD-10 Code E83.40: Disorders of magnesium metabolism, unspecified

B) Clinical condition	To detect and measure magnesium to aid in the diagnosis of clinical conditions (e.g. kidney disorders, primary infantile hypomagnesemia, etc.) associated with abnormal magnesium levels in the body, hyper / hypomagnesemia.	• 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.	specificity, precision, etc.	Diagnostic/clinical sensitivity and specificity to detect specific clinical conditions
C) Clinical condition **'therapy monitoring'**	To detect and measure magnesium to monitor therapeutic levels of drugs (e.g. proton pump inhibitors, diuretics, cytotoxic drugs), or 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 intended to detect and measure C-reactive protein (CRP)

Products different in principles	**Intended Purpose / Intended Use(function and specific medical purpose only)**	**Scientific Validity**	**Analytical Performance**	**Clinical Performance**
A) Physiological state	To detect and measure C-reactive protein to aid in the diagnosis and /or monitoring 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 measures of C-reactive protein (method comparison), using standardised reference material. Reference ranges appropriate to the clinical condition and target population could be included either from literature or a new study.
B) Clinical condition	To detect and measure C-reactive protein to aid in the diagnosis and/or monitoring sepsis.	Determination of CRP is clinically useful to screen for organic disease, to assess activity of inflammatory diseases such as sepsis, rheumatoid arthritis, to detect intercurrent infection in systemic lupus erythematosus, in leukaemia or after surgery.		Diagnostic/clinical sensitivity and specificity to aid in the diagnosis of sepsis.

C) Clinical condition **'therapy monitoring'**	To detect and measure C-reactive protein to monitor efficacy of drugs which are known to suppress or prevent inflammatory processes (e.g. ISDs, anti-inflammatory drugs) known to alter C-reactive protein levels.	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 intended to measure Troponin T

Products different in principles	**Intended Purpose / Intended Use** **(Function and specific information)**	**Scientific Validity**	**Analytical Performance**	**Clinical Performance**
A) Clinical condition	To determine cardiac troponin T levels in human serum and plasma to aid in the diagnosis of clinical conditions (e.g. to rule out acute myocardial infarction) and risk associated with cardiomyocyte damage.	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. It is useful in risk stratification in patients with ACS or cardiac risk in patients with renal disease. Determination of TnT aids in early diagnosis (PoC). Measurement of troponin T in serum and plasma aids in therapy selection in patients with elevated Troponin T levels.	Quantitative determination of the troponin T concentration in human serum, and plasma with appropriate analytical sensitivity,	Diagnostic/clinical sensitivity and specificity to detect specific clinical condition, (e.g. to rule out acute myocardial infarction) and hazard ratio to assess associated risk.

B) 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.	specificity, precision, etc.	Appropriate diagnostic/ clinical sensitivity and specificity to monitor troponin T levels in order to adjust or induce appropriate treatment.

Example 4: IVD intended to measure glucose

Products different in principles	**Intended Purpose / Intended Use** **(Function and specific medical purpose)**	**Scientific Validity**	**Analytical Performance**	**Clinical Performance**
A) Physiological state	To determine glucose levels in human serum, plasma and urine to aid in the diagnosis of clinical causes of hypoglycaemia.	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. Measuring glucose levels 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.	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). Reference ranges appropriate to the clinical condition and target population could be included either from literature or a new study.

B) Clinical condition	To determine glucose levels in humans for the diagnosis of diabetes mellitus as part of an oral glucose tolerance test.	Determination of glucose in serum, plasma and urine is useful in diagnosis of diabetes.		Diagnostic/clinical sensitivity and specificity to diagnose diabetes as part of an oral glucose tolerance test.
C) 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

Products different in principles	**Intended Purpose / Intended Use** **(Function and specific information)**	**Scientific Validity**	**Analytical Performance**	**Clinical Performance**
A) Pathological state	To detect specific mutations in the KRAS gene in the DNA of cancer cells and tissue of patients diagnosed with metastatic colorectal cancer to characterise disease prognosis (this intended purpose is not to be confused with the indication associated with companion diagnostics).	Somatic mutation in the KRAS gene is an essential step in the development of colorectal cancer. KRAS mutations are prognostic of clinical outcomes and may help in the understanding of the cancer and the selection of general treatment options.	Qualitative detection of somatic mutations in the KRAS gene using extracted DNA from FFPE samples of CRC with appropriate analytical sensitivity, specificity, precision etc.	Clinical performance can be demonstrated through a review of the literature or from a method comparison study using samples from subjects in the Intended Purpose population'. For KRAS codons 12 and 13 WHO reference panel NIBSC 16/250 is available.

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

Products different in principles	**Intended Purpose / Intended Use** **(Function and specific information)**	**Scientific Validity**	**Analytical Performance**	**Clinical Performance**

A) 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 to monitor response to treatment and check for treatment-resistant mutations.	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. Clinical performance can be demonstrated through a review of the literature or from a method comparison study using samples from subjects in the Intended Purpose population'.
B) Companion diagnostic	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	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.

	monitoring of treatment-free remission.			

Example 7: IVD intended to diagnose COVID-19 infections

Products different in principles	**Intended Purpose/ Intended Use** **(Function and specific information)**	**Scientific Validity**	**Analytical Performance**	**Clinical Performance**
A) Clinical condition	Near patient test to detect SARS-CoV-2 antigens to diagnose (rule in) COVID-19 infection.	SARS-CoV-2 antigens are a marker of COVID-19 infection	Detection of SARS-CoV-2 antigen in relevant sample type with appropriate analytical sensitivity, specificity, precision, etc. Reference material can be used to establish performance,	Study of a sufficient number of positive and negative samples from subjects in the Intended Purpose population' from people with a range of viral loads in comparison with a composite reference method or an established laboratory method in current clinical use when used by the intended user. The required clinical performance reflects the

			including standard validation panels, quality control materials and proficiency testing materials.	function of the device (to diagnose (rule in) COVID-19 infection)

References:

1. Directive (EC) 98/79 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. ISO 18113-1:2022 *In vitro* diagnostic medical devices. Information supplied by the manufacturer (labelling). Part 1: Terms, definitions and general requirements. Definition 3.1.37 intended use/intended purpose.

Chapter 2 - Clinical Evidence

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.

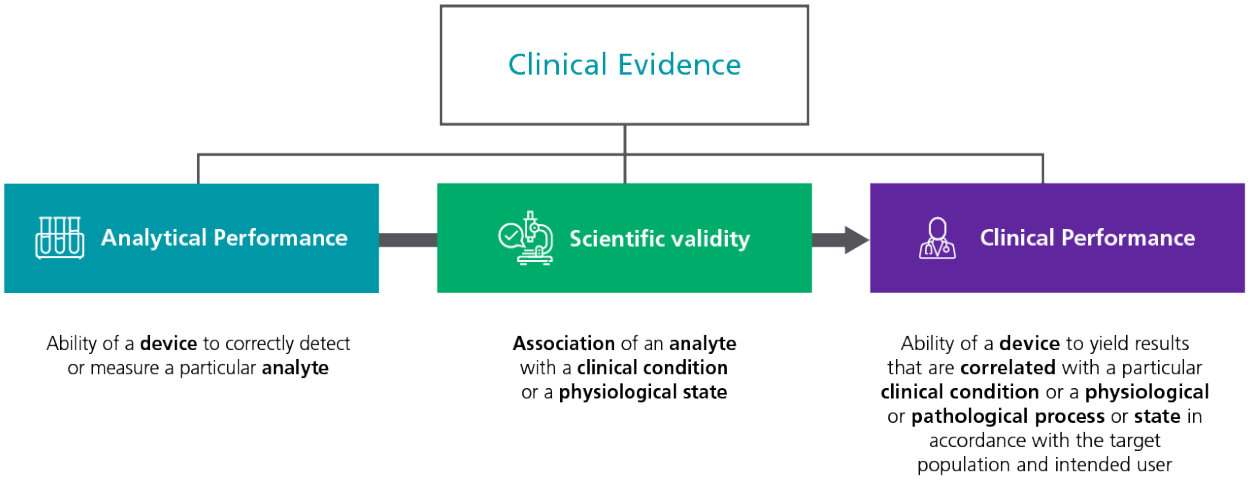


Figure 1. Components of clinical evidence according to IVDR 2017/746

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, and expected values in normal and affected populations.

Annex I, Chapter 2, Section 9.1 and Annex II, Section 6.1.

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 requirements can be omitted and justified as not applicable.

1) What is Scientific Validity?

Scientific validity is a new term and requirement that has been introduced in the IVD Regulation.

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 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 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.

2) 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:

II. 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;

III. appraise all relevant data by evaluating their suitability for establishing the safety and performance of the device;

IV. generate any new or additional data necessary to address outstanding issues.

b. 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.'

3) What is the relationship between scientific validity and clinical utility?

The IVDR does not mention or define **clinical utility**.

- In IMDRF document GHTF/SG5/N6:2012, 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** through 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 1.1 I).

As described below in Chapter 2, in general the demonstration of clinical utility is not a requirement according to the IVDR.

4) 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].

5) 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 3 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 patient samples (sometimes called “clinical samples”), independently 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

- 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.

Clinical Performance

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 3. 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.

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 stardard), **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

Table 4. 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.

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 in terms 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 specifications[11], and others. These are established guidelines that could be considered, but it is beyond the scope of this brochure to provide a comprehensive overview.

The study population should be representative of the EU population. For further details, refer to Chapter 13 of this eBook: *Use of Clinical Data from Outside the European Union*.

6) 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.

7) 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, contracting an investigator. Depending on the type of study, ethics approval may be needed. For clinical performance studies, in addition to the Clinical Performance Study Plan requirements from Annex XIII 2.3.2), also see EN 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 maximise 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.'

8) What are the clinical performance expectations for Therapeutic Drug Monitoring devices?

Therapeutic Drug Monitoring (TDM) devices are *in vitro* diagnostic (IVD) medical devices designed to measure the concentration of a drug and/or its metabolites in biological fluids, such as blood or urine. These devices do not establish medical decision points themselves; Their clinical benefit lies primarily in providing accurate, quantitative data that clinicians use to assess whether drug levels fall within a therapeutic window defined by the drug manufacturer, depending on the clinical context.

Under IVDR Article 2(41), clinical performance refers to a device's ability to yield results correlated with a particular clinical condition or physiological or pathological process or state, taking into account the target population and intended user. This definition presents a challenge for TDM devices, which are not directly correlated with a clinical condition but instead measure drug concentrations which are quantitative parameters influenced by numerous patient-specific and treatment-specific variables.

Drug concentrations can exhibit substantial intra- and inter-patient variability based on factors such as time since administration, organ function, drug metabolism, concomitant medications, and toxicity profiles. Importantly, the therapeutic range, as well as toxic or subtherapeutic thresholds, are not established by IVD manufacturers but by the drug manufacturers themselves, based on clinical evidence gathered during drug development. As a consequence, diagnostic sensitivity or specificity data cannot be generated for TDM devices. Alternatively, reference ranges can represent meaningful performance data. Depending on the interpretation, such data may represent analytical or clinical performance.

References:

1. ISO 18113:2022, Parts 1 to 5 In vitro diagnostic medical devices — Information supplied by the manufacturer (labelling)

2. TGA guidance “Clinical evidence guidelines supplement In vitro diagnostic (IVD) medical devices Version 1.0 March 2020”

3. ISO 20916:2019 In vitro diagnostic medical devices — Clinical performance studies using specimens from human subjects — Good study practice

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

5. 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

6. 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

7. CLSI document EP06-Ed2: Evaluation of Linearity of Quantitative Measurement Procedures; 2nd Ed. CLSI EP06. Clinical Laboratory and Standards Institute, 2020

8. 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

9. CLSI document EP17-A2: Protocols for Determination of Limits of Detection and Limits of Quantification; Approved Guideline. Wayne, PA : s.n., 2012

10. CLSI document EP21-A Estimation of Total Analytical Error for Clinical Laboratory Methods. CLSI. 20, s.l. : CLSI, 2012, Vol. 23

11. CLSI document EP24-A2: Assessment of the diagnostic accuracy of laboratory tests using receiver operating curves; approved guideline - second edition, 2011

12. 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*

13. 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 – State of the art (in medicine)

1) Did the concept of state of the art change from the Directive to the Regulation?

The concept of state of the art has been a core element of the essential requirements of IVD Directive (EC) 98/79 (IVDD)[1] and remains such of the general safety and performance requirements of IVD European Regulation 2017/746 (IVDR)[2]

The IVDR (Annex I, Section 9) stipulates that "Devices shall be designed and manufactured in such a way that they are suitable for the purposes [..], as specified by the manufacturer, and suitable with regard to the performance they are intended to achieve, taking into account of the generally acknowledged state of the art".

Hence, manufacturers must adopt solutions to design a safe and effective device, where benefits to the patients outweigh any residual risks associated with the use of this device. These solutions shall take into account the generally acknowledged state of the art.

Further to the requirements of the IVDD, the IVDR puts a lot of emphasis on the clinical relevance of the diagnostic device. Thus, in addition to the generally acknowledged state of the art of devices, the performance of a device, particularly the clinical evidence and the clinical benefit, shall take into account state of the art in medicine.

As per Article 56 of the IVDR: "The clinical evidence shall be such as to scientifically demonstrate, by reference to state of the art in medicine, that the intended clinical benefit(s) will be achieved, and that the device is safe."

2) What is 'state of the art'?

There is no definition in the IVD Regulation itself, however EU guidance MDCG 2022-2 provides the following definition: "Developed stage of current technical capability and/or accepted clinical practice in regard to products, processes and patient management, based on the relevant consolidated findings of science, technology and experience. Note: The state-of-the-art embodies what is currently and generally accepted as good practice in technology and medicine. The state-of-the-art does not necessarily imply the most technologically advanced solution. The state-of the-art described here is sometimes referred to as the "generally acknowledged state-of-the-art". The same guidance also mentions that "the performance evaluation of an IVD must consider the benefit-risk ratio in light of the state-of-the-art". The regulation uses the term 'state of the art', whereas the guidance refers to it as 'state-of-the-art'. Despite the difference in formatting, both expressions are understood to carry the same meaning.

IMDRF/GRRP WG/N47 and EN ISO 14971:2019/A11:2021 - "Medical Devices-Application of risk management to medical devices"[3] provide a similar interpretation.

Indeed, the state of the art does not necessarily imply the most technologically advanced solution but is sometimes also referred to as the "generally acknowledged and accepted good practice in technology and medicine", as illustrated in the examples below.
This standard also gives a number of methods that can be leveraged to determine 'state of the art' for a device, which may include:

- Standards used for the same or similar devices;
- Best practices as used in other devices of the same or similar type;
- Results of accepted scientific research;
- Publications from authorities, or additional information for similar other products;
- Comparison of the benefits and risks of the device under development with the benefits and risks of similar devices available on the market.

Moreover, European guidelines can also be added as a valuable source of reference.

'State of the art' can be interpreted in some contexts as the 'cutting edge or leading edge' and refers to the 'highest level of general development'[4] of a device. However, this is a marketing perspective and not a regulatory definition.

Based on the foregoing, the concept 'state of the art' is usually used to describe all knowledge accumulated to date and practice in general terms (including but not limited to clinical practice, conceptual thinking in the scientific / clinical field, consensus guidelines, the latest versions of the inter- / national standards and regulations, etc.) on a subject and products to minimise user and patient risk in balance to its benefits. It shall be noted that the concept of generally acknowledged state of the art implies general acceptance as such, rather than individual or regional interpretation.

A device satisfies the 'state of the art' criteria when it has been designed and manufactured to reflect and incorporate that knowledge and practice. The determination of what is the current state of knowledge may always be a matter on which there are different views. Still, it is based on the robust evidence at that point in time (as opposed for example to hypotheses, speculation, etc.). As 'state of the art' reflects the thinking at a point in time, the state of the art of a specific device may not be static and can change as current knowledge and practice changes.

Since standards and 'Common (Technical) Specifications' are the result of the collaborative work of experts in the field, they are likely at least when they are adopted, to reflect the 'state of the art' on that particular subject.

Similarly, EU Reference Laboratories can provide scientific advice regarding the state of the art in relation to specific devices, or a category or group of devices (see IVDR art 100.2. (d)

3) What is 'state of the art in medicine'?

[4] https://en.wikipedia.org/wiki/State_of_the_art

There is no MDCG guidance or definition in the IVD Regulation of 'state of the art in medicine' itself. In the absence of any official reference, 'state of the art in medicine' can be defined as currently accepted medical or diagnostic practice(s) based on current EU clinical guidelines.

IVDR article 56(3) describes 'state of the art in medicine' in relation to the performance evaluation concept for the demonstration of scientific validity, analytical performance and clinical performance. The data and conclusions, as output from the assessment of those elements, constitutes the clinical evidence. By reference to the 'state of the art in medicine', the clinical evidence demonstrates scientifically that the intended clinical benefit will be achieved.

This is often referred to as the standard of care that is defined as "a diagnostic and treatment process that a clinician should follow for a certain type of patient, illness, or clinical circumstances"[5]. Similarly, the' state of the art in medicine' derives from current knowledge and clinical practice taking account of the available diagnostic and therapeutic options.

References to 'state of the art in medicine' at any point in time can be found, e.g. in:

- Medical textbooks;
- Clinical guidelines;
- Peer-reviewed literature;
- Recommendations from medical and / or laboratory associations

4) Changes to the state of the art – what should be considered?

In light of ongoing technological developments and adoption of innovative medical solutions, the evolution of state of the art is inevitable. In such cases, manufacturers should evaluate the intended purpose, the acceptability of the benefit-risk ratio and the clinical benefit assessments, to verify whether the device can continue to be regarded as state of the art. This is particularly relevant for the first IVDR assessment of a device placed on the market under the IVD Directive a long time ago. Furthermore, TR 24971 advises the manufacturers to consider the availability or non-availability of adequate diagnostic alternatives for the clinical condition in the intended population as well as the associated risks and benefits.

Although state of the art refers to current knowledge and practice, this does not mean that state of the art in medicine must always evolve rapidly. At the same time, for certain technologies the state of the art can change more rapidly. Therefore, manufacturers will monitor the state of the art as part of the device's continuous and proactive post-market surveillance activities to mitigate the risks by reducing levels of uncertainty. Individual IVD devices have occasionally been questioned about still being state of the art in medicine, although they are still part of the clinical routine in Europe and elsewhere. Routine uses of state of the art devices according to the intended use in EU healthcare facilities in line with current clinical practice can help illustrate what state of the art means. Examples of such devices are shown in the table below along with the rationale whereby they are still state of the art and in clinical practice. Any changes to the manufacturer's intended purpose need to be supported by clinical evidence. Similarly, changes to the manufacturer's intended purpose of a product may help ensure that the device remains state of the art. IFU should reflect the updated intended purpose.

Examples	Rationale
Creatine Kinase (CK-MB)	Troponins (T or I isoform) have replaced CK-MB for the diagnosis of acute myocardial infarction (AMI). This could lead to the view that CK-MB is no longer state of the art in medicine. However, CK-MB is clinically still useful and routinely used 1) in hospitals that have no access to troponins and 2) in hospitals applying troponins to assess re-infarction, i.e. a 2nd AMI episode that is challenging to diagnose due to the longer half-life of troponins. A CK-MB assay with a revised intended purpose may reflect state of the art provided this is supported by sufficient clinical evidence.
Conventional troponin (non-high sensitivity)	High sensitivity troponin assays have become the gold standard for the diagnosis of acute myocardial infarction (AMI). In conjunction with other medical information, they allow for early rule in / out of AMI. This could lead to the view that conventional troponins devices are no longer state of the art in medicine. However, conventional troponin is clinically still useful and routinely used in some settings where high-sensitivity troponins are not available, e.g. Point of Care settings, and particularly for ruling in AMI. A conventional troponin assay labelled to reflect this new intended purpose might be considered to be state of the art.
Antimicrobial Sensibility Testing (AST)	Agar dilution or broth microdilution are well established methods for the purpose of determination of the minimum inhibitory concentration. The breakpoints for such change regularly according to CLSI and EUCAST guidelines. Manufacturers are required to be vigilant and assess how the new breakpoints influence the test results. If the interpretation of the test result is irrespective of the new breakpoint, the device continues to be state of the art.

Table 5. Examples of devices that represent state of the art in medicine

References:

1. Directive (EC) 98/79 of the European parliament and of the council of October 27, 1998 on in vitro diagnostic medical devices

2. Regulation 2017/746/ EU of the European parliament and of the council of April 5, 2017 on in vitro diagnostic medical devices

3. EN ISO 14971:2019/A11:2021 Medical devices – Application of risk management to medical devices

4. ISO TR 24971:2020 Medical devices – Guidance on the application of ISO 14971

5. IMDRF/GRRP WG/N47 FINAL:2018 – Essential Principles of Safety and Performance of Medical Devices and IVD Medical Devices

6. Stöppler, Melissa Conrad. *Medical Definition of Standard of care*, MedicineNet, 29.03.2021, https://www.medicinenet.com/script/main/art.asp?articlekey=33263

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 it is 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:

1) scientific validity (as defined in Art. 2 (39));
2) analytical performance (as defined in Art. 2 (40));
3) 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 clinical evidence levels. 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 medicine[2], 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?

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). In the case of TDMs, reference ranges can represent meaningful performance data. Depending on the interpretation, such data may represent analytical or clinical performance (see also footnote 5 below). 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. MTE proposes that clinical performance levels are proportionate to risk classification and intended purpose. 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 typically follow the following pattern: class B < class C < class D devices (see Figure 2 below). However, manufacturers need to assess it case by case. Consideration should be given to the novelty of the device, in addition to its intended purpose. For a more comprehensive understanding of evidence levels and the process of building clinical performance, please refer to Figure 3 and Figure 4.

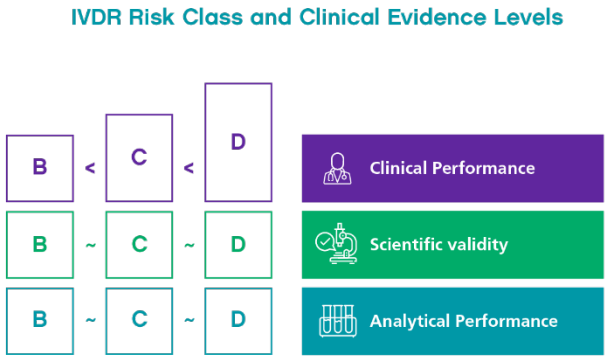


Figure 2. Risk-based evidence levels for analytical performance, scientific validity, and clinical performance

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 (including Real World Evidence)

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 in the technical documentation 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

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[5]	Not applicable, unless reference ranges (if applicable) can be used as clinical performance. Omission to be justified in the respective Clinical Performance section of Performance Evaluation Plan and Report

Table 6. 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.

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 on published experience gained by routine

diagnostic testing and the chapter on Real-World Evidence). 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 with the objective of establishing equivalence with a selected reference device that has a published and accepted body of strong clinical evidence.

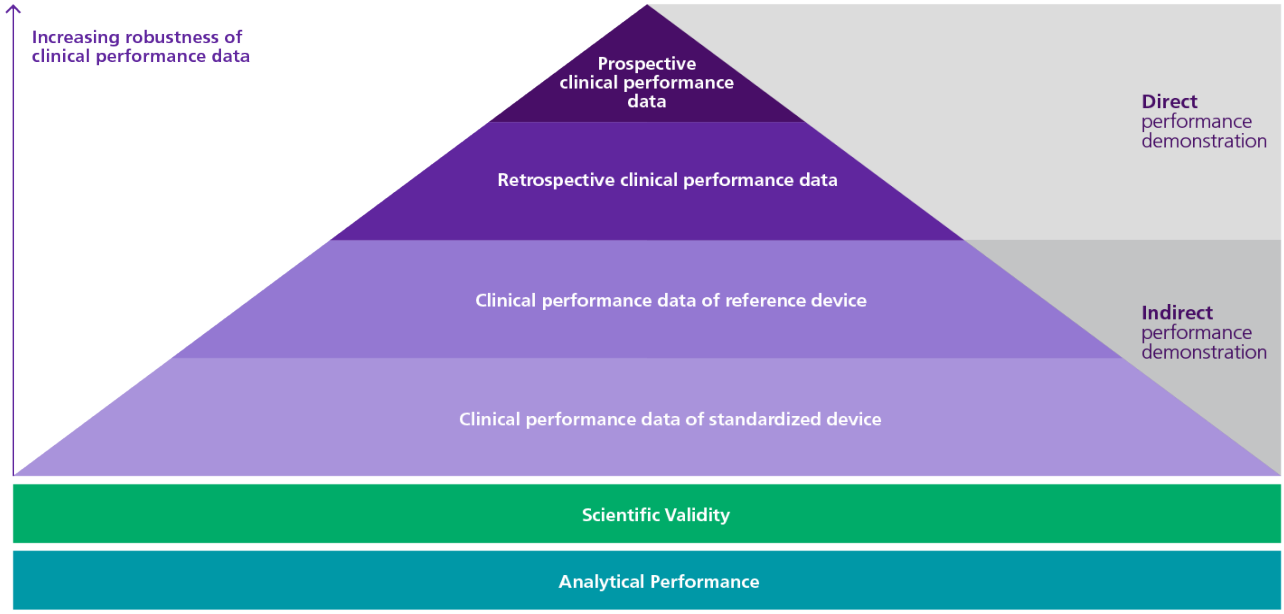


Figure 3. Clinical evidence levels for IVD classes B, C, and D

It should be noted that multiple general evidence grading systems exist (e.g. GRADE[5]), QUADAS-2[6], Hayes[7]) 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 Harmonisation Task Force (GHTF)3
  a) established, standardised device
  b) established, non-standardised device
  c) novel device
III) IVDR class

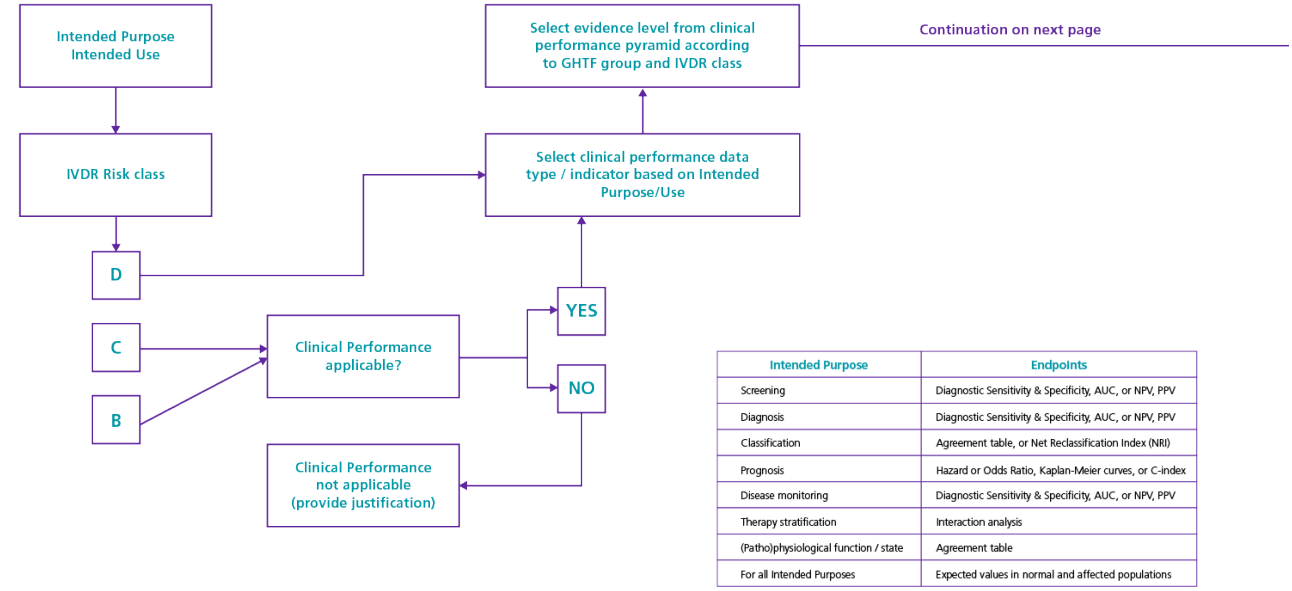
Determining clinical performance indicators and study endpoints

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), 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, for instance stability and integrity of 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 and equivalence established via published data from reference devices, provided the analytical performance determination is performed using standardised device and reference material. However, it is essential that some data using material with demonstrable clinical provenance ("clinical samples") is generated for the device under application to support the claim.



Intended Purpose	Endpoints
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	Interaction analysis
(Patho)physiological function / state	Agreement table
For all Intended Purposes	Expected values in normal and affected populations

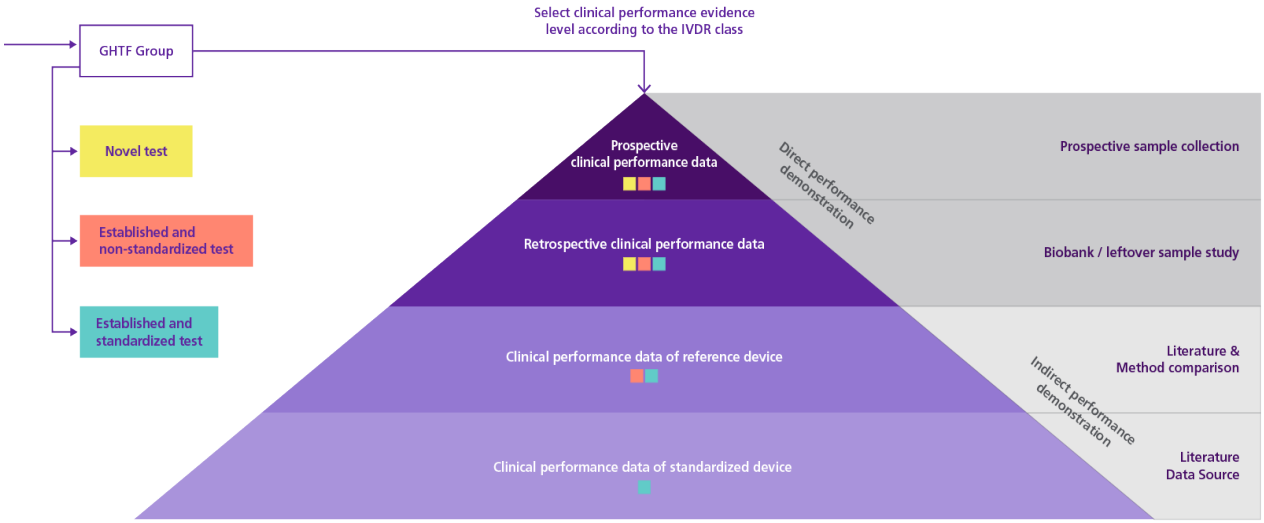


Figure 4. Flowchart for Clinical Performance

**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*

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 6 - How to demonstrate evidence gained from 'published/documented routine testing' and Chapter 9 – 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 the 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/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/documented 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 (EMA) 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[6] 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[7]
- 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)

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

[6] Modified from *Dictionary.com*
[7] Might be free of charge (e.g. website from clinical labs)

3) What does published experience refer to?

Any document or set(s) of data coming from the use of the device and which 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 kinds of data are included in published experience gained by routine diagnostic testing?

Routine diagnostic testing may involve various data sources, as outlined below. The data must come from a specific device that is identical, similar, or equivalent to the device in question. This device can be CE marked under IVDD or IVDR. The level of equivalence must be determined before the data can be used. If the data is from a similar or equivalent device, equivalence must be established first to ensure its relevance.

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

6) Searching for published literature

Published data can be collected according to scientific principles using predefined search terms with a qualified assessment of the search results. Relevant data should have been generated using the device according to its intended purpose (e.g. instrumentation, target population, use environment etc.).

Chapter 6 – Equivalence and similarity concepts in the IVDR

1) What are the concepts of equivalence and similarity as used in IVDR?

With respect to performance evaluation, equivalence and similarity are connected terms. Clinical evidence for a device can be based partly or totally on clinical evidence from an equivalent or similar device. The suitability, relevance and adequacy of the claim for equivalence is assessed by the Notified Body. [Annex IX part 4.5]

2) 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:	
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: • Validity of **equivalence** claimed in relation to other devices, the demonstration of **equivalence**, the suitability and conclusions data from **equivalent** and **similar** devices
Annex IX: Conformity Assessment based on a Quality Management System and on assessment of Technical Documentation	Chapter 1: Quality Management System • 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 • 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 2: 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 • 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: • 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 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 7. Compilation of references of terms 'equivalence', 'equivalent', similar' throughout the IVDR related to performance evaluation

3) 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 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. A justification should be included for why a particular comparator device is sufficiently similar.

- '**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 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). Yet, it remains to be seen whether biological, technical and clinical characteristics will become part of the definition of 'equivalence' for IVDR.

- 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 and the intended purpose of the product.

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

Table 8 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.**

The concept of equivalence and similarity apply to performance evaluation and PMS/PMPF but it may be more challenging to conclude equivalence in a PMS/PMPF setting because the information on the comparator device(s) on the market is limited.

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

Measures of safety:					
Test limitations					
Risks					
Summary of Safety and Performance					
Other measures of safety?					
Measures of performance:					
Analytical performance characteristics (Annex I, Chapter 2, 9.1 and Annex II, Section 6.1)					
Clinical performance Annex I, Chapter 2, 9.1 (b)					
Scientific validity					
Intended purpose/use					
(i) what is detected and/or measured;					
(ii) its function (e.g. screening, monitoring, diagnosis or aid to diagnosis, prognosis, prediction, companion diagnostic);					
(iii) the specific disorder, condition or risk factor of interest that it is intended to detect, define or differentiate;					

(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) the intended user;					
(ix) in addition, for companion diagnostics, the relevant target population and the associated medicinal product(s).					
Design Information:					
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))					
Biological controls (metrological traceability)					

Antibodies (polyclonal/monoclonal)					
Clinical benefits to patients.					

Table 8. 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.

5) How to use this table?

The terms 'equivalence' and 'clinically significant difference' should be pre-specified by the manufacturer. The table lists possible technical, analytical, biological and clinical characteristics of an IVD device in terms of safety and performance. 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 clinical evaluation of the significance of the difference.

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 (CDx) described in the IVDR?

Recitals 10 to 12 and Article 2 (f) of the IVDR introduce a 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(s) 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;

It is understood that the device manufacturer is responsible for determining whether the device is considered '*essential*'. This decision may be subject to assessment by the Notified Body and can be supported by various sources, including the drug's labelling or Summary of Product Characteristics (SmPC), clinical trial reports of the medicinal product, the IVD's scientific validity report, clinical performance study report, or an opinion from a medicines authority.

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".
On the other hand, 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.

Examples include:

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

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

in the need to measure cyclosporine blood concentrations to enable the dose of the drug to be individualised to the patient. When done correctly, therapeutic efficacy can be maximised while toxicity is kept to a minimum[1].

8) Such a device intended to monitor levels of medicinal products, substances or biological components, is classified under IVDR Annex VIII, rule 3 (j). For more information, please go to chapter 1 "'Intended Purpose / Use'" for the question 8. Clinical performance expectations for Therapeutic Drug Monitoring devices?.

– **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 the case 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 factors[2].

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

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

It should be noted that in cases where it is not clear if a device has a CDx indication, the European Medicines Agency (EMA) European public assessment report (EPAR) and the summary of product characteristics (SmPC) should be consulted. These resources should indicate if a medicinal product's indication depends on a specific gene/phenotype or biomarker-based patient selection[3].

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

Article 58 (2) describes the following: 'performance studies involving companion diagnostics shall be subject to the same requirements as the performance studies listed in Article 58 paragraph (1)'.
CDx performance studies often will be 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 usually must meet:

- 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[8] are used, in a non-interventional setting, the requirements under Article 58 paragraph (1) do not apply to such performance studies. Such studies must, however, be notified to the competent authority.

A study concept with leftover or archived samples may play a role in diagnostic 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.

An overview of the IVDR general and additional requirements in relation to CDx performance studies is shown in Figure 5 below.

---

[8] How are leftover & archived specimens defined[4]?

- 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 defines these terms as follows:

'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 defines archived samples as follows

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).

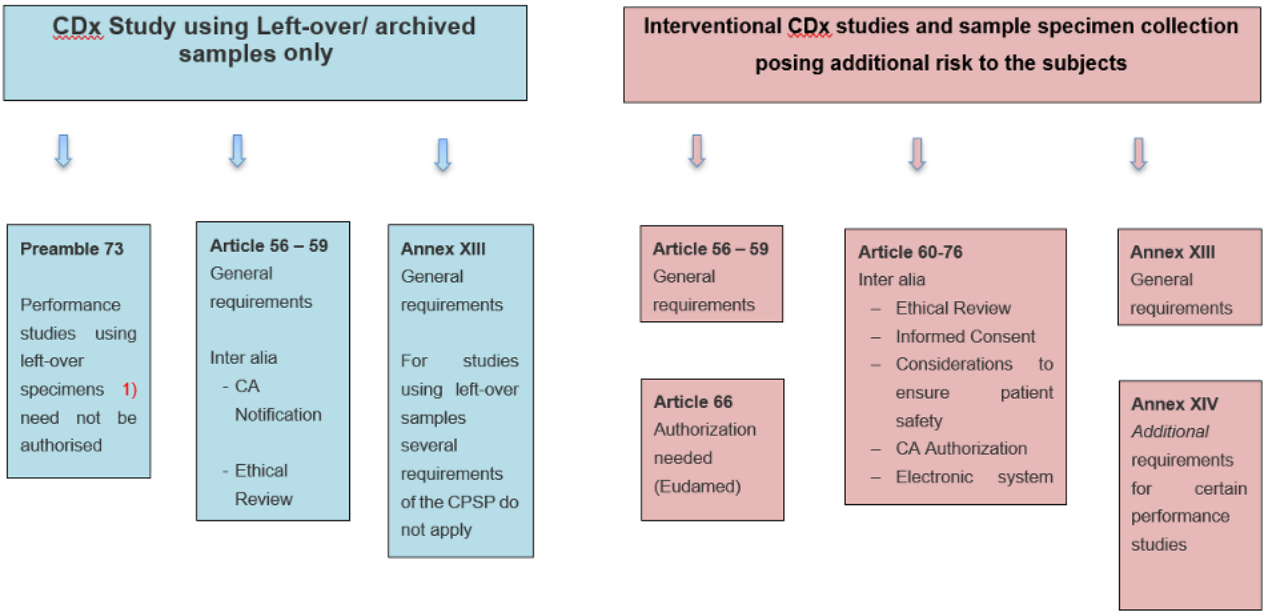


Figure 5. IVDR Requirements for CDx study using leftover/archived samples compared to interventional CDx study and specimen collection with additional risk to the subjects

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). In case of co-development, the application for authorisation of the CDx performance study ideally takes place at the same time as the application for authorisation of the associated clinical drug trial of the medicinal product. For combined studies, the IVDR requirements have to be met.
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 (Part A) and Chapter 1 of Annex XIV.

Based on Article 66, the Notified Body is not involved in 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 European database on medical devices (EUDAMED) system (Article 69).

The Member States notify the sponsor of the authorisation. 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 of the application for an interventional CDx performance study based on Articles 66, 67 and 71 is displayed in Figures 6 and 7 below.

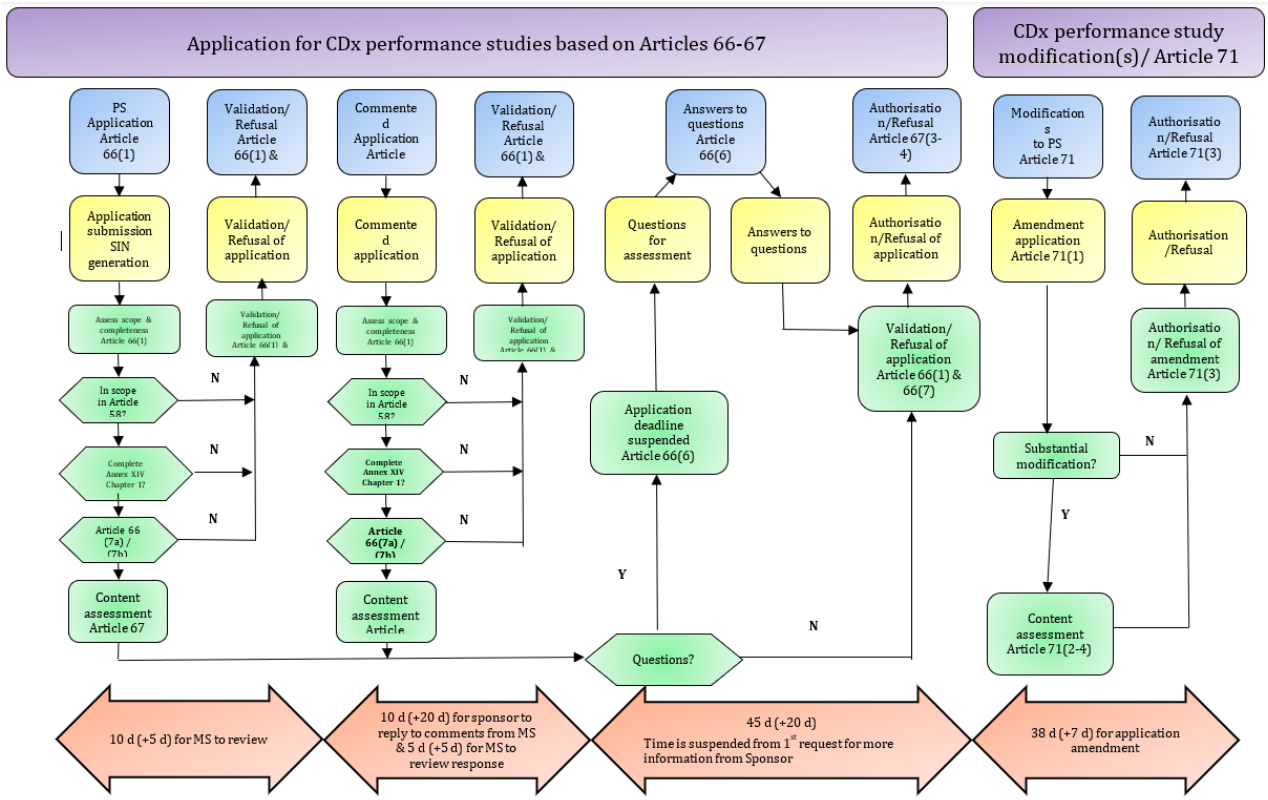


Figure 6. Process flow of the application for an interventional CDx study and related timelines based on Articles 66, 67 and 71.

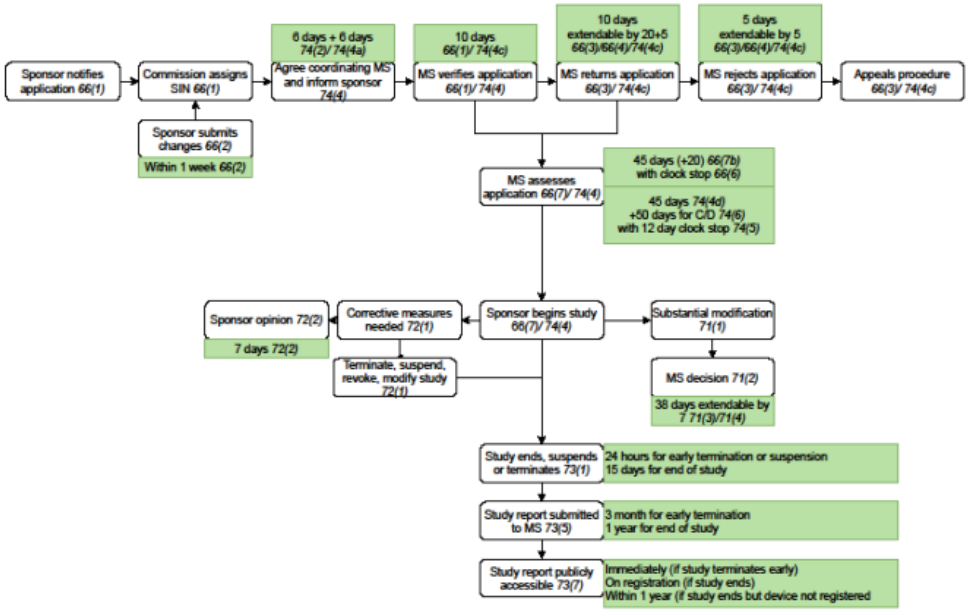


Figure 7. Process flow of the application for an interventional CDx study and related timelines *(provided as courtesy by Steve Lee)*

5) When can a non-interventional 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 authorisation, it is unclear if this notification is planned as a national notification or if it will be done over the EUDAMED electronic system on performance studies (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 CDx devices used performance studies?

Regardless, if they are used in an interventional or non-interventional performance study using leftover samples only, the CDx device 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). By exception, if the device already bears the CE marking and therefore falls under the provisions of Article 70(1) of the IVDR ('PMPF study'), then it may continue to bear that CE marking.

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 eBook. 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. It should be noted that under IVDR contemporaneous approval of a companion medicine is not required. However, for a CDx to be scientifically valid the companion medicinal product needs to be approved in Europe.

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 the clinical performance indicator (see Table 9 below).

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

- 'therapy prediction' (also known as 'therapy response prediction', or 'predictive CDx Intended Use' in other references), applied when the study design includes both marker positive and marker negative patients, or

– ‘therapy selection’ (also known as ‘selective CDx Intended Use’ similar to therapy prediction, but applied when a ‘marker positive only’ study design is used).

No other Intended Purpose/Use than CDx 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 question 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 to generate appropriate clinical evidence for a CDx device to predict or select a specific therapy. This is possible during co-development of IVD CDx and therapeutic or after development of the therapeutic.

A co-developed CDx (as defined by the EMA) is a device that is developed in a clinical development programme together with the concerned medicinal product, either in view of an initial marketing authorisation or a change of the indication. This can mean that the device was developed in the framework of a pivotal clinical trial with the concerned medicinal product or of a bridging study assessing the concordance of the CDx and the device used in the pivotal clinical trial of the corresponding medicinal product[5]. In case of a bridging study, sufficient documentation needs to be provided to conclude that the performance compares to the device used in the pivotal clinical trial of the corresponding medicinal product and that there is no impact on clinical performance that would be incompatible with the safe and effective use of the medicinal products described in the SmPC.

Another example of CDx development after launch of a therapeutic is a follow-on CDx device. A follow-on CDx (as defined by the European Medicines Agency) is a device that seeks the same indication in its intended use as the co-developed CDx (hereafter, original CDx). The follow-on CDx targets the same biomarker but is not developed in parallel with the clinical development programme of the medicinal product and is not necessarily based on the same technology as the original CDx. The analytical and clinical performance of a follow-on CDx, and the consequential safety and effectiveness of the associated medicinal product, should therefore be highly comparable to the original CDx. For follow-on devices, sufficient documentation needs to be provided to conclude that the analytical performance compares to the original CDx and that there is no impact on clinical performance that would be incompatible with the safe and effective use of the medicinal product as described in the SmPC. In addition, the scientific rationale/justification for the comparator (original or predicate) device and how cutoffs have been established should be included in the technical documentation and Summary of Safety and Performance (SSP).

In any case, a corresponding study and analysis needs to show that the proposed CDx device is able to predict or select the patients into likely responders or on-responders (see Table 9), and subsequently also show that the group of patients that was characterised as likely responders were also the ones that benefitted the most from the treatment and/or show favourable safety[6]. 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 randomised 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

- 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.

Clinical Performance

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 9. Possible examples of analytical and clinical performance indicators based on the intended purpose. Therapy prediction 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 stardard), **NPV, PPV**	Diseased patients with or without treatment	Prospective or retrospective observational longitudinal study	Glucose, PSA
Therapy prediction (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

Table 10. 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 the 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 (EU) 2017/746. For the CDx Intended Use of Therapy Prediction or Therapy Selection, a clinical outcome study may be involved in 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 discussed in the earlier chapters, the IVDR mentions cut-offs 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 clinical (or surrogate) outcome data arising from prospective or retrospective trial data involving the therapeutic 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 chapter 1 of this eBook - 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 8 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(s) 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 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 prediction and/or therapy selection.

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

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 trastuzumab 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.

MDCG 2022-10 "*Q&A on the interface between Regulation (EU) 536/2014 on clinical trials for medicinal products for human use (CTR) and Regulation (EU) 2017/746"* gives the following clarifications:

Device for performance study: according to IVDR Article 2 (45) means a device intended by the manufacturer to be used in a performance study. A device intended to be used for research purposes, without any medical objective, shall not be deemed to be a device for performance study.

Stratification: Within clinical trials, a method used in the randomisation to ensure equal distribution of chosen variables between treatment arms.

Clinical Benefit Assessment of a BRAF CDx Device (therapy prediction 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.

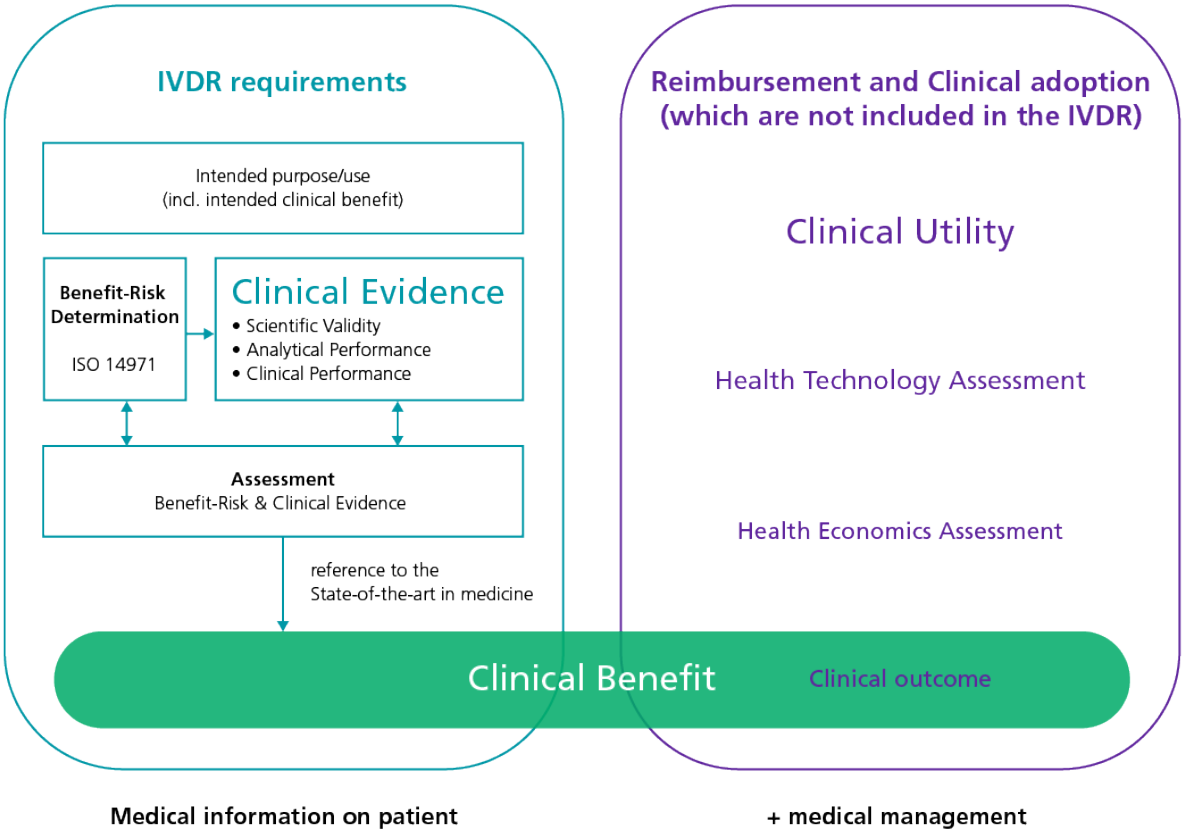


Figure 8. Clinical benefit and clinical utility concepts under the IVDR for CDx devices

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 IVDs, 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 the IVDR. They are required for the therapeutic product.

12) What are 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 on the related therapeutic, the scientific validity of the test, the availability of similar 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 the 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 predict or select the therapeutic product. As mentioned earlier, under the IVDR, the medicinal product needs to be approved in Europe to be scientifically valid as a CDx. For co-developed devices, the medicine may still be under review at the EMA so the technical documentation should document if that is the case.

14) What are the sources for clinical performance data?

Based on the Intended Purpose/Use of therapy prediction, CDx devices always require clinical performance data (omission cannot be justified). Specifically, they require evidence demonstrating that the CDx can successfully predict 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 predict 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 predicted 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 Table 11 below.

IVD CDx Device	Function/Intended Purpose/Intended Use	Clinical Performance
PD-L1	Therapy prediction: aid in the assessment of lung cancer patients for whom anti-PD-L1 therapy is considered	Interaction analysis demonstrating that the CDx can successfully stratify the patients into responders or likely non-responders.
HER2	Therapy selection: aid in the assessment of breast and gastric cancer patients for whom Anti-HER2 therapy is considered.	Observed clinical efficacy in the population(s) selected using the CDx.
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 11. 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.

15) What is a Follow-On CDx?

A follow-on CDx is an IVD that is similar to an original (predicate) CDx for the same corresponding medicinal product (e.g. the original CDx developed during the clinical trial of the corresponding medicinal product). See also chapter 6 on equivalence.
The manufacturer of a follow-on CDx device might not have a therapeutic partner to conduct a new clinical trial or might lack the patient samples from the original clinical trial where the original CDx and therapeutic product were evaluated. As such, an external comparison study is conducted to assess the similarity between the original and the follow-on device. The therapeutic efficacy for the corresponding medicinal product, when used with the CDx in the intended use population, should be similar between the follow-on and earlier comparator companion diagnostic device[7].

16) What is a Follow-On CDx concordance study?

Although the terms ‘concordance’ and ‘bridging’ are not terms found in the IVDR, for the purposes of this guidance, a ‘concordance’ or ’bridging’ study can be used to assess the similarity between the earlier comparator CDx (also referred to as the original or predicate device) and the follow-on CDx device. To

support the same intended purpose, the safety and effectiveness of the comparator and follow-on CDx should be similar and meet predefined acceptance criteria.

Relying on a simple method comparison study between the approved CDx and its follow-on CDx to assess comparability between these two devices is generally not acceptable for approval, because it is unknown how different levels of analytical comparability between the two CDx would translate into clinical performance of the follow-on CDx. Therefore, the regulatory review of the follow-on CDx generally may also be expected to include some type of assessment of clinical performance to ensure that the use of the follow-on CDx would not alter the established therapeutic efficacy and safety profile (derived from FDA published literature)[8]. As stated above (Q15), an external comparison study using a dedicated design and methodology may be considered to assess the 'concordance' between the original and the follow-on device[9]. Figure 9 below summarises the level of clinical evidence according to the availability of direct clinical outcome measures compared to external method comparison/concordance studies.

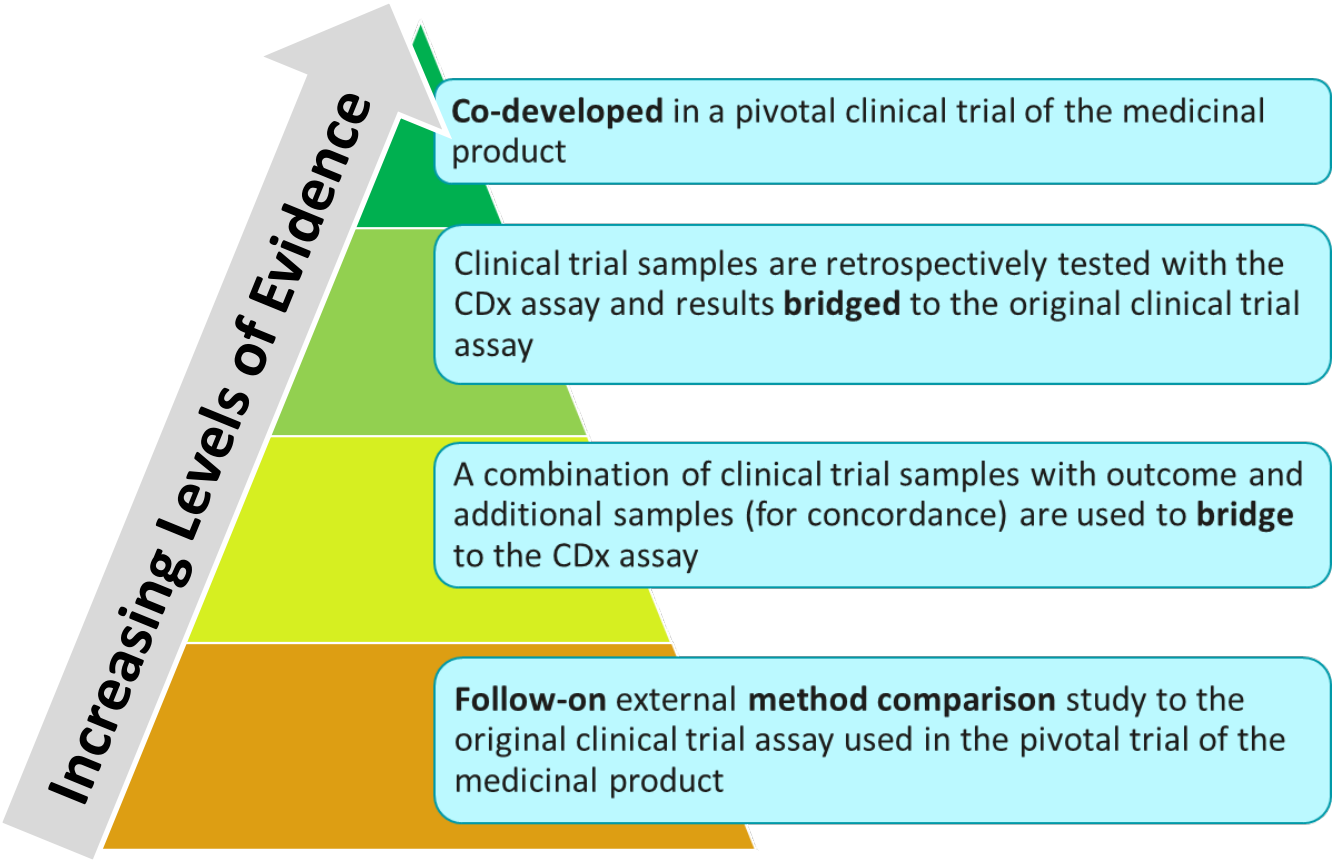


Figure 9. Different ways to leverage the evidence for clinical performance of a CDx

When clinical trial samples directly linked to clinical efficacy are not available (e.g. due to limited specimens, or ethical considerations for conducting a new clinical trial, etc.) a follow-on approach may be the only way to demonstrate clinical performance for the proposed CDx device.

References:

1. Jorga A Holt DW, Johnston A. Therapeutic drug monitoring of cyclosporine

2. 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

3. Garcia LPO et al, Front. Med., 31 October 2021Sec. Regulatory Science Volume 8 - 2021 | https://doi.org/10.3389/fmed.2021.753187

4. Status of Companion and Complementary Diagnostics: Strategic Considerations for Development and Launch https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5355969/

5. Meijuan Li, Statistical Consideration and Challenges in Bridging Study of Personalised Medicine, Journal of Biopharmaceutical Statistics, 2015 25:3, 397-407

6. JT Jørgensen and M. Hersom (Ann Transl Med.) 2016 Dec; 4(24): 482

7. 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

8. Kalavar S., Philip R., IVDs and FDA Marketing Authorisations: A General Overview of FDA Approval Process of an IVD Companion Diagnostic Device in Oncology. In: Badve S., Kumar G. (eds) Predictive Biomarkers in Oncology. Springer, Cham, 2019

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

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 10), the required frequency for updating the reports, and seeks to clarify elements of the wording.

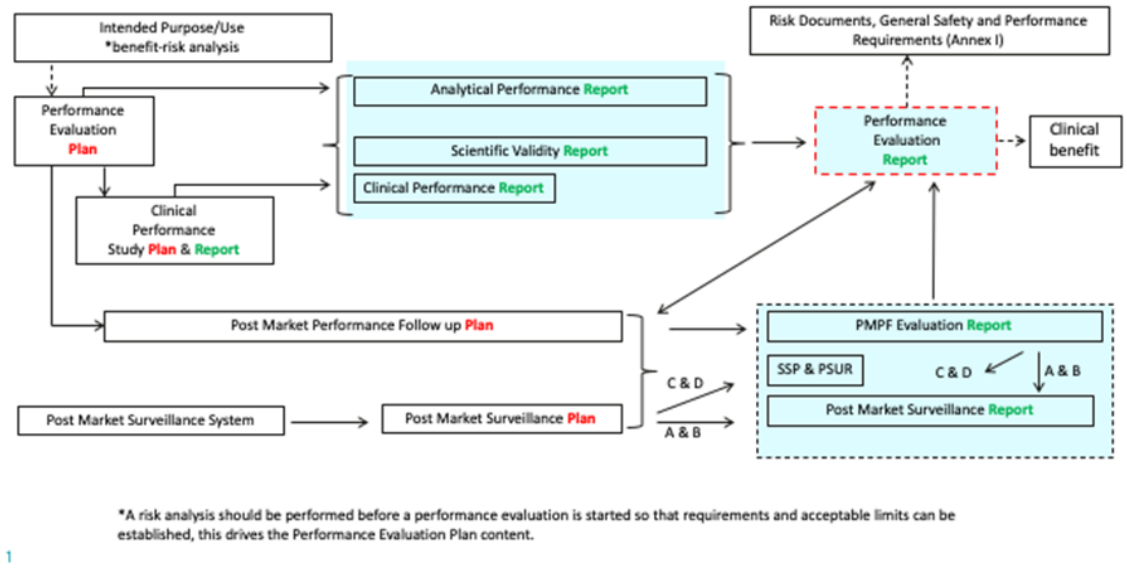


Figure 10. Flow of Plans and Reports for Performance Evaluation.

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.

Analytical performance documentation requirements are: Annex I (9.1 a), ISO13484 and Annex II 6.1.2. Annex XIII does not refer to the analytical performance plan requirement, however Section 3 refers to studies other than clinical performance studies which shall be documented in the same way. Analytical performance study documentation may be included in the performance evaluation report and may be therefore addressed in a similar manner as the clinical performance study plan and report. This open approach leaves it up to the manufacturer to use this concept for other studies, such as feasibility studies.

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. The conclusions of the performance evaluation report may lead to changes to the intended purpose or performance evaluation plan.

Table 12 below provides an overview of the 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 12. 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)?

A benefit-risk analysis is to aid the determination of what specific performance evaluation studies should be carried out and what data may be required prior to any interventional study.
As per the similar requirement in ISO 20916 section B7 "Risks and benefits of the IVD medical device under investigation and clinical performance study:

a) Anticipated adverse device effects
b) Anticipated adverse events associated with the study other than those associated with the IVD medical device, e.g. during specimen collection
c) Residual risks associated with the study, as identified in the risk analysis report

d) Steps that will be taken to control or mitigate the risks
e) Risk-to-benefit rationale."

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 20916[2] 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 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 EN ISO 14971[3].

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 General Data Protection Regulation (GDPR)[3] 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 the 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 20916:2020 In vitro diagnostic medical devices – Clinical performance studies using specimens from human subjects – Good study practice

3. EN ISO 14971:2019 Medical Devices – Application of risk management to medical devices

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

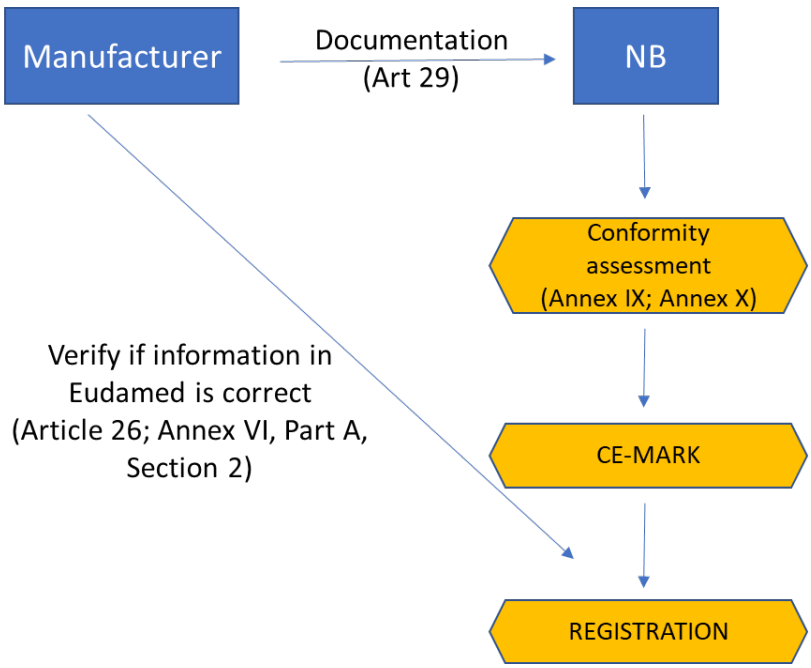
The Summary of Safety and Performance (SSP) is one of the requirements of the new Regulation, specific for class C and D devices, to enhance transparency and adequate access to information. It intends to provide public access to summarised data on the safety and performance of class C and class D IVD 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 does not - by no means - replace 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.

1) Where to find the SSP templates?

MDCG templates for SSP are available, please see guidance MDCG 2022-9.

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).

Notified Bodies will only validate and upload the SSP for devices they are required to assess during the conformity assessment process. In cases where the assessment is based on sampling—meaning only representative devices are reviewed rather than every individual file—the Notified Body will not upload the master SSP for devices not individually evaluated. Similarly, translated versions of the SSPs for these non-

assessed devices will not be uploaded by the Notified Body. These responsibilities, including uploading both master and translated SSPs, will eventually be transferred to manufacturers via the EUDAMED system.

See in MDCG 2022-12 *"The SSP shall be made available to the public upon request without undue delay, or the manufacturer shall specify where it is made available to the public.*
Note: The functionality is available in Eudamed. The system may be used (on voluntary basis) for the upload of the SSP even before the notice of full functionality of Eudamed has been published"

3) What is the frequency of updates?

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. For further details of what should be included in the SSP for a CDx, please refer to Chapter 7.

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. PMPF and PMS may help the manufacturer to update a product according to the state of the art by closely monitoring the market and following the scientific and clinical progress. Figure 11 describes how PMPF relates to other elements of the IVDR.

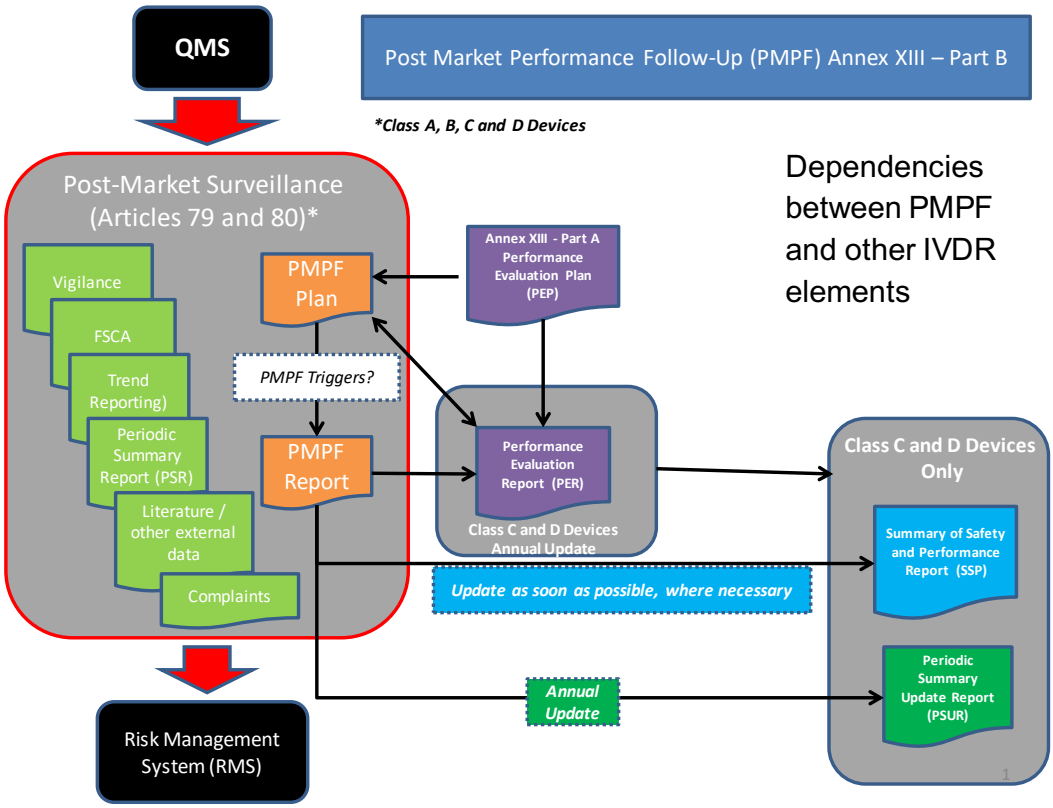


Figure 11. Dependencies between PMPF and other IVDR elements

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 13 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 of the relevant Common Specifications and harmonised standards consulted and relevant PMPF guidance should also be listed, as well as a reference to the relevant sections 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.

The overall objectives of the PMPF are to:

- confirm the safety, performance and scientific validity of the device throughout the expected lifetime;
- identify systematic misuse[9];
- 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.

Notified bodies may require manufacturers to undertake specific PMPF studies (IVDR Article 51 (3)).

[9] 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 guidance[3].

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 new risks. Identify new limitations and contra-indications.	Product class-dependent. (TBD by the manufacturer).
Gathering of clinical experience gained	Conduct post-market studies on data generation.	Post-market studies will allow further collection of safety and	Verify that product claims are met	Product class-dependent. (TBD by the manufacturer).

	Conduct company-sponsored or investigator-initiated post-market studies. Evaluate patient registers where applicable.	performance data, including large-scale data where 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.	Identify safety issues. Analyse the benefit/risk ratio. Identify new risks. Identify new limitations and contra-indications.	
	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).

Table 13. 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.
^ 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 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 annually[4] to include important developments and the PMPF key 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?

In addition to specific Notified Body requests for PMPF (see art 51 (3)), 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 or 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 evaluation[5]. 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 transparency[6].

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 11 and Table 14 in this eBook. The chapter on Documentation further describes the flow of plans and reports.

	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
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

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

Submitted to EUDAMED

Table 14. PMPF and PMS requirements

5) In what instances is PMPF not deemed a requirement?

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:
  - Justification: Performance is typically related to reagents running on the instruments; other PMS activities (see Table 14) should be sufficient to monitor performance
- Class A - Washing solution – separate, not included in IVD test/kit:
  - 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 Standardised tests on the market:
  - Justification: Sufficient data from other devices available to mitigate the risk so that other PMS activities should be sufficient to monitor performance

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 analyse 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:

- Commission Implementing Regulation (EU) 2022/1107 of 4 July 2022 laying down common specifications in accordance with Regulation (EU) 2017/746
- Standards:

PMPF Time Schedule

The data will be reviewed each year and gathered in a report according to the table below (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	Evaluate 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	Evaluate the specificity and results	If complaints linked to specificity performance
Clinical experience gained	Conducting a post-market clinical study according to Annex XIII IVDR /ISO 20/916	To collect new performance information on the product	Evaluate 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	Evaluate 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	Dependent on occurrence of the event

Table 15. PMPF plan example 1

^ 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 analyse 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)

Examples - General Methods and Procedures	Specific methods and Procedures	Rationale for method and procedure appropriateness	Objectives	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 Analyse the benefit/risk ratio Identify new risks Identify new limitations	Customer complaint data will be monitored continuously through PMS activities

Table 16. PMPF plan example 2

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

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

7. Q&A on the interface between Regulation (EU) 536/2014 on clinical trials for medicinal products for human use (CTR) and Regulation (EU) 2017/746 on in vitro diagnostic medical devices (IVDR)

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 20916:2020 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

Chapter 11 – Benefit-Risk Requirements & Potential Approaches under the IVDR

Scope

This chapter is intended to assist in understanding the requirements of the IVD Regulation[1] with respect to capturing 'clinical benefit' when carrying out benefit-risk assessments. Approaches to capture the benefit-risk assessment are also considered. The IVD Regulation takes precedence with respect to benefit-risk. In this chapter, attention is also drawn to other recognised guidance documents with particular reference to EN ISO 14971:2019[2], the risk management standard for medical devices.

Key definitions from the IVDR

Benefit-Risk (IVDR: Article 2 (17))[1] - 'benefit-risk determination' means the analysis of all assessments of benefit and risk of possible relevance for the use of the device for the intended purpose, when used in accordance with the intended purpose given by the manufacturer.

Clinical Benefit (IVDR; article 2(37))[1] - 'clinical benefit' means the positive impact of a device related to its function, such as that of screening, monitoring, diagnosis or aid to diagnosis of patients, or a positive impact on patient management or public health.

- IVDR, Recital 64[1] - It should be recognised that the concept of clinical benefit for *in vitro* diagnostic medical devices is fundamentally different from that which applies in the case of pharmaceuticals or of therapeutic medical devices, since the benefit of *in vitro* diagnostic medical devices lies in providing accurate medical information on patients, where appropriate, assessed against medical information obtained through the use of other diagnostic options and technologies, whereas the final clinical outcome for the patient is dependent on further diagnostic and/or therapeutic options which could be available.

Clinical evidence (IVDR: Article 2 (36))[1] 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.

Additional definitions from EN ISO 14971:2019

Benefit – positive impact or desirable outcome of the use of a medical device on the health of an individual, or a positive impact on patient management or public health.

Note: Benefits can include positive impact on clinical outcome, the patient's quality of life, outcomes related to diagnosis, positive impact from diagnostic devices on clinical outcomes, or public health impact.
Risk – a combination of the probability of occurrence of harm and the severity of that harm

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

Unless specific patient management steps are included in the manufacturer's intended purpose, then clinical benefit refers only to the function of the device not including the potential benefits that may arise as a result of patient management (i.e. clinical utility).

The clinical benefit of an IVD may be unrelated to the final clinical outcome for the patient, and so focuses on the 'accurate medical information' output of an IVD device, in the context of the intended purpose and claimed performance of the device as defined by the manufacturer and in conjunction with other medical information.

However, in some instances, the clinical benefit of an IVD may be related to the positive impact of the IVD result on patient management or public health. **Where a specific patient management decision is part of the manufacturer's intended purpose, then the impact of patient management is an essential element of clinical benefit. For example, companion diagnostic IVDs are linked to a specific therapeutic outcome.**

2) 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 performance of the device) 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[10] 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.

3) How is clinical utility related to clinical benefit?

In general, a manufacturer is not required to demonstrate elements of clinical utility in pre- or post-market phases. Patient outcomes, cost and cost effectiveness are outside the scope of the IVDR.

As an exception, where the intended purpose of the IVD is linked to a specific patient management decision (for example a companion diagnostic IVD), clinical utility can be demonstrated in patient outcome studies. For a companion diagnostic IVD claiming therapy stratification and concomitantly improved patient outcomes

[10] There are currently no generally accepted quantitative, structured methods for assessing benefit.

in the Intended Purpose, evidence should usually be generated from pharmaceutical trials investigating therapeutic regimens together with the companion diagnostic.

In the case of blood glucose testing, the clinical utility could be described should the manufacture wish to do so, for example 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. However, it is not an IVDR requirement to demonstrate the clinical outcome and/or health economic benefits of a glucose testing device as long as the Intended Purpose is limited to diagnosing or monitoring diabetes.

4) What are the general requirements for addressing Benefit-Risk?

Benefit-risk assessment is a qualified assessment of the corresponding clinical evidence and acceptability of the benefit-risk ratio of the intended purpose as to whether the clinical benefit is achieved. For IVDs, the clinical benefit is the extent of accurate medical information on patients (IVDR; recital 64)[1], any other benefit to the patient should also be considered in benefit-risk assessments (IVDR; Annex I (1))[1]. The positive impact to the patient, including any benefits to patient management and public health is therefore the overall benefit (IVDR; Article 2, Definitions (37))[1] to be compared to a product's known and foreseeable risks when used for its intended purpose during normal conditions of use (IVDR; Annex I (8))[1]. Undesirable effects shall be minimised and be acceptable when weighed against the evaluated potential benefits to the patients and/or the user arising from the intended performance of the device during normal conditions of use (IVDR; Annex 1 (8))[1].

Rather than requiring each individual benefit and risk to be compared against one another, the Regulation defines the benefit-risk determination to be the overall benefit-risk determination (IVDR; Article 2, Definitions (17))[1]. Where an individual critical risk may not meet the initial acceptance criteria, this residual risk must be justified and be addressed accordingly. This is in line with EN ISO 14971:2019, section 8, which states that the overall residual risks should be compared against the benefits of the device to evaluate whether a high-risk but highly beneficial medical device should be marketed.

Example 1 – HIV lateral flow self-tests:
The risk that a self-test user will receive a false negative result cannot be fully mitigated. However now that HIV ART can result in undetectable viral load, meaning HIV cannot be transmitted, not knowing your HIV status is the main barrier to reducing disease transmission rates in Europe. Therefore, access to testing outweighs the risk of false negative result (assuming the overall approach to design, manufacture, risk management and PMS is appropriate).

Practicability[11] is also taken into consideration, the IVDR states: 'risks are to be reduced as far as possible without adversely affecting the benefit-risk ratio' (Annex I (2))[1]. It is reasonable to interpret that the economic practicability in such decisions includes reference to the benefits for public health and for society as a whole. However, section C4 of ISO/TR 24971[4] goes on to state that the 'economic practicability should not be used as a rationale for the acceptance of unnecessary risk'.

The IVDR is clear in stating that the benefit-risk assessment should be carried out under normal conditions of the intended use of the device (IVDR; Article 56 (1))[1] and (IVDR; article 57 (2).[1] It is therefore important to identify the hazards from normal use[12], see table below.

Use errors may occur during normal use, see Table 1.

ISO/TDR 24971 Section H [4]	Hazard Identification	Examples
2.3.3	From normal use	Inherent false positive/negative rates, measurement uncertainty, within/outside normal range when using 95% normal range, known interference, biological variation, matrix effects, instrument reliability
2.3.4	From use errors	Performing operations out of sequence due to unclear instructions, data entry errors, applying insufficient volume manually or through automation

Table 16. Hazard identification examples in normal use and from use errors (modified from ISO/TR 24971)[3].

When planning clinical studies, it may be beneficial to define the conditions of normal use within the clinical study documentation and include the justification as to how the study itself represents this use. Where the study does not reflect the 'normal use situation' then further justification or evidence may be useful.

Documentation of benefit-risk ratio is required under the IVDR[1] as part of the general safety and performance requirements (Annex I (1); Annex I (8))[1], as part of the general risk management system (Annex I (3e))[1], and as part of the technical documentation (Annex II (5a)[1].

5) Are there any specific requirements for companion diagnostics (CDx)?

Within the IVDR[1], 'companion diagnostic' means a device which is essential for the safe and effective use of a corresponding medicinal product.

[11] Practicability has two considerations: technical practicability and economic practicability (ISO/TR24971:2020 Medical devices – Guidance on the application of ISO 14971], section C3)[4]. Technical practicability refers to the ability to reduce the risk regardless of cost. Whereas the economic practicability refers to the ability to reduce the risk without making the medical device an unsound economic proposition, because the risk control measure(s) would make the medical device too expensive for widespread use.

[12] 'Normal use' is not defined in the Regulation. In EN ISO 14971, section 6.2[3], it is understood as being used for the intended use. A further definition is found in IEC 62366-1, section 3.9[5]. Here, normal conditions are understood to mean according to the intended use and instructions for use.

To meet the general requirement for performance evaluation and clinical evidence (Article 56 & 57)[1], CDx performance evaluation studies may require studying the IVD device in relation to and/or together (co-developed) with the corresponding drug or therapy to determine the efficacy and safety of the drug or therapy. As such, the intended purpose/use, medical treatment and outcome of the patient need to be taken into consideration for studies involving CDx. Additionally, Article 58[1] & Annex XIV[1] may be applicable to clinical performance studies aimed at demonstrating the clinical benefit of a CDx.

For follow-on CDx devices seeking the same intended use as the co-developed CDx based on analytical concordance studies, an assessment should be performed to ensure there is no impact on the safe and effective use of the companion therapy.

6) What are the requirements for addressing benefit-risk prior to product launch (performance evaluation)?

The IVDR (Article 56 & 57)[1] requires "…confirmation of conformity with relevant general safety and performance requirements as set out in Annex I [1]. Annex 1 (8)[1], requires 'all known and foreseeable risks, and any undesirable effects to be minimised and be acceptable when weighed against the evaluated potential benefits' and 'the intended performance of the device during normal conditions of use'. For this to be achieved, sufficient clinical evidence[13] is required within the performance evaluation and 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.

As per the IVDR, Annex XIII (1.1)[1], the performance evaluation plan shall include acceptability parameters of the benefit-risk ratio for the intended purpose and performance of the device, see Figure 1. Also, the method of this assessment should be included. A possible approach could be to use a risk acceptability table. The purpose of such a table would be to document the probability of harm vs. the severity of harm for each risk, the acceptability of which is driven by the benefits of the device.

A description of the expected benefits and risk is to be documented as part of the clinical performance study plan (Annex XIII (2.3.2h)[1], and sufficient data demonstrating that the device achieves the intended clinical benefit(s) and is safe is to be documented as part of the clinical evidence and performance evaluation report (Annex XIII (1.3.1))[1]. With respect to Figure 1, outputs of the performance evaluation would be considered in the first diamond (left-hand side).

For CDx, Article 58[1] & Annex XIV[1] may be applicable to clinical evidence generation to demonstrate the clinical benefit of the device

7) Are there performance study specific requirements for subject participation where benefit-risk should be considered?

[13] The required level of clinical evidence is outside the scope of this document and is addressed in the MedTech Europe WG guidance document titled 'Clinical evidence levels under the Regulation 2017/746 on in vitro diagnostic medical devices'.

There are additional requirements for certain 'higher risk' studies, as set out in Article 58[1], and detailed in Annex XIV[1]. This is a separate aspect of benefit-risk as it considers the risks and benefits for a representative population and forms part of the performance study plan.

8) What are the requirements for addressing benefit-risk post product launch?

During post-market surveillance, the benefit-risk assessment shall be updated actively and systematically (Article 78 (3a) reading onto Article 78 (2))[1]. The meaning of the term 'actively and systematically' is interpreted as being defined by the manufacturer in the post-market surveillance plan (PMSP) and is expected to include the defined depth and frequency of review, see Figure 1. The requirement in Annex XIII (4)[1] 'benefit-risk ratio is to be continuously monitored' may be interpreted similarly.

Per Article 78 (1)[1] the PMSP should be proportionate to the risk class and appropriate for the type of device. For the higher risk classifications, class C and D devices, the periodic safety update report (PSUR) should be updated throughout the lifetime of the device, and the conclusions of the risk-benefit assessment shall be set out (IVDR; Article 81 (1a))[1].

If the benefit-risk assessment changes significantly, and has the potential to lead to unacceptable risk, then it should be reported (IVDR; Recital 82)[1], see Figure 1. To allow determination of reportability, the PMSP shall describe suitable threshold values/parameters for continuous assessment to determine if action should be taken (IVDR; Annex III (1b))[1].

Where defined thresholds are crossed, manufacturers should report this by means of the electronic system. This should also apply to any statistically significant increase in the frequency or severity of incidents that are not serious incidents that could have a significant impact on the benefit-risk (Article 83)[1].

For CDx, related serious incidents and field safety corrective actions (FSCA) associated with the drug should be considered. This is addressed in Article 84 (6)[1]: "In the case of companion diagnostic, the evaluating competent authority or the coordinating competent authority referred to in paragraph 9 of this Article shall, depending on whether the relevant competent authority of the Member State that authorised the medicinal products or the EMA was consulted by the notified body in accordance with the procedures set out in Section 5.2 of Annex IX[1] and Section 3.11 of Annex X[1], inform that national competent authority or the EMA, as appropriate".

9) What are the Notified Body (NB) considerations?

The IVDR requires the NB to verify the adequacy of the benefit-risk determination through assessment of the technical documentation (Annex IX (4.6))[1].

10) Is there guidance for carrying out the assessment of benefit-risk and how might this relate to IVDR requirements?

Article 2 (37)[1]. EN ISO 14971:2019[2], alongside ISO/TR 24971:2020, are helpful pieces of information on benefit-risk analysis. The IVDR applies a specific meaning to the concept of clinical benefit, Recital 64 IVDR[1], which differs from therapeutic devices. Both medical devices and IVDs are within the scope of EN ISO 14971:2019[2] and there are elements of the standard relating to medical devices which would not be relevant for IVDs. Unless present in an intended purpose or other product claim, the downstream clinical benefits on the final patient outcome are not taken into consideration in a benefit-risk analysis. However, downstream risks should be considered as part of the overall risks.

The standard does not outline the criteria for benefit-risk judgement as they would be specific to the product in question and its anticipated conditions of use. Criteria are therefore left to those writing the benefit-risk statements as they are best informed of the detailed performance of the device; examples of risks and benefits that may be considered are provided in Annex I of this document. As stated in question 1 of this chapter, residual risks may be justified in the risk-benefit analysis once all practicable measures to reduce risk have been applied. Verification of the anticipated performance or effectiveness through a simulation study or a (clinical) investigation may be useful where significant residual risks are present to confirm that the benefit-risk balance is as expected and to prevent unwarranted exposure of patients to a large residual risk.' In the context of the IVDR[1] this requirement may be addressed by the objectives of the post-market performance follow up (Annex XIII, Part B section 4) where the plans may be aimed at reducing the uncertainty of risk estimation by carrying out further studies.

Direct comparisons of benefit and risk can only be achieved if they are on a common scale. If a common scale is used, then the benefit- risk assessment may be quantitative. For IVDs, however, it is more likely that indirect benefit-risk assessments are made, and these are qualitative and not quantitative. Comparisons may be achieved using information available in the literature, comparison to current technology and data from clinical studies. Where risks are known, a measure of the benefit may be established from the reverse of the risk, for example, by comparing benefit of the availability of the device compared to the risks incurred due to its unavailability.

Some benefits may only be for a proportion of the patient population, for example the subset of the population where the IVD provides an increased sensitivity for a condition. Also, an improved precision of an assay may benefit the population as it may allow resources to be focused in a more efficient manner, or on an individual level may allow the patient to move more quickly down the right patient management pathway.

Within the benefit-risk assessment it may be helpful to include characterisation of the disease or condition of the patient, and for high-benefit/high-risk devices the labelling should include adequate information to users and patients of significant residual risks in the accompanying documentation (EN ISO 14971:2019 Section 5)[2].
Review of other sources of information on risk-benefit decisions identified several examples from US FDA guidance documents, which provide helpful insights but are not legally binding for the European Regulation:

A) Guidance for Industry and Food and Drug Administration Staff: Benefit-Risk Factors to Consider When Determining Substantial Equivalence in Premarket Notifications (510(k)) with Different Technological Characteristics; Document issued on September 25, 2018.[5]

For diagnostic devices specifically, this Guidance discusses benefit(s) in reference to the nature of the public health impact, and could be based on a number of factors including:

- Identification of a specific disease;
- Provision of diagnosis at different stages of a disease;
- Prediction of future disease onset;
- Improvement of patient workflow;
- Increase in efficiency or examination;
- Provision of reproducible and quantifiable results contributing to the optimization of therapy and treatment; and
- Improvement of patient outcome (e.g., well-being, health status, safety of patients) by:
    - Facilitating fewer missed diagnoses (or the right diagnosis the first time, hence the correct treatment plan) and/or
    - Identification of patients likely to respond to a given therapy and therefore enable treatment of the disease or reduce/prevent its spread, which can often be measured through the use of patient-reported outcomes (PROs)

Guidance for Industry and Food and Drug Administration Staff: Factors to Consider Regarding Benefit-Risk in Medical Device Product Availability, Compliance, and Enforcement Decisions; Document issued on December 27, 2016.[5]

This Guidance document assesses Information Concerning Extent of Probable Benefit(s) by considering the following factors individually and in aggregate:

- Magnitude of the Benefit
    - Defined by the accuracy and reproducibility of test results and by the expected effect of clinically applying those results
- Probability of the Patient Experiencing One or More Benefit(s)
    - Which patients may experience a benefit (patient subgroups may experience different benefits or different levels of benefits)
    - Large benefit may be experienced by a small proportion of participants vs. small benefit experienced by a large proportion of participants
- Duration of Effect(s)
    - How long can the benefit be expected to last for the patient? Does the treatment need to be repeated?

An aspect of the PMSP is to carry out reviews and updates of the benefit-risk analysis (Figure 1). Here, it is important to identify any new or unanticipated risks. It is also important to confirm that the anticipated benefits are achieved and whether any additional benefits are observed. The following table may provide a useful approach when reviewing the benefit through the post-market surveillance process.

Anticipated benefit	Initial assessment during pre-launch	Current assessment	Does the marketed device/product achieve the anticipated benefits?
Type of benefits	What is the medical device's anticipated impact on clinical management and patient health? What benefits were initially anticipated? What benefits were expected based on similar devices?	Using real-world data or other available data, what is the medical device's impact on clinical management and patient health? Have additional benefits been observed?	

Table 17. reviewing the benefit through the post-market surveillance process

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. EN ISO 14971:2019 Medical devices - Application of risk management to medical devices
3. ISO/TR 24971:2019 Medical devices -- Guidance on the application of ISO 14971
4. EN IEC 62366-1:2015 Medical devices – Part 1: Application of usability engineering to medical devices

   Guidance for Industry and Food and Drug Administration Staff: Benefit-Risk Factors to Consider When Determining Substantial Equivalence in Premarket Notifications (510(k)) with Different Technological Characteristics; Document issued on September 25, 2018 https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM404773.pdf
5. Guidance for Industry and Food and Drug Administration Staff: Factors to Consider Regarding Benefit-Risk in Medical Device Product Availability, Compliance, and Enforcement Decisions; Document issued on December 27, 2016. https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm506679.pdf

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 the 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 the 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.

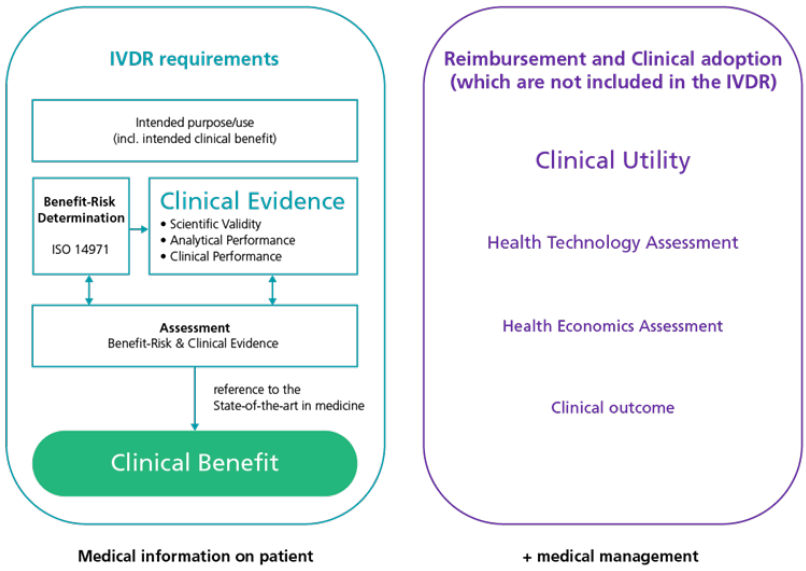


Figure 12. Clinical benefit concept under the IVDR and its distinction from clinical utility

NOTE: 'Clinical benefit' refers to the positive impact of a device related to its function in providing accurate medical information on patients.

Where specific patient management steps are included in the manufacturer's intended purpose (for example companion diagnostic IVD), then 'clinical benefit' may also refer to the benefits that arise as a result of that patient management.

The below table lists the descriptions of the common test purposes for IVDs as defined in GHTF/SG5/N8:2012. The considerations in determining the benefits and risks are also provided. Table indicating how benefit-risk varies across product groups:

Test Purpose	Description	Benefit	Risk
Diagnosis	A common test purpose or function for an *in vitro* diagnostic medical device, whereby the test is used solely or principally to determine, verify or confirm a patient's current clinical condition. Note: Adapted from GHTF SG5 N8R3[7] In addition, the IVDR includes consideration of physiological or pathological process or state. Where an assay is used in diagnosis it may either be used in isolation or form an essential element (for example as part of an algorithm or guideline) that allows a diagnosis to be made.	Provides accurate information on the patient's status that allows the treating clinician to make a diagnosis (determine, verify or confirm a patient's condition) that may be used in isolation or as an essential element alongside additional available information (or patient management decision if stated in the manufacturer's intended purpose and claims).	An erroneous result may lead to an incorrect diagnosis, inability to diagnose, or delay in reaching the correct diagnosis. Depending on the urgency of the result, unavailability of an assay may lead to a delay in reaching the correct diagnosis, where this results in an unmet need. Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as false results. Such tests may be urgent and thus delayed or unavailable results could result in less-informed patient management decisions.
Aid to Diagnosis	A common test purpose or function for an in vitro diagnostic medical device, whereby the test is used to provide additional information to assist in the determination or verification of a patient's clinical status. NOTE: Adapted from GHTF SG5 N8R3[7]	Tests provide accurate information on the relevant biological/congenital/physical parameters that facilitate interpretation, diagnosis and related	An erroneous result may lead to a delay in reaching the correct diagnosis while other assessments are carried out. The clinician may have to explain the incongruous result to the patient.

	Aid to diagnosis tests are used to provide additional information to assist/facilitate in the determination or verification of a patient's clinical status, physiological or pathological process, or state/congenital physical or mental impairments. The test is not the sole determinant. These tests are designed to evaluate a patient's current state.	patient management decisions while taking into account the overall clinical picture.	Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as erroneous results. Such tests may be urgent and thus delayed or unavailable results could result in less informed patient management decisions.
Screening	A common test purpose or function for an *in vitro* diagnostic medical device, whereby the test is used to detect the presence or absence of an analyte (measurand) in asymptomatic patients. NOTE: examples include tests for genetic screening, tests for early detection of disease, and tests used to reduce the risk of infectious disease transmission, such as assays for prenatal screening and donor screening (transfusion or transplantation). NOTE: Depending on the nature of the condition and the targeted patient population, screening tests may be used routinely or may be restricted to "at risk" patients. NOTE: Adapted from GHTF SG5 N8R3[7]	Provides additional insight regarding the patient's status to the patient management team. Although screening is not necessarily diagnostic, it may lead to a more efficient patient pathway, and subsequent appropriate diagnostic pathway that could lead to public cost or health benefits as well as individual benefit.	Erroneous results (e.g. a device not meeting claimed performance) may lead to delays in the patient following the most appropriate patient pathway, potentially leading to delayed diagnosis. Erroneous results may lead to further unnecessary follow up which may worry the patient and lead to unnecessary costs. Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as erroneous results.
Monitoring	A common test purpose or function for an *in vitro* diagnostic medical device, whereby the test is used for serial measurement of the analyte (measurand) levels in order to detect/assess	May allow more appropriate/effective treatment or patient management decisions.	False results may lead to inappropriate or less effective patient management or interventions.

	disease progression, regression, recurrence, minimal residual disease and/or response or resistance to therapy. NOTE: These tests are designed to evaluate changes in a patient's state. NOTE: adapted from GHTF SG N8R3[7] Monitoring tests are used for the measurement of analyte levels for the purpose of adjusting treatments/interventions as required.	For example, this may contribute to better and stable physiological status of the patient (e.g. diabetic or HIV-1 suppression).	Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as erroneous results.
Predisposition	A common test purpose or function for an *in vitro* diagnostic medical device, whereby the test is used to determine the likelihood of disease onset (i.e. assessing the risk of developing the disease in the future) in pre-symptomatic patients. NOTE: For patients at sufficient risk (as determined by test results), preventive interventions may be taken. NOTE: These tests are designed to evaluate a patient's future state. NOTE: Adapted from GHTF SG5 N8R3[7]	May allow decisions to be taken on lifestyle changes and treatment options, benefit may have the potential to be both personal, familial and to overall public health. May allow decisions on closer monitoring that could facilitate an improved patient workflow. For patients at sufficient risk, as indicated in medical guidelines or from clinical evidence,	False positive results could introduce undue concern and unnecessary treatment or monitoring. False negative results could lead to less efficient patient workflow. Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as false results. Generally, not considered urgent tests and thus delayed or unavailable results would have negligible risks provided the result was not irreplaceably lost.

	IVDR, Article 2, (2)[14] to determine the predisposition to a medical condition or a disease.	preventive interventions may be taken. Negative results may reduce worry for the individual and their family.	
Prediction (of Treatment Response or Reaction)	A common test purpose or function for an *in vitro* diagnostic medical device, whereby the test is used to measure factors that determine the likelihood of patient responses or adverse reactions to a specific therapy. NOTE: These tests are designed to evaluate a patient's future state. NOTE: Adapted from GHTF SG5 N8R3[7] IVDR, Article 2, (2) (e)[15] to predict treatment response or reactions Predictive tests designed specifically for use with a targeted therapy are sometimes termed	Provide accurate information to the physician to make informed decisions on patient management. This may lead to more effective patient management and reduction of patient risk by reducing the impact or side effects of non/less effective patient management (treatment) strategies. Some (e.g. CDx are guiding patient	An erroneous result (e.g. a device not meeting claimed performance) may lead to the wrong/less effective patient management strategy. In the case of a CDx an erroneous result may lead to less appropriate or inappropriate treatment. Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as false results. Such tests may be urgent and thus delayed or unavailable results could result in less informed patient management decisions.

[14] Products may not fall into neat categories or may fall across several categories.

	'companion diagnostics' (CDx) or 'personalized medicine'.	management, e.g. therapy) Others are more predictive or prognostic and thus enable the physician to take informed decisions on patient management.	
Prognosis	A common test purpose or function for an in vitro diagnostic medical device, whereby the test is used to measure factors linked to clinical outcome irrespective of treatment. Such tests may be used to estimate the natural progression of a disease (i.e. outcome in the absence of treatment), or to determine the likelihood of a clinical outcome irrespective of therapeutic intervention. NOTE: These tests are designed to evaluate a patient's future state. NOTE: Adapted from GHTF SG5 N8R37 A subset of prognosis may be the Risk assessment. This is considered a separate test purpose heading by the FDA and is described as the purpose '*to determine the risk for progression to a particular pathological or physical status within a short timeframe while*	May allow the individual, family or patient management team to take more informed decisions on the potential clinical pathway. It may prepare the subject, family or patient management team for the likely progression of the condition. *May be used by the patient management team to determine the risk of progression to a particular pathological or physical status within a short timeframe while under treatment/assessment for another condition.*	False results may lead to more poorly informed decisions on the possible clinical pathways. False results may incorrectly prepare the subject, family or patient management team for a progression of the condition. Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as false results. Generally not considered urgent tests and thus delayed or unavailable results would have negligible risks.

	under treatment/assessment for another condition.'		
Determination of physiological status	A common test purpose or function for an in vitro diagnostic medical device, whereby the test is used to evaluate the physiological state of an individual for the purpose of identifying a human condition or characteristic. NOTE: These tests are designed to evaluate a patient's current state. NOTE: Adapted from GHTF SG5 N8R37 IVDR, Article 2, (2) (a)[1] concerning a physiological or pathological process or state. E.g. hCG test for the determination of pregnancy.	The physiological state may aid in the identification of the individual's condition or characteristic. This may help point the patient management team towards the underlying cause of presenting symptoms. This may alert the patient's management team of an underlying abnormal condition or status, which may contribute to appropriate intervention or patient management decisions.	Erroneous results (e.g. a device not meeting claimed performance) may contribute to patient management decisions that could have the potential to further exacerbate a patient's abnormal physiological state. Incorrect labelling may result in correct results being interpreted incorrectly and thus have similar effects as false results. Such tests may be urgent and thus delayed or unavailable results could result in less informed patient management decisions.

Table 18. Benefit-risk differences across common IVD purposes

Footnotes:

1. Products may not fall into neat categories or may fall across several categories.
    a. For example, glucose assessments may be discrete assays used for a single determination. Or they may be used in monitoring; such monitoring may be discrete assessments or continuous monitoring.
2. The details on the benefits and the risks are product-specific as they will be dependent on the intended use/purpose and the extent of the claims within this. Aspects that may be considered include analytical and clinical performance, for example false positive and false negative incidence under normal conditions could be used to numerically estimate the incidence of benefits and risks.
3. The above benefits and risks are in relation to application of the assay result and not the use of the IVD.
    a. There are other potential benefits for the user and public health such as ease of use, cost, time, environmental etc.
    b. There are other potential risks to the user such as chemical, biological and physical

Chapter 12 – Near-Patient Testing (NPT)

I) Definition of NPT

1) How is NPT defined?

The IVDR defines a device for near-patient (NPT) testing as follows:
Article 2 (6) 'device for near-patient testing' means any device that is not intended for self-testing but is intended to perform testing outside a laboratory environment, generally near to, or at the side of, the patient by a health professional.

2) How does POC differ from NPT?

Point of care testing (POC or POCT) is a term in the IVD industry, referring to smaller devices used by healthcare professionals and employed near the patient.

As of today, POCT is not defined in any regulation that addresses the provision of devices to the market but rather by standards or guidelines that target quality practices in laboratories.

EN ISO 22870:2016 (1) provides one definition for POCT and NPT. It defines testing that is performed near or at the side of a patient with the result leading to a possible change in the care of the patient, suggesting that both terms can be used interchangeably. This standard is addressed to facilities working with such devices and is foreseen to be used in conjunction with EN ISO 15189 (2) and has no direct impact on IVD manufacturers.

With IVDR, the term NPT is introduced into a regulation. IVDR distinguishes in its definition of NPT only between the different environments of use, not between different health professional users.

However, IVDR demands that NPTs are accompanied by instructions where the manufacturer should make clear the level of training, qualifications and/or experience required by the user.

Therefore, during development and validation testing, manufacturers need to decide on the environments in which the product is intended to be used as well as the intended users. Based on this decision, the manufacturer will aim to fulfil the NPT requirements or not – if, for example, the product will be used only in a laboratory environment by laboratory professionals. The intended user and use environment will be specified in the technical documentation of the device.

From the manufacturer's point of view, both terms NPT and POC can be seen as synonyms considering the requirements for NPTs coming from ISO Standards, as applicable, for design input requirements.

3) Within the European Union, what does NPT mean, how is this different from US CLIA waived tests?

In Europe, an NPT must only be operated by a healthcare professional as defined by national authorities (e.g. doctors, nurses, paramedics), whereas in the US, CLIA guidance allows the use of POC tests by either trained or untrained operators. Trained operators may include clinical laboratory professionals, doctors, nurses, medical assistants, among others. Untrained users are office assistant type staff, teachers, or prison wards, etc. For further information on the qualification of an NPT, please refer to point 5 of the chapter.

4) Which are the main standards specific to the point of care testing?

a. EN 13532:2002 General Requirements for IVD medical devices for Self-Testing – Not updated to reflect IVDR.
b. EN 13612:2002 Performance Evaluation for IVD medical devices including Self-Test – Not updated to reflect the IVDR.

Three Standards regarding end user requirements to set up and run a POC/NPT Testing service:

c. ISO 15189:2012 – Medical Laboratories. Requirements for Quality and Competence – this standard can be used by medical laboratories in developing their quality management systems and assessing their own competence. It also touches upon POCT provision as part of a laboratory service. The associated ISO 22870 goes further in stating the requirements to establish POCT provision (under laboratory supervision) and should be read alongside ISO 15189. These two standards are increasingly being used for accreditation of laboratory and laboratory supervised POCT services, although alternative national requirements exist in many countries.
d. ISO 22870:2016 Point of Care Testing (POCT). Requirements for quality and competence.
e. PD ISO/TS 22583:2019 Guidance for Supervisors and operators of point of care (POCT) devices.
f. ISO/IEEE 11073 Health informatics — Point-of-care medical device communication series

At this time, unlike for medical devices and IVDs for self-testing, there are currently no NPT specific standards which take into account the specific design requirements and the working environment which NPT equipment can be used. Current standards are focused on IVD use in the laboratory setting. Companies may wish to take insight from these other standards, which have already identified a number of critical factors associated with NPT settings (home, ambulance, air ambulance).

g. ISO 15197:2013 In vitro diagnostic test systems — Requirements for blood-glucose monitoring systems for self-testing in managing diabetes mellitus.
h. ISO 17593:2007 Clinical laboratory testing and in vitro medical devices — Requirements for in vitro monitoring systems for self-testing of oral anticoagulant therapy.

i. see ISO 18113:2022-1-5
j. CLSI
   - POCT series – Mainly guidance for end-users in the USA. It includes widely accepted industry standards such as POCT-1-A2 POCT instrument interface standard, which replaced the previous ASTM standard.
   - EP series for performance evaluation – aimed at both industry and end user verification.
   - GP series including GP42 7th Ed. on Capillary Sampling.
k. FDA Guidance
   - Clinical Laboratory Improvement Amendments (CLIA)
   - Blood Glucose Monitoring Test Systems for Prescription Point-of-Care Use
   - Self-Monitoring Blood Glucose Test Systems for Over-the-Counter Use
l. MedTech Europe guidance on Annex I of IVDR (note: document only available to MedTech Europe members)

II) NPT user definition & training

5) How is the NPT user defined? How is the user qualified? And how is the profile different from a trained lab technician?

According to IVDR, the user of an NPT is a healthcare professional. The criteria and qualifications for *healthcare professionals* in the near-patient setting will likely come from local and member state requirements and regulations and may or may not include laboratory training. Requirements for a *trained laboratory technician* will also likely come from local and member state requirements but do include laboratory training. Moreover, ISO 15189 requires that laboratory personnel must be trained in the following areas with a periodic review of their skills to ensure their skills remain effective:

- the quality management system
- assigned work processes and procedures
- the applicable laboratory information system
- health and safety, including the prevention or containment of the effects of adverse incidents
- ethics
- confidentiality of patient information

Other relevant definitions from ISO 18113-1 (3) are:

3.1.28 healthcare provider - individual authorised to deliver health services to a patient
EXAMPLES Physician, nurse, ambulance attendant, dentist, diabetes educator, laboratory technician, medical assistant, medical specialist, respiratory care practitioner.

3.1.68 professional use - designation that an IVD medical device is intended for personnel who are qualified to perform IVD examinations through special education and training

6) How can the manufacturer best instruct on appropriate specimen collection and testing, taking into consideration the educational/training level of the NPT user?

According to IVDR, the user of an NPT should be a healthcare professional. It can be assumed that the healthcare professional user has some level of education or training which equips them to work in this field. Therefore, the manufacturer must determine who is the appropriate target user for their NPT and write instructions for specimen collection and testing accordingly.

7) What training on the device could be allowed, if any?

According to ISO 22870:2016, the laboratory director or another qualified person is responsible for appointing the person responsible for training and competency assessment. That being said, if the manufacturer wants to provide training materials, it is likely that this will be helpful to the person in charge of training.

From training, the user must attain the appropriate knowledge and skill requirements to understand the appropriate use of the device, including, where applicable:

- a specimen collection,
- its clinical utility and limitations,
- expertise in the analytical procedure,
- reagent storage,
- quality control and quality assurance,
- technical limitations of the device,
- response to results that fall outside of predefined limits,
- infection control practices, and
- correct documentation and maintenance of the results.

At a minimum, the user can be directed to read the instructions for use, but again, this will be at the discretion of the person responsible for training in the lab.

8) Is e-training sufficient in those situations where training is needed and allowed?

From MTE guidance on changes under IVDR which impact labelling: “According to the definition of a device for near-patient testing, the user of the device is a healthcare professional (ref. IVDR Art. 2(6)). This excludes laypersons, and it can be assumed that the healthcare professional user has some level of education or training which equips them to work in this field.”

Given the lack of standardisation in qualifications throughout Europe and the rest of the world, it may be challenging to cite a degree level. e.g., the UK and Germany take different approaches to education in nursing.

At a minimum, if no specific training is needed, the user may be directed to read the instructions for use. e.g., a rapid test intended to give a qualitative result/diagnosis for HIV is designed to be used in the field by a local healthcare worker who is not required to have specific training or qualifications; they should be guided to read the instructions for use before administering the test. The scope of the training will depend on the specific device and qualification of the user. In the case of a complex device, a detailed training should be provided, whereas for a simple rapid test, e-training which may include reading of the IFU, should be sufficient by agreement with the clinical institution.

Periodic training and update training should be considered based on device complexity, operating procedure complexity, risk profile of the device and end user circumstance, e.g., end user staff turn over and who is responsible for end user training. Consideration needs to be made whether it is the manufacturer or the distributor responsible for providing the training. Update training also needs to be considered where there is a significant change to the device, its operation, and or instructions for use. The need for post-market performance follow-up studies following significant change must be assessed, decision recorded and, where no studies are deemed necessary, justified in the technical documentation.

In the context of gathering clinical performance data, the training given to users during the study must reflect the normal training that will be offered to end users. I.e., staff operating the device during a clinical performance study must not be trained or coached beyond the expected level of training that will be routinely given to intended end users. The operators during the clinical performance studies must reflect the typical intended users, operating the device under the intended use conditions and intended use environment.

If some specific knowledge or training is required then this should be specified, e.g., the user needs to know how to use specific equipment such as a centrifuge or be qualified to take blood in order to use the device. The instructions for use may also indicate that specific training in accordance with the manufacturer's instructions for use is required. For example: a device intended for testing of cardiac markers in an emergency room will require the user to have specific training to use that device.

Finally, based on the manufacturer's risk management and the device intended purpose, it may be appropriate to note that results from use of the device must go to through a physician or that they must be sent to a clinical laboratory for further analysis.

Local requirements for training and access to a facility where questions can be asked and answered in an interactive manner, should be considered.

III) NPT testing location/environment

9) In the EU, what defines a Laboratory Environment? Is a GP Laboratory an NPT Environment?

IVDR defines devices for near-patient testing as any devices which are intended to perform testing outside a laboratory environment, generally near to, or at the side of, the patient by a healthcare professional. Outside

of the laboratory environment should be understood as outside of an accredited laboratory (based on national provisions); this could be an intensive care unit, emergency department or primary care settings such as a GP's office (4). Testing is performed by clinical staff (physicians, nurses), who are usually not laboratory trained (5). Additionally, it should be noted that, unlike the central labs, the GP's laboratory may not have sophisticated or automated equipment hence such GP laboratories should be considered as NPT environments.

10) What other standards/guidance can be used to help define testing locations of NPT under IVDR?

- MHRA "Management and use of IVD point of care test devices" 2013
- Point of care testing in primary care in the Netherlands document
- Larsson, A. et al The state of point-of-care testing: a European perspective 26/01/2015 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4389002/
- ISO/TS 22583:2019 "Circumstances where POCT testing can occur include but are not limited to hospitals, medical practices, pharmacies, paramedics, long-term care facilities, outreach clinics in remote and rural settings, in emergency and natural disasters and community settings such as law enforcement, workplace health and safety, sporting facilities, academia, the military and public areas such as shopping centres."
- IMDRF GRRP WG/N47 FINAL: 2018 "Near-patient testing: testing that is performed near a patient and outside of centralised laboratory testing facilities"
- EU working document on COVID testing kit performance "...in terms of location of testing, devices can be either laboratory-based or near-patient, also termed point-of-care, i.e. performed near a patient and outside of laboratory testing facilities. In the EU, near-patient tests are intended to be used only by a healthcare professional."

IV) NPT labelling requirements

The IVDR provides new labelling requirements for NPTs. These often mirror the requirements for self-tests.

- The label of the device needs to indicate if the device is intended for near-patient testing. This can be indicated by a symbol as appropriate (6);

- Language requirements for the label and instructions for use can be defined by Member States;

- In the instructions for use, the device intended purpose must include all the elements specified under Annex I, 20.4.1(c). The testing population must be specified here, along with the specimen. It is worth noting that the intended user must also be specified in the instructions for use (if not formally as part of the intended purpose); here, a broad and non-specific user group can be given, e.g., near-patient use, healthcare professionals, provided there is sufficient evidence to support their inclusion;

- The medium, format, content, legibility and location of the label and instructions for use must be appropriate to the device, its intended use and the technical knowledge, experience, education or training of the intended user(s) (ref. Annex I, 20.1 (a)). For devices intended for near-patient testing, the information given should be appropriate to the training of the intended user and the experience needed to use the device as intended (see also Annex I.19.1).

- The instructions for use cannot be provided solely in electronic format for near-patient testing (Annex I Chapter 3 20.1 (f)). Furthermore, there is a derogation that when supplying multiple devices intended for professional use within the laboratory environment to a single user and/or location, a single IFU may be provided by agreement with the user. This is specifically not the case for NPTs, meaning that instructions for use must always accompany each device. However, where multiple NPTs are provided to a single user or location, e.g. 25 rapid tests, the manufacturer, based on risk-management assessment and if duly justified, could provide a full copy of the instructions for use and 24 abbreviated operating manuals (also see Annex I 20.1d). In this case, the manufacturer should still be able to provide additional copies of the full instructions for use upon request, free of charge.

V) NPT performance indicators

11) Is the performance standard different for “near-patient” tests than “laboratory tests”? Does this depend on the analyte?

No, regardless of the testing location, tests must meet minimum requirements, which are dependent on associated guidelines and common specifications where applicable.

The performance criteria should support the stated intended purpose. For example, based on a ‘screening’ intended use confirmatory testing might be needed as a follow-up.

Also see chapters of this eBook: ‘state of the art in medicine’ and ‘analytical and clinical performance indicators’ under IVD Regulation 2017/746.

12) Can the same Product be used in both NPT and Laboratory Environments and have one Conformity Assessment?

Yes. Conformity should be assessed in its own right (Annex VIII Rule 4b).

One device can have a dual intended purpose. In this case, the device would be intended for use in different environments, both in the laboratory environment and by a healthcare professional outside of the laboratory environment. One conformity assessment is possible: the notified body will need to cover both the general device requirements as well as ‘additional’ requirements which relate to the different environments of use including NPTs. The requirements relating to NPTs are specified under IVDR Annex I:

- Section 19.2 provides requirements for NPTs

- Section 20 provides requirements for labelling. The labelling provided will need to be appropriate to both user environments. There are further specific requirements for NPTs.

In addition to the conformity assessment requirements for the class B or C device, the device will need to follow the procedures for technical documentation assessment set out in Section 5.1 of Annex IX.

13) Is the Conformity Assessment Route of Class A NPT product the same as for higher risk classes? (Combination of class A analyser with class B/C/D strips/reagents etc.)

No. All class A devices follow the 'self-declaration' route laid out under Article 48(10).

The class A device intended for near-patient testing does not require a notified body to conduct conformity assessments (unless sterility is claimed), nor does it need to follow the procedures for technical documentation assessment set out in Section 5.1 of Annex IX.

In general, instruments are expected to be class A (unless the instrument has an independent measuring function which does not use any additional reagents, e.g., instruments measuring blood gases or glucose via its sensors). Due to their interdependence, the notified body will assess the performance of the reagent on the instrument as part of the conformity assessment of the reagent. The manufacturer will be expected to provide evidence to support the use in combination claim between all devices used in combination (e.g., analyser and the software driving and influencing it, reagents, calibrators, controls, buffer/ washing solutions, etc.).

(Refer to MDCG 2020-16 Guidance for Classification rules and MDCG 2019-11 Software guidance)

14) How do the Instructions for Use and Intended Purpose requirements for NPTs translate to clinical performance studies? Specifically, will manufacturers have to do multiple clinical performance studies for different testing environments/locations, testing populations, and intended users, respectively?

The testing environment / location, testing population, and intended users are features that shall be included in the instructions for use and intended purpose/use for NPTs (IVDR Annex I, Chapter 2, Section 9.4 (b), Chapter 3, Section 20.4.1, (c) (vii), (e), respectively).

For qualification of users of the NPTs and streamlining of user skills and trainings, users of NPTs could be divided into two broad categories/groups:

- Users in routine professional care environments: Here, the training and user skills required are lower, and this group includes users in hospital wards, clinics, general practitioners' offices, pharmacies, retirement homes, rehab clinics etc.

- Users in critical care environments: Here, the training and user skills required are higher, and this group includes users in intensive care units, emergency units, urgent care centres, operating rooms, ambulances, etc.

This grouping is meant exclusively for the qualification of users.

For testing populations and testing locations, however, the performance indicators from one testing location within the same category cannot be grouped with or inferred from/transferred to another testing location in the same category. In other words, data from one routine professional care location (e.g., GP) cannot be grouped with or inferred from/transferred to other routine professional care locations (e.g., retirement home) without appropriate justification. Similarly, data from one critical care location (e.g., emergency room) cannot be grouped with or inferred from/transferred to other critical care locations (e.g., operating room). This is particularly true for analytes where performance indicators are already known to differ substantially between testing populations or among testing locations (e.g., troponins).

Thus, for each testing population and testing location claimed, the corresponding performance data will need to be provided unless duly justified, for example, in cases where it can be demonstrated that the skill level of the operator and the characteristics of the target of the test are substantially similar in the different NPT environments. It is conceivable that manufacturers launch NPTs with narrow and precise intended purpose claims based on clinical evidence generated in one testing population and location. Post-launch studies, including real-world evidence, could also help expand intended purpose claims to additional testing populations and locations.

15) Does Clinical testing have to take place solely in the anticipated “environment of use” if so, to what other setting or user group is the test clinically tested in this environment compared in order to determine performance claims? Can claims from one testing environment be transferred to another?

For each testing location or environment of use claimed, the corresponding performance data (analytical and clinical) will need to be provided. If equivalence between environments of use is established (through clinical performance studies and/or published literature), performance data, and therefore claims, can be transferred. Justification of equivalence must be included in the technical documentation of the device.

For analytical and clinical performance studies, a lab-based assay with similar intended uses can be used as a comparator, and data demonstrating operation by the intended users should also be generated.

16) What time-effective and cost-effective studies are required to provide suitable evidence for NPT devices?

The IVDR does not mention or define clinical utility.

Cost-effectiveness and time-effectiveness are related to clinical adoption and reimbursement; they are not required by IVDR for CE marking.

See page 22 of this eBook:
----"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." ----

----"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;)."

If samples or patients are difficult to obtain for the study, testing can be done on the manufacturer's premises or under other simulated conditions.

Other cost-effective approaches that can be considered include the use of data from non-EU studies that represent the intended use and EU population, and bridging studies where changes to the intended purpose increase scope.

17) What additional studies/evidence is required to differentiate between professional lab-based tests and NPTs?

An IVD is required to function in the use environment and by the user defined by the manufacturer. This functionality is required to be demonstrated in the use environment by the intended users by following the instructions for use. In addition, analytical performance studies and, in some cases, clinical performance studies need to be performed.

In addition, the usability and the use environment need to be taken into consideration when creating the evaluation/study protocols. This also means that analytical performance studies, i.e., the intended users and sites, need to be considered (physician offices, ambulances, hospital near patient testing, elderly homes, emergency rooms, etc.) when selecting testing sites.

When a test is intended to be used in the laboratory environment, the intended user group is laboratory professionals. Manufacturers providing the evidence may have their own product development groups that include laboratory professionals testing and verifying performance. Whether the manufacturer's own laboratory professionals represent the intended end user group in the verification and validation group and whether there is a necessity to perform external evaluation studies should be evaluated.

18) What are the key differences between usability and clinical performance studies for NPTs? And what does adequate usability documentation consist of?

Usability studies and testing are important processes within the product development process meant to verify the effectiveness of the design and to evaluate the ease of use of a product. Formative usability testing is done early in product development to help develop the product's shape and design. The goal is to detect issues and eliminate usability problems before a product is fully developed. It is crucial to observe and understand the users' thought processes and their actions resulting from them. The data collected during formative usability testing is observational in nature.

Summative usability testing is usually performed later in the product development process when the product is fully developed. It is often conducted when a design is reasonably complete and involves **evaluating the design against quantitative goals or competitor's products**.
The purpose of a "usability study" is to confirm correct operation of the device in the hands of the intended operator and effectiveness of the instructions for use. Any subsequent amendments in the IFU should be minor in nature and not impact the validity of the usability data obtained in the study.

Summative usability testing is typically carried out as a part of performance studies of the NPT. The spectrum of possible use sites and the level of education/training of the end-users should be taken into consideration when planning usability testing. Also, an NPT should be easy to use, and this aspect should be considered in design and usability.

If specimens, patients or study sites are difficult to obtain, testing can be done on the manufacturer's premises or under other simulated conditions.

Harmonised standard (EN BS EN 62366-1:2015+A1:2020 Medical devices - Application of usability engineering to medical devices EN) can be used to comply with documentation requirements by regulatory authorities.
The purpose of the clinical performance studies is to establish or confirm aspects of device performance, which cannot be determined by analytical performance studies, literature and/or previous experience gained by routine diagnostic testing (IVDR Annex XIII, 2.1.). Typically, clinical performance studies are studies in which diagnosis is available (through the clinical performance study or, e.g., through biobank samples) and can be used to calculate different diagnostic parameters for the test in question, e.g., diagnostic sensitivity and specificity and negative and positive predictive values.

19) Which reference methods are most appropriate for NPTs, US vs Europe?

A reference method is a scientifically established/recognised and standardised method for certain analytes and is selected according to the analyte in question. A comparative method is a method for a similar device on the market. The difference in the analytical performance data analysis of these two methods can be found in the different publications.

A reference method or reference material is required (IVDR Annex II, sec. 6.1.2.1 Accuracy of the measurement) to establish the traceability and trueness of a method. This rule applies to all IVDs, not only NPTs. If a reference method or reference material does not exist, traceability cannot be established. In this case, comparative methods accompanied by justification of the selected method may be used to establish the required performance.

The predicate method is a term used in US submissions for FDA marketing clearance. This term refers to a similar device (or test) already cleared for the US market. Predicate device comparison includes, e.g. information on similar devices and test performance. The test meant to be cleared in the state of art and risk-benefit sections is compared to the predicate device.

If the test in question has been cleared for the US market, information in the FDA database can be a useful starting point to identify potential systems to support equivalence. Method comparison to the predicate method can be utilised when establishing the state of art and risk-benefit. Further, the similarity table used in the predicate method could be utilised for legacy products.

20) Can participants be compensated in the EU? (recruiting patients for NPT device studies can be difficult)

Interventional specimen-taking procedures should be considered separately (differentiate between interventional study design).

Small compensation, e.g., travel expenses according tocountry-specific principles, lunch or coffee stamps etc., are allowed. When such studies in which an ethical committee statement is needed, compensation needs to be described in the study protocol (as in all studies), and the ethics committee will make an assessment if the compensation is appropriate. This is a general principle independent from the study type in question.

21) What are the specimen types that should be included in performance studies for NPTs (leftover samples vs fresh samples vs banked samples)?

Specimen type ultimately depends on the intended purpose of the device and could include, e.g., urine or blood (venous, arterial or capillary blood) specimens. Considering the settings where NPT are deployed and the turnaround time, fresh specimens are generally the most favourable specimen. For example, if NPT devices require the use of capillary or arterial blood, a fresh specimen should be taken for the purpose of a study. However, for devices utilising venous blood, leftover/banked samples or specimens may be considered for clinical performance studies, provided that they are deemed suitable for the analysis, e.g., heparinised/non-coagulated blood.
The study protocol should reflect the use case laid out in the instructions for use unless an appropriate justification for any deviation is provided.

If fresh specimens are collected prospectively, the following should be considered:

a. IVDR Articles 58 A & C: where the conduct of the study involves blood sampling and additional invasive procedures- venous blood sampling is now a high-risk procedure

b. ISO 20916 (7) 5.3: Design of clinical studies

22) Patient self-sampling (consider self-test requirements)

IVDR defines devices for self-testing as "any device intended by the manufacturer to be used by laypersons, including devices used for testing services offered to laypersons by means of information society services". According to the EU borderline manual (8), for a device to be considered a self-testing device, the lay user's action shall result directly in a test result or the lay user must manipulate the collected specimen before it is dispatched to a laboratory.

References:

1. EN ISO 22870:2016 Point-of-care testing (POCT) — Requirements for quality and competence

2. EN ISO 15189 Medical laboratories – Requirements for quality and competence

3. ISO 18113-1:2022 *In vitro* diagnostic medical devices — Information supplied by the manufacturer (labelling) — Part 1: Terms, definitions and general requirements

4. K Patel and B Suh Lailam, Implementation of point-of-care testing in a pediatric healthcare setting, https://pubmed.ncbi.nlm.nih.gov/30973797/

5. J.Shaw Practical challenges related to point of care testing https://pubmed.ncbi.nlm.nih.gov/28856189/

6. New IVD symbols for compliance with the IVDR https://www.medtecheurope.org/resource-library/new-ivd-symbols-for-compliance-with-the-ivdr/

7. ISO 20916:2019 *In vitro* diagnostic medical devices — Clinical performance studies using specimens from human subjects — Good study practice.

8. Manual on Borderline and Classification in the Community Regulatory Framework for Medical Devices (Version 1.22 (05-2019)

Chapter 13 – Use of Clinical Data from Outside the European Union

It is common practice today for clinical data coming from outside of the EU to be used to support performance evaluation claims for devices on the EU market[16]. For example, a multi-country performance study may have been run to develop data for a device that is intended to be placed on the market in a range of jurisdictions, including the EU. Or a device may be placed on the market of a non-EU country before it is introduced onto the EU market. In the latter case, the evidence collected to support the device will often be based on studies conducted outside of the EU. Depending on the intended purpose of the device, this data may be sufficient and can be justified without further studies being necessary. In other cases, a bridging study may be needed. This chapter discusses selected questions regarding the use of third-country[17] clinical data for the *In Vitro* Diagnostic Medical Devices Regulation (EU) 2017/746 (IVDR).

There are incentives for both industry and authorities to allow clinical data gathered outside of the EU to be used for the EU clinical data package:

1. Minimise duplication of performance studies,
2. Make new diagnostic tools accessible to patients faster,
3. Avoid wasting development resources.

The use of clinical data from outside the EU can only be made if that "data package" meets the local regulatory requirements while, however, fulfilling EU ethical standards.

1) Does the IVDR permit the use of clinical data collected outside of the EU?

Yes, the IVDR allows the use of clinical data collected outside of the EU.
For a list of references in the IVDR, see APPENDIX 11.1.

2) What is meant by the target population?

Under the IVDR, where applicable, the testing or target population is required to be specified as part of the device's intended purpose under IVDR Annex I. For example, a study design may include methods for determining the assay cut-off, which could include considerations around the target population. For CE-marking and in line with the product claims, the subjects of the performance study must be a representative sample of the target/testing population of the final CE-marked device.
Under the IVD Directive, Common Technical Specifications [1] prescribe the use of an equivalent European population to conduct a performance evaluation study for an IVDD Annex II List A device:

"3.1.6 Performance evaluations shall be performed on a population equivalent to the European population."

[16] For the purpose of this discussion the term 'EU market' is defined to be countries of the European Union, Switzerland and EEA countries.
[17] Third country here means a country that is not in the European Union, Switzerland or EEA.

ISO 20916 [2] provides considerations for how the clinical performance studies can be designed; this includes consideration of the target population. Examples of target population include age, race, gender, geography, clinical condition, and treatment status (reference ISO 20916:2019 - 5.3 Design of the clinical performance study 5.3 C 2) [AR1] [AR2]

Considering the element of geography, the manufacturer should check if clinical guidelines published by European medical societies need to be taken into account when using the data. Consequently, if there is an impact, an adjustment or bridging study needs to be considered.

3) What can we do with established (approved under IVDD) devices versus devices which will develop evidence entirely under the new performance evaluation procedures of the IVDR?

Devices on the market that already have CE marking under the IVD Directive will have followed the analytical performance requirements of that Directive.. Some level of clinical performance [SR3] will have been established in this regard, e.g., for diagnostic sensitivity and specificity. If needed, refer to Chapter 4 “Clinical Evidence Levels”.

A manufacturer can conduct studies under the IVD Directive and use the data also to demonstrate clinical evidence under the IVD Regulation. This is permitted until 26 May 2022, when the IVD Directive ceases to be applicable: the fact that a study is designed under the Directive or Regulation does not prevent the data from being used to meet clinical evidence requirements. The Regulation accepts many sources of data aside from clinical performance studies. Even if the studies were conducted outside of the EU, the transition from the IVDD to the IVDR does not per se require an amendment to the study protocol as long as the safety and performance of those devices regarding the European population can be demonstrated. So-called ‘legacy’ data are not excluded, and a retrospective amendment of the study protocol is not necessary. Data collected before the application of the IVDR, either within the manufacturing facility or published by scientists, collected considering the ethical and standard criteria should also be considered. Also, see Chapter 5 on “How to demonstrate evidence gained from published routine diagnostic testing”.

For devices that have no CE-marking under the IVDD (“novel devices”), it is recommended to follow the analytical and clinical performance study requirements under the IVDR, including design and documentation of the study to the extent possible. It should be noted that certain provisions set out in the IVDR for performance studies as per Article 58, such as notification and/or authorisation via EUDAMED, are only applicable to studies conducted in the EU Member States and the EFTA countries.

In the case where a performance study is needed, the use of data from outside of the EU is permitted as long as the study design and documentation requirements are fulfilled, provided that the study population is comparable to the intended European testing population of the device. The rationale for the study design should be provided as part of the benefit-risk determination under the clinical performance study protocol; this will be reviewed as part of the conformity assessment process by the notified body.

4) What are some ethnic factors which should be considered when using clinical data generated outside the EU?

Depending on the device in question, it may be necessary to consider genetic or physiologic factors (intrinsic factors), and cultural and environmental characteristics (extrinsic factors) when assessing the value and completeness of using clinical data generated outside the EU.

Genetic or physiologic factors:

To consider: is the analyte the same across populations in different geographies? Meaning, can data collected in one population be transferred to a different geographical region?
Below are examples of analytes illustrating the use of clinical data generated outside of the EU to support the intended clinical benefit of the test:

- **Alzheimer's disease, as detected by Abeta**

The cut-off limit (Abeta 42 over Tau) was established in a North European population and later verified in the United States. These populations represent different ethnic make-ups.

Consideration: Is Abeta equally presented in the North European- versus US population?

Published literature shows that Abeta is equally presented in both populations and clinical data from these populations is transferable. Also, medical practice in both regions is comparable. Furthermore, appropriate patients in the appropriate settings are not easily obtained, further adding justification for using non-EU cohorts.

- **Cytokeratin-19 for detection of cancer cells**

The clinical cut-off of Cytokeratin-19 expression was established in a Japanese population.

Consideration: is this clinical cut-off established in Japan also applicable to the EU population?

Published studies demonstrate that the expression level of this gene in tumour cells is identical among different ethnical populations. Therefore, the cut-off value for this gene is applicable to the EU population. Moreover, the study performed in Japan is in line with the EU requirements, and clinical practice between the two regions is comparable.

HBV genotype distribution

Consideration: Is the below clinical performance study conducted outside of the EU also applicable to the EU population?

Published literature demonstrates a wide distribution of HBV genotypes around the world, underscoring the need to ensure that clinical performance studies address the HBV genotype coverage specific to the EU population.

Analytical performance studies shall demonstrate that the device can detect all HBV genotypes (A-J) if the device's intended purpose claims to detect all genotypes.

Clinical performance studies were conducted outside of the EU in geographies with a similar but not identical prevalence of HBV genotypes. Combined with analytical performance, literature reviews (showing common genotypes between EU and outside of the EU) as well as peer-reviewed published literature demonstrating clinical performance using the device from various geographical locations that have genotypes common to the EU, provided support for the device's intended use. See Figure 13 for the distribution of HBV genotypes by country.

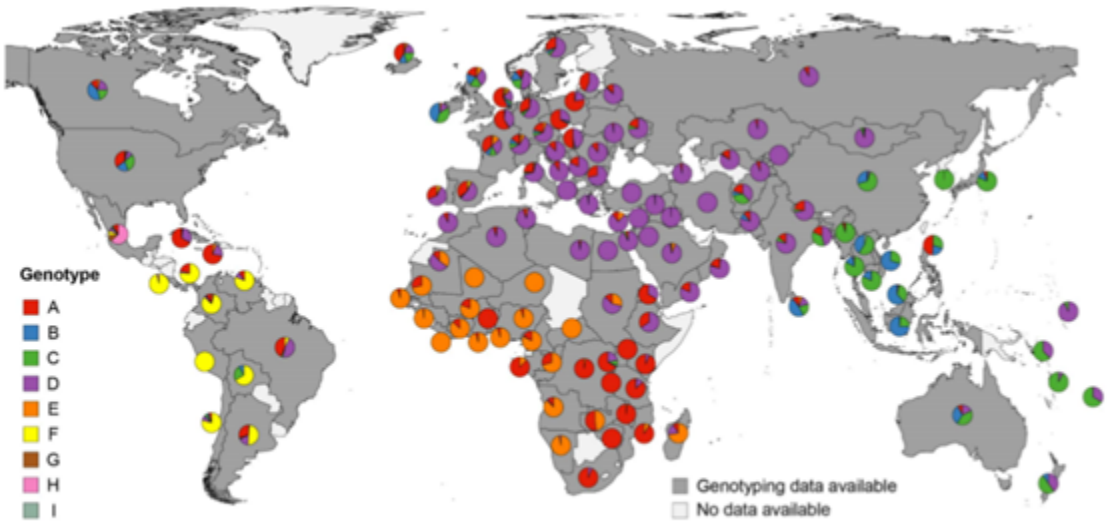


Figure 3. Distribution of HBV genotypes by country. Pie charts indicate proportional HBV genotype distributions in the respective countries. Genotype distributions within samples with successful genotyping are presented, excluding inter-genotype recombinant viruses, co-infections with more than one HBV genotype or undefined infections. Underlying literature sources and number of sequenced isolates are given in Table S1 (Supplementary Materials).

Figure 13. Distribution of HBV genotypes by country. Pie charts indicate proportional HBV genotype distributions in the respective countries [3]

Cultural and environmental characteristics:

To consider: examples of extrinsic factors include social and cultural aspects of a region such as medical practice, diet; and particularly important to the reliance on studies from a different region, practices in clinical trial design and conduct. Although it is the manufacturer's responsibility to ensure the clinical study is designed and conducted according to EU requirements, it is recognised that study sites with global variation may demonstrate an unconscious bias to the interpretation of the clinical study protocol provided by the manufacturer, conducting the study according to local cultural/environmental norms. This could lead to the

practical application of the study protocol/training as provided by the manufacturer to be somewhat different to the original intent.

Equitable selection of patients The UK Government has conducted an independent review for equity in medical devices (citation, March 2024). The review investigated equity principles for approved medical devices, such that they function to the same high standard and quality for all relevant population groups. This concept recognises that there are unavoidable differences in performance in relation to certain ethnic groups. As such, this principle can be applied to clinical performance studies, particularly those involving the application of polygenic risk score (PRS). The planning of such studies can be widened beyond genetic diversity to include the contribution of social determinants of health, including lifestyle, living and working conditions and environmental factors.

Medical practice

Medical practice in different regions needs to be considered in an early phase when the clinical performance study protocol is designed. Co-medication and invasive procedures might differ across regions, particularly in the critical care setting. When these aspects of the clinical performance protocol are defined, proactively researching the local clinical practice guidelines can reduce unnecessary exclusion of patients once the study is running and result in a more reliable estimate of the number of enrolled study patients.

Definition of clinical conditions might also be a complicating and confounding factor introducing bias in the clinical data. Even though well defined in study protocols, some heterogeneous conditions might still be defined differently around the world. Also, the treatment of these conditions (including medication) might vary and be influenced by historical medical practice.

Patients available for clinical performance studies might also represent different severity and clinical stages. This can be due to a lack of standardisation, or different scales or scoring practices.

Definition of the clinical cut-off might differ from region to region.

Clinical cut-off might be defined differently in different regions. The underlying reason for this difference might be as simple as different units are preferred (e.g., see Cholesterol below). In some countries, the cut-offs are influenced by limitations to the medical system, pushing out cut-offs to include only more advanced clinical conditions.

Example: Cut-off for total Cholesterol in the EU vs the US

The cut-off definition for desirable and borderline high Cholesterol differs slightly in the EU vs the US. This difference is driven by the units preferred in the two regions. The most suitable cut-off (number) is used to define the clinical condition, based on mmol/L in the EU or mg/dL in the US. This results in different cut-offs based on units:

Total Cholesterol (desirable/borderline high)

- **200 mg/dL** (5.18 mmol/L) National Cholesterol Education Program (NCEP), USA
- 5 mmol/L (**190 mg/dL**) European Society of Cardiology (ESC) and the European Atherosclerosis Society (EAS)

Distribution of eligible subjects

Prevalence of the disease, mutation, or infections might force a sponsor to search outside of the EU to find suitable patients.

Example: Due to extensive HPV vaccination in the EU, this results in a low prevalence of women suffering from cervical cancer. Therefore, HPV patients may need to be sourced outside the EU.

Dietary differences

Geographic differences in nutritional habits can impact IVD testing. An example of this was the increased use of Biotin as a nutritional supplement, which had a negative impact on the performance of IVD tests using biotin-streptavidin binding technology. Moreover, other interfering substances should be considered.

APPENDIX 11.1 – In Vitro Diagnostic Medical Devices Regulation (EU) 2017/746 (IVDR) – relevant references

***Article 56* Performance evaluation and clinical evidence**

1. 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.

To that end, manufacturers shall plan, conduct and document a performance evaluation in accordance with this Article and with Part A of Annexe XIII

Annex I CHAPTER 2

REQUIREMENTS REGARDING PERFORMANCE, DESIGN AND MANUFACTURE

9. Performance characteristics

9.1. Devices shall be designed and manufactured in such a way that they are suitable for the purposes referred to in point (2) of Article 2, as specified by the manufacturer, and suitable with regard to the performance they are intended to achieve, taking account of the generally acknowledged state of the art. They shall achieve the performances, as 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, measuring 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; and*

(b) the clinical performance, such as diagnostic sensitivity, diagnostic specificity, positive predictive value, negative predictive value, likelihood ratio, and expected values in normal and affected populations.

20.4. Information in the instructions for use

20.4.1. The instructions for use shall contain all of the following particulars:

(a) the name or trade name of the device;

(b) the details strictly necessary for the user to uniquely identify the device;

(c) the device's intended purpose:

(i) what is detected and/or measured;

(ii) its 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:

— 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; and

(viii) for companion diagnostics, the International Non-proprietary Name (INN) of the associated medicinal product for which it is a companion test.

Annex II:

"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.

Annex XIII

2.3.2. Clinical Performance Study Plan

(m) information on the performance study population: specifications of the subjects, selection criteria, size of the performance study population, representativity of the target population, and, if applicable, information on vulnerable subjects involved, such as children, pregnant women, immuno-compromised or elderly subjects;

References:

1. Decision (EC) 2002/364 Commission Decision of 7 May 2002 on common technical specifications for *in vitro* -diagnostic medical devices

2. ISO 20916:2019 *In vitro* diagnostic medical devices — Clinical performance studies using specimens from human subjects — Good study practice

3. Velkov, S.; Ott, J.J.; Protzer, U.; Michler, T. The Global Hepatitis B Virus Genotype Distribution Approximated from Available Genotyping Data. *Genes* **2018**, *9*, 495. https://doi.org/10.3390/genes9100495

4. UK Government Independent Report Equity in medical devices: independent review - summary report - GOV.UK (www.gov.uk) 2024.

Chapter 14 – Clinical evidence for NGS-based IVD assays – focus on oncology

In recent years, many gene assays have been replaced by multi gene assays that utilise next generation sequencing (NGS). These allow simultaneous detection and evaluation of genetic variations across multiple genes or genomic regions. Human genetic variations or variants can be germline or somatic in origin (see definitions below).

In general, NGS-based assays for clinical use include consumables, reagents, instruments, and software. Any two NGS-based assays with similar intended purposes may differ in their design and workflow since they will require a particular set of reagents, consumables, instruments and software attributes in order to function and achieve their intended purposes.

NGS-based assays differ from traditional IVD tests in many ways:

- NGS-based assays may identify thousands or even millions of genetic variants depending on the size of the region targeted
- NGS-based assays may detect previously unobserved genetic variations
- NGS-based assays may detect multiple alteration classes simultaneously such as single nucleotide variants (SNVs), insertions and deletions (indels), copy number alterations such as copy number losses and gains, genomic rearrangements, and various chromosomal abnormalities
- NGS-based assays may detect genetic alterations of various levels of evidence or clinical significance (e.g. strong clinical significance, potential clinical significance, variants of unknown significance, no clinical significance)
- NGS-based assays may have an extensive intended purpose as it may cover multiple test functions and claims (e.g. diagnostic, prognostic, predictive, risk assessment) and target multiple intended use populations (e.g. multiple solid tumour types, average and high-risk individuals for a given disease)
- NGS workflows consist of a range of technically complicated steps from specimen collection and processing, through to library preparation and sequencing, as well as variant calling and reporting
- The term NGS-based assays covers a range of sequencing technologies and method(s) used to capture genomic content. These include targeted methods, where specific regions are enriched during library preparation by using specific probes or primers, as well as broader approaches like whole exome and whole genome sequencing.. In whole genome sequencing, no specific enrichment is applied; instead, the entire genome is sequenced, and regions of interest may be selected during data analysis. Regardless of the method, reporting can be limited to a predefined subset of the sequenced data based on the clinical application.
-
- NGS-based assays employ sophisticated algorithms for secondary analysis, also referred to as variant calling, which require their own training, verification and validation upstream prior to end-to-end assay validation

- NGS-based assays generate many output files that include vast information at sample and variant levels

1) Which type of NGS-based assays are in the scope of this chapter?

In scope:

IVD* NGS assays for somatic (DNA and/or RNA) and germline (DNA) testing performed on appropriate samples, with appropriate intended purpose.

**As within the definition of an IVD described in EU Regulation 2017/746[1], Article 2.2.*

Out of Scope:

- Companion diagnostics (CDx)
- Products intended to be research use only (RUO)
- Laboratory developed assays that fall under the Healthcare Institution Exemption described in Article 5.5*
- Assays for pathogen detection, microbiology diagnostic tests

**While this* chapter *is directed towards devices with CE-marking, some considerations may be helpful also for devices falling under Article 5.5.*

Many NGS-based assays have one or more CDx related claims. However, this chapter will focus on NGS-based assay considerations for non-CDx claims that include test functions such as screening, risk assessment, diagnosis, prognosis, and/or monitoring. For CDx biomarker validation, chapter 7 of this eBook should be reviewed.

The determination of whether a product is an IVD with CE-marking, a device for performance study, an NGS assay manufactured under the healthcare institution exemption, or a product intended for research use only, should be conducted in line with IVD Regulation (EU) 2017/746, and consider published guidance such as MDCG 2022-10. This is ultimately a decision for the manufacturer of the product.

2) How are NGS-based IVD assays defined?

For the purpose of this eBook, we propose the following definitions:

NGS-based IVD – an IVD which detects a set of clinically relevant alterations and/or signatures in genes or genomic regions, and accurately reports findings (in a quantitative, semi quantitative or qualitative manner) to be used by qualified healthcare professionals in accordance with professional guidelines.

Biomarker – an individual alteration (e.g. KRAS G12C), chromosomal abnormality (e.g. chromosome 21 trisomy) or a specific analyte (e.g. circulating tumour DNA) with established utility for a given clinical indication".

Germline variant – a genetic alteration that occurs within the germ cells that can be passed to subsequent generations.

Somatic variant – a genetic alteration that occurs in any somatic cell (non-germ cell) and is not passed through the generations".

Such assays are likely to be considered Class C IVDs under Rule 3 of Annex VIII (genetic testing).

3) How should the intended purpose of the NGS-based assay be defined?

For general guidance on intended purpose, see chapter 1 of this eBook.

The manufacturer must identify and specify the disease or other condition of interest, define the clinical function of the test, and define the population that the test is intended to be used for. Clinical functions of NGS-based assays are broad and may include disease screening, diagnosis aids, prognosis, determination of pathological state, disease monitoring, or response to treatment. If additional clinical evidence is generated over the lifecycle of an IVD assay, the intended purpose must be revised accordingly.

In addition to standard critical components of a general intended purpose, NGS-based assay intended purpose statements must include a list of genes (these can be in table format within an appendix) where applicable, as well as interrogated regions and types of genetic alterations included in the IVD (this has been demonstrated by Table 19 in the below examples).

The intended purpose must specify what sample types are expected as input (e.g. DNA, RNA, DNA/RNA extracted from FFPE tumour tissue or fresh tumour tissue, whole blood and its derivatives), the nature of the test (quantitative, semi-quantitative or qualitative) and what it intends to report after secondary and tertiary analyses are performed.

It is the responsibility of the manufacturer to mitigate foreseeable misuse of the device, for example through inclusion of limitations (e.g. limited coverage of certain regions, lack of clinical evidence for certain alterations detected but not reported) and critical warnings.

Table 19 below shows examples of intended purpose.

Principle	Intended purpose	Target population	Type of test	Measurand	Quantitative /qualitative	Specimen type	Nucleic acids	Number of genes, list of genes
Pathological state	Detect genetic alterations (SNV, Indels, Rearrangements, CNAs) and signatures (MSI, TMB, HRD) across 600 genes in solid malignant neoplasms to establish disease stage	Patients with prior diagnosis of solid tumours	NGS, Targeted hybrid capture	SNV, indels, CNAs, MSI, TMB, HRD across 300 genes and rearrangements (RNA only) across 50 genes,()	Qualitative	FFPET	DNA and RNA	300, Table X
Diagnosis	Aid in diagnosis of lysosomal storage disorders	Patients with signs and symptoms of mucopolysaccharidosis/ mucolipidoses/ oligosaccharidoses	NGS, Targeted	SNP, insertions and deletions <15bp, exon-level deletions and duplications (germline) in specified genes	Qualitative	Buccal swab, saliva, peripheral blood	DNA	65, Table X
Prognosis	Assess a risk to develop distant metastasis 7-10 years after diagnosis of breast cancer	Individuals with established breast cancer diagnosis of stage I and II, with tumour size ≤5.0 cm and negative lymph nodes	RNA-seq, gene expression analysis	RNA transcripts for specified 180 genes	Quantitative	FFPET	RNA	180, Table X
Monitoring/ MRD	Assess minimal residual disease (MRD) in colorectal cancer (stage II, III) patients after curative intent surgery	CRC (stage II, III) patients undergoing surgical resection	High throughput hybrid capture NGS	Detection of SNVs and Indels across 500 genes measured units (e.g mutant molecules per mL (MMPM))	Semi-quantitative	Blood	Cell free DNA	500, Table X
Screening	Prenatal screening for trisomies in chromosomes 13, 18, and 21	Pregnant individuals	Whole genome sequencing	Chromosomal copy numbers	Semi-quantitative	Maternal blood	Cell free DNA	N/A (whole chromosome)

Table 19. An example of the list of genes that could be included for an NGS-based assay

4) Can the NGS-based assay generate data on genes which are not included within the intended purpose?

NGS-based assays often cover a large number of genes and associated alterations. For any results to be reported for clinical use, those alterations or signatures need to be validated as part of the performance evaluation and included in the intended purpose.

For example, an NGS-based liquid biopsy assay that measures circulating tumour DNA in peripheral blood to assess the presence or absence of minimal residual disease, may detect all classes of genetic alterations across 500 genes. However, the manufacturer may assess only those alterations (e.g. SNV, indels, CNA) that are validated for minimal residual disease assessment in a particular indication (e.g. colorectal cancer stage II, III).

Data generated beyond validated genes and associated alterations must not be included in the final static test results report (TRR).

Similarly, a Whole Genome Sequencing assay used for Non-Invasive Prenatal Testing could detect many genetic alterations and chromosomal abnormalities. However, the manufacturer should restrict the clinical

report to the specific alterations associated with specific medical intended purposes that have been validated (e.g. trisomy of chromosome 13, 18 and 21).

This should be made clear by the manufacturer in the instructions for use and should be appropriately risk assessed.

5) How should I generate clinical evidence for an NGS-based IVD assay?

Clinical evidence is composed of three main pillars: analytical performance, scientific validity and clinical performance. More information on clinical evidence as a concept can be found in chapter 2 of this book.

Clinical evidence for a given assay must support the intended purpose, which can be further expanded in line with appropriate evidence generation.

In scenarios where manufacturers can justify clinical evidence without clinical performance studies, they must provide an appropriate justification why they are not required and instead provide clinical performance data via scientific peer reviewed literature and/or published experience of routine diagnostic testing.

6) How should I define scientific validity for the NGS-based assay?

Scientific validity as a concept is explained in chapter 4 of this book.

A manufacturer must document the literature search methodology, literature search protocol and generate a literature search report. These should be included within the performance evaluation report.

Defining literature search methodology helps to assess the nature and extent of scientific validity. This process requires clear inclusion and exclusion criteria for literature relevance evaluation. Furthermore, defining screening and selection criteria help with identification of adequate articles. Manufacturers must ensure that data utilised from these sources are appropriate.

For NGS-based assays (WGS, WES, targeted CGP panels), reported gene alterations (e.g. SNV in EGFR gene, Deletions in BRCA genes) and complex biomarkers (e.g. TMB, MRD, HRD signatures) supporting the intended purpose must be reviewed and cited in its scientific validity report. Manufacturers of NGS-based assays may utilise online databases such as MEDLINE®, BIOSIS Previews®, and Embase®, among potentially others. These databases can be accessed via search engines such as, but not limited to, PubMed and Scopus. For variant curation, several public databases are available, such as COSMIC, TCGA, ExAC, ClinVar, and dbNSFP, and there may be others. MedTech Europe does not endorse any specific solution mentioned in this document. The reader is encouraged to perform their own commercial research.

If this approach is used and the benefit/risk analysis is positive, Post-Market Performance Follow-up (PMPF) is necessary to address potential risks associated with reporting genetic alterations and biomarkers of uncertain scientific validity, ensuring patient safety.

As well as these sources, scientific validity can be generated by the manufacturer themselves through clinical performance studies or their own published literature. Examples are shown in Table 20.

Principle	Intended Use	Scientific Validity
Pathological state	Detect genetic alterations (SNV, Indels, Rearrangements, CNAs) and signatures (MSI, TMB, HRD) across 600 genes in solid malignant neoplasms to establish disease stage	Literature review for evidence demonstrating genetic alterations in specified 600 genes associated with tumourigenesis, disease state and severity, sensitivity or resistance to treatment
Diagnosis	Aid in diagnosis of lysosomal storage disorders	Scientific validity can be supported by the evidence demonstrating association between germline alterations in specified 65 genes and phenotypic manifestation of various subtypes of mucopolysaccharidosis, mucolipidoses and oligosaccharidoses
Prognosis	Assess a risk to develop distant metastasis 7-10 years after diagnosis of breast cancer in individuals (stage I and II) with tumour size ≤5.0 cm and negative lymph nodes by measuring gene expression across 180 genes	Scientific validity can be supported by the evidence demonstrating association between gene expression across specified 180 genes and distant breast cancer metastases - such evidence could be based on literature or generated by the manufacturer
Monitoring for treatment escalation/de-escalation	Assess ctDNA based minimal residual disease (MRD) in colorectal cancer (stage II, III) patients after curative intent surgery for treatment escalation or de-escalation	Scientific validity can be supported by the evidence on circulating tumour DNA in peripheral blood as a surrogate marker for MRD. Furthermore, intended use must be supported by the evidence on association between MRD positivity and treatment escalation status post surgery in stage II, III patients. This could be based on the combination of literature as well as additional evidence generated by the manufacturer.
Screening	Detect trisomies in chromosomes 13, 18 and 21 in peripheral blood of pregnant individuals	Scientific validity can be supported by the literature based evidence on association between chromosomal abnormalities observed in cfDNA of pregnant individuals and trisomies 13, 18 and 21

Table 20. Scientific validity generated from manufacturer clinical performance studies or their own published literature

7) How should I define analytical performance for the NGS-based IVD assays?

During the design phase, the manufacturer must identify a specific sequencing approach and the regions covered (e.g. whether all exons of a particular gene are covered or a subset) that are most suitable for the intended purpose, and determine exonic and intronic coverages across the genomic regions. They also must determine the alteration types that are relevant for a specific indication or indications. Such genetic alterations, chromosomal abnormalities, and signatures include but are not limited to:

- single nucleotide variants

- insertions and deletions
- copy number alterations (amplifications or losses)
- structural variants (inversions, translocations, conversions)
- short tandem repeats
- microsatellite instability
- tumour mutational burden
- genomic loss of heterozygosity
- homologous recombination deficiency

Prior to analytical verification, feasibility studies must be performed on appropriate samples to ensure the assay can meet predefined acceptance criteria. During the feasibility assessment, manufacturers must set parameters and acceptance criteria on quality control thresholds, sample pass rates, positive and negative call rates, flag any excluded regions (genomic coordinates) of the gene where making variant calls is not possible (e.g. homopolymer, heteropolymer, high/low GC content, repeat sequences).

Analytical performance as a concept is explained in chapters 2 and 6 of this eBook.

Analytical verification is the characterisation and assessment of analytical assay performance characteristics and aims to demonstrate whether the test performance meets predefined performance specifications, and – in the case of the NGS assay – successfully identifies and measures the presence or absence of a genetic alteration and chromosomal abnormalities that are relevant for a specific indication. This typically involves studies to assess analytical sensitivity, analytical specificity, trueness (bias), precision (repeatability and reproducibility), accuracy (resulting from trueness and precision), limits of detection and quantitation, measurement range, cut-off, including determination of appropriate criteria for specimen collection and handling and control of known relevant endogenous and exogenous interference, cross-contamination, reagent stability, specimen stability, and guardbanding.

Organisations conducting analytical performance studies need to determine the appropriate methods and acceptance criteria, working within the appropriate regulatory requirements. It is for the manufacturer to specify the extent of the data required and ensure the analytical performance is sufficient.

In relation to analytical validation, the following should be considered:

- Genomic alterations can be analytically validated either themselves or by representative approach when there is justifiable rationale. For example, if an NGS-based assay covers hundreds of SNVs, not all SNVs need to be represented in analytical validation studies. These can be binned based on class of alteration provided sufficient coverage is achieved across the regions of interest to detect variants at the desired variant allele frequency (sensitivity). For example, SNVs resulting in KRAS G12C and PIK3CA E545K alterations may be grouped under the same SNV group/bin. Similarly, indels can be grouped based on the insertion and deletion sizes:
    - Once the variants are analysed based on classes of alterations, binning can also be considered based on variant type in terms of whether variants are reported in public databases (e.g. COSMIC) as known or likely pathogenic, as well as based on levels of evidence (LOE) per European Society for Medical Oncology (ESMO) Scale for Clinical

Actionability of molecular Targets (ESCAT) or Association for Molecular Pathology (AMP)/American Society of Clinical Oncology (/ASCO)/College of American Pathologists (CAP) guidelines. Such LOE-based analyses are especially meaningful when conducting reproducibility, limit of detection (LOD) and accuracy studies. The bin level analysis allows grouping of variants based on the commonality of whether it is an alteration class, subclass or LOE

- For precision and reproducibility studies:
    - Maximisation of variant representation within each individual unique sample is beneficial. For example, using a unique breast cancer sample that harbours concurrent alterations (e.g. CDx biomarkers PIK3CA mutations and ERBB2 amplification)
    - If running multiple samples representing the same indication but harbouring different alterations, each of them can be used as wild type (negative control) for the others. Follow CLSI guidelines for information on number on reproducibility study guidance in terms of number of days, reagent lots, operator or consult with the local regulatory requirements
    - To supplement the LOD study sample set, select samples that have close to LOD

- For samples in analytical validation studies:
    - Clinical samples should be considered whenever possible
    - Contrived samples and cell lines can be used for rare variants (if applicable) – If contrived samples/cell lines are used, performance equivalence must be established
    - Representative set of variants must be selected to maximise the variation across all target sequences

Specimen requirements (Annex II, Section 6.1.1) - Although multiple specimen and collection types may be appropriate for any given NGS-based assay, each type should be validated for use in producing nucleic acids of the appropriate quality and quantity required for overall test performance. Minimum nucleic acid content for input samples must be established and verified, and subsequently described in the acceptance criteria in instructions for use. Examples of specimen types include, but are not limited to:

- formalin-fixed paraffin-embedded (FFPE) tissue
- fresh frozen tissue
- fine needle aspirate (liquid form)
- whole blood or other peripheral blood derivatives (e.g. plasma, peripheral blood mononuclear cells)
- saliva
- swabs (e.g. buccal, vaginal)
- biological fluids (e.g. cerebrospinal fluid, ascitic fluid, pleural fluid, bronchoalveolar lavage)

Table 21 below includes analytical validation examples.

Principle	Intended Use	Analytical Performance
Pathological state	Detect genetic alterations (SNV, Indels, Rearrangements, CNAs)	Qualitative detection of somatic alterations (SNV, Indels, rearrangements, CNAs) and

	and signatures (MSI, TMB, HRD) across 600 genes in solid malignant neoplasms to establish disease stage	signatures (MSI, TMB, HRD) across select regions of 600 genes using DNA and RNA extracted from FFPE tumour tissue specimens from patients with prior diagnosis of solid tumours with appropriate analytical sensitivity, specificity, precision etc.
Diagnosis	Aid in diagnosis of lysosomal storage disorders	Qualitative detection of germline alterations (SNP, insertions and deletions <15bp, exon-level deletions and duplications) across select regions of specified 65 genes using genomic DNA isolated from buccal swab, saliva or peripheral blood mononuclear cells obtained from patient with established diagnosis of lysosomal storage disorders yielding appropriate analytical sensitivity, specificity, precision etc.
Prognosis	Assess a risk to develop distant metastasis 7-10 years after diagnosis of breast cancer in individuals (stage I and II) with tumour size ≤5.0 cm and negative lymph nodes, by measuring gene expression across 180 genes	Quantitative detection of RNA transcripts in specified 180 genes by RNA sequencing using RNA extracted from FFPE tumour tissue specimens with appropriate analytical sensitivity, specificity, precision etc.
Monitoring for treatment escalation/de-escalation	Assess ctDNA based minimal residual disease (MRD) in colorectal cancer (stage II, III) patients after curative intent surgery for treatment escalation or de-escalation	Quantitative measurement of ctDNA levels in peripheral blood samples from CRC patients with appropriate analytical sensitivity, specificity, precision etc.
Screening	Detect foetal trisomies in chromosomes 13, 18 and 21 in peripheral blood of pregnant individuals	Quantitative measurement of chromosomal copy numbers in peripheral blood from pregnant individuals with a sufficient foetal fraction to suggest a foetal aneuploidy

Table 21. Analytical performance of NGS-based assays

8) How should I define clinical performance for an NGS-based IVD assay?

Clinical performance as a concept is explained in chapters 2 and 6 of this book.

The recommendations within these sections are applicable for IVDs, with very little difference specifically for NGS-based assays. Examples are shown in Table 22 below.

Principle	Intended Use	Clinical Performance
Pathological state	Detect genetic alterations (SNV, Indels, Rearrangements, CNAs) and signatures (MSI, TMB, HRD) across 600 genes in solid malignant neoplasms to establish disease stage	Clinical performance must be established to derive diagnostic performance metrics (e.g. diagnostic sensitivity and specificity through method comparison study) using samples from the Intended Purpose population. In scenarios when biomarkers are novel, literature review in combination with analytical performance may be sufficient. PMPF studies should be planned to mitigate the risk of patient harm arising from delivering patient reports including non-clinically validated genetic alterations.

Diagnosis	Aid in diagnosis of lysosomal storage disorders	Clinical performance can be demonstrated by assessing diagnostic sensitivity and specificity using appropriate samples from the intended purpose population
Prognosis	Assess a risk to develop distant metastasis 7-10 years after diagnosis of breast cancer in individuals (stage I and II) with tumour size ≤5.0 cm and negative lymph nodes, by measuring gene expression across 180 genes	Clinical performance can be demonstrated by establishing association between gene expression patterns and occurrence of distant metastases (Hazard Ratio). Furthermore, association between recurrence free survival at 7-10 years and gene expression profile for low versus high risk must be established (HR, KP).
Monitoring/MRD	Assess ctDNA based minimal residual disease (MRD) in colorectal cancer (stage II, III) patients after curative intent surgery for treatment escalation	Clinical performance can be generated by first establishing association between postsurgical MRD positivity and disease recurrence (Hazard Ratio). Furthermore, benefits derived by escalating treatment in patients with MRD positivity status post surgery must be assessed (HR, KP).
Non-invasive Prenatal Testing (NIPT) Diagnostic	Prenatal screening for trisomies in chromosomes 13, 18 and 21	Assess diagnostic sensitivity and specificity of the investigational device predicting phenotypic expression of a chromosomal aneuploidy from tested maternal blood and newborn clinical diagnosis of an aneuploid-associated condition

Table 22. Clinical performance for an NGS-based assay

9) Test result reporting

NGS-based assays can generate a number of output files (e.g. BAM, JSON, VCF) that contain sample and variant level information. However, manufacturers are obligated to generate a test results report (TRR) that is considered as a final static report and lists only those detected biomarkers that have been validated for a given assay as per IVD requirements.

Although intermediate sequencing files contain many pieces of quality control related information, it is desirable that test manufacturers list sample specific critical quality control metrics (e.g. coverage) in the TRR. In addition, the TRR must disclose limitations of a test (e.g. test has a low threshold for contamination leading to failed status of a sample). If applicable, the TRR may also host the list of genes and associated regions covered.

For somatic oncology, specifically NGS-based assays that are intended for pathological state determination of solid or haematological malignancies, manufacturers may consider applying levels of evidence rules that are established by professional societies e.g. ESMO/ESCAT, AMP/ASCO.

Table 23 below shows examples of test result reporting.

Principle	Intended Use	Test Results Report
Pathological state	Detect genetic alterations (SNV, Indels,	TRR must report only those genetic alterations and signatures clinical evidence of which have

	Rearrangements, CNAs) and signatures (MSI, TMB, HRD) across 600 genes in solid malignant neoplasms to establish disease stage	been established by analytical, clinical performance studies and scientific validity. Detected alterations and signatures can be tiered based on levels of evidence rules as per ESMO/ESCAT or AMP/ASCO or levels of evidence rules that correlate with tiering system by professional societies.
Diagnosis	Aid in diagnosis of lysosomal storage disorders	For this specific example, TRR must report only those genes and associated germline alterations, which were analytically and clinically validated with specific types of lysosomal storage disorders.
Prognosis	Assess a risk to develop distant metastasis 7-10 years after diagnosis of breast cancer in individuals (stage I and II) with tumour size ≤5.0 cm and negative lymph nodes, by measuring gene expression across 180 genes	To support this specific example of prognostic claim, the TRR may include a 7-10 years recurrence risk score for developing distant metastases
Monitoring/MRD	Assess ctDNA based minimal residual disease (MRD) in colorectal cancer (stage II, III) patients after curative intent surgery for treatment escalation	This particular example is based on a semi-quantitative assay for MRD. Hence TRR may report MRD in MMPM as noted in the intended purpose.
Non-invasive Prenatal Testing (NIPT) Diagnostic	Prenatal screening for trisomies in chromosomes 13, 18 and 21	For this example, TRR must report specific trisomies as specified in the intended use

Table 23. Example of test result reporting

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 15 – Real World Evidence

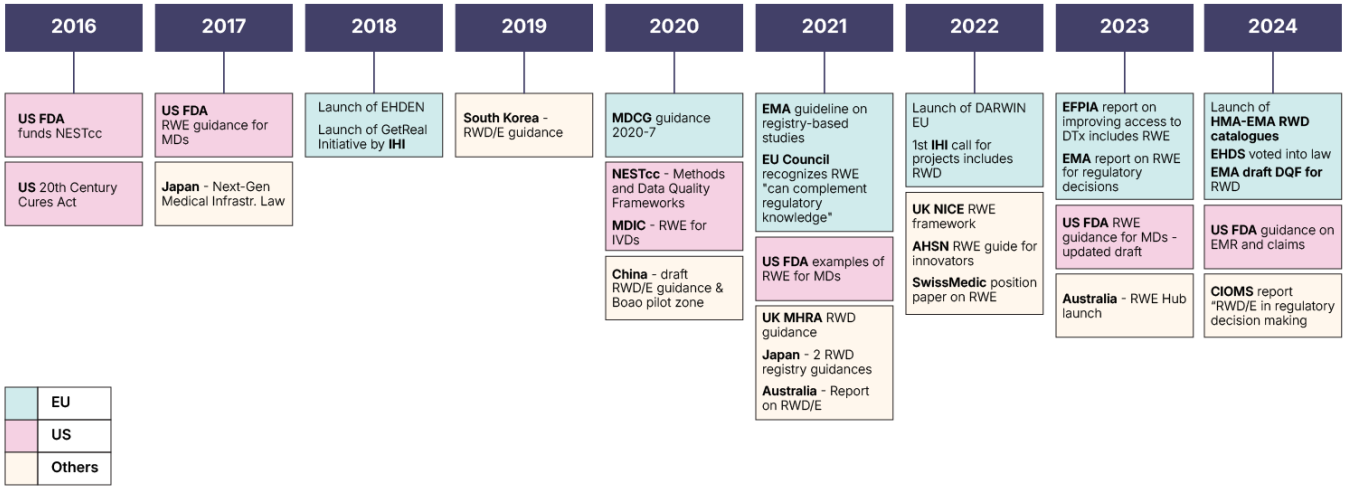
While clinical studies are the established gold standard for generating clinical evidence for in vitro diagnostics, another type of evidence is slowly becoming more prominent – real-world evidence (RWE), which is generated from real-world data (RWD). Put simply, RWD is healthcare data derived from sources other than traditional controlled studies or clinical trials. Particularly in the pharmaceutical industry, RWD has been used to support product registration and post-market surveillance for quite some time. The majority of progress in using RWD and RWE to support regulatory submissions for IVDs has been driven by the US Food and Drug Administration (FDA). This is primarily due to the much easier access to patient data for research purposes (enabled by less restrictive data protection regulations like the Health Insurance Portability and Accountability Act (HIPAA)[1] and a higher degree of healthcare data digitisation (driven at least in part by the Health Information Technology for Economic and Clinical Health Act (HITECH) Act of 2009[2], which aimed to incentivise the use of electronic health records across the US.

Within the last decade the EU[3], the US FDA[4] and multiple Asia-Pacific (APAC) authorities[5,6,7] have been launching initiatives and releasing guidance on various aspects of RWE generation across the pharmaceutical, MD, and IVD industries, as well as sharing examples of successful use cases[8,9] . For a list of some relevant publications, see Figure 14.

Additionally, there appears to be an increasing interest and motivation to promote the secondary use of healthcare data for both research and regulatory purposes, as evidenced by initiatives like the European Health Data Space (EHDS)[10], Data Analysis and Real World Interrogation Network (DARWIN EU)[11], or the FDA's active surveillance system Sentinel Initiative[12,13].

The key challenge in utilising RWE to meet clinical evidence requirements for IVDs, particularly under regulations like the IVDR, is that most available guidance and frameworks are pharmaceutical industry-oriented, and while somewhat more experience has accumulated in the medical device space thanks to the use of device registries for safety monitoring, the use of RWE for IVDs is still in its early phase. Consequently, many of the examples in this chapter are related to medications or medical devices; while these medical products are quite distinct from IVDs, the general RWD-related concepts are very similar and remain broadly applicable.

For an even more in-depth overview of this field, the Duke-Margolis International Harmonization of Real World Evidence Standards Dashboard is useful[14].



AHSN, Academic Health Science Network; CIOMS, Council for International Organizations of Medical Sciences; DQF, data quality framework; DTx, digital therapeutics; EFPI, European Federation of Pharmaceutical Industries and Associations; EHOEN, European Health Data Évidence Network; EHDS, European Health Data Space; 1H, Innovative Health; Initiative:MDOG, Medical Device Coordination Group; MHRA, Medicines and Healthcare products Regulatory Agency (UK); NESTcc, National Evaluation System for health Technology Coordinating Center (US); NICE, The National Institute for Health and Care Excelence (UK).

Figure 14: Global regulatory guideline development for RWD/RWE

1) What is the definition of real-world data?

While regulatory agencies in Europe and around the world have their own RWD definitions, which vary slightly, they are generally consistent. These definitions share the common concept of RWD being collected *during routine care* administration, and *outside* of a traditional controlled environment of a clinical trial/study. Some examples are below.

European Medicines Agency: "Routinely collected data relating to patient health status or the delivery of health care from a variety of sources other than traditional clinical trials." "RWD can be used to achieve better informed and more efficient regulatory decision-making as a complement to existing evidence."[15]

UK National Institute for Health and Care Excellence (NICE): "Data collected outside the context of a highly controlled clinical trial. Real-world data can be routinely collected during the delivery of health or social care. It can also be collected prospectively, to address one or more specific research questions."[16,17]

Swissmedic: "All data other than those collected through a clinical trial."[18]

US FDA: "Data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources".[19]

Beyond these definitional nuances, agencies can also differ significantly in their experience with and acceptance of RWD and RWE. Navigating this still-evolving landscape requires ongoing dialogue, experience sharing, and concerted efforts towards establishing common frameworks and standards for real-world data and real-world evidence across the entire IVD industry. Collaboration and knowledge sharing among the industry, scientific, and regulatory community can help drive the development of clear and harmonised definitions, guidelines, and best practices for the use of RWD and RWE in the IVD space.

2) What are the most common sources of RWD?

Established RWD sources

Electronic health/medical records (EHR/EMR)

A collection of electronic health data related to a patient, utilised for healthcare purposes[20]. A typical EHR/EMR includes a patient's medical history in form of diagnoses, treatments (e.g. medications or procedures), preventive care (e.g. screening or immunisation), and laboratory and other test results. This data is usually stored in tabular format, but may also contain images (e.g. radiology) or unstructured physician notes[21].

Medical claims or administrative data

Information shared by care providers with insurers to receive reimbursement; typically includes details on care appointments, insurance enrolment, and costs. It can come from individual hospitals or healthcare networks, health insurance providers, clearinghouses, pharmacy software, or other sources[22].

Patient, product, or device registries

Organised health information systems using observational methods to collect clinical and other related data in a standardised format, focusing on a population defined by a particular disease, condition, device, or exposure[23]. They may be established for safety monitoring or research purposes, or be completely or partially related to clinical studies[24]

Lab and instrument data

Data obtained directly from laboratory instruments and laboratory information systems, consisting of test results as well as quality and instrument performance data. It is frequently integrated with EMR data[25].

Spontaneous reporting systems

Passive surveillance method where healthcare providers, patients, manufacturers, or other stakeholders voluntarily report adverse events or issues related to medications, vaccines, or medical devices. Examples include the EU EudraVigilance[26], UK Yellow Card scheme[27], field safety corrective actions (FSCAs) for medical devices such as SwissMedic's[28], or the US FDA's Adverse Event Reporting System (FAERS), Vaccine Adverse Event Reporting System (VAERS)[29], and Manufacturer and User Facility Device Experience (MAUDE) for medical devices[30].

Emerging RWD sources

Genetic and genomic data

Data consisting of biomarker, gene, or genome testing results of an individual, derived from the analysis of the genetic material[31].

Patient-generated health data (biosensor data)

Data generated by implantable or external physiological monitoring medical devices, consumer wearables, or digital health devices (e.g. health and wellness apps), gathered directly from/by the patient outside of clinical care settings[32].

Patient-reported outcomes (PROs)

Outcomes reported directly by the patient, pertaining to their symptoms, treatment, health status, and quality of life[33]. They may be collected as part of a clinical study, during routine care, or even directly by the patient, either on paper or electronically (ePROs). Patient surveys can be another similar source of RWD.

Public health surveillance data

Data collected with the purpose of public health surveillance, e.g., of COVID-19 or reportable diseases like tuberculosis[34], used as a form of passive surveillance.

Clinically annotated biobanks

Repositories of biological specimen samples, accompanied by clinical information about the subjects. Biobanks (and the corresponding data) could in some cases be connected to a traditional clinical study, but may be considered an RWE source if they are embedded in routine clinical practice. A great example is the UK Biobank[35].

Medical imaging and other device data repositories

Repositories of clinical imaging and other visual or trace data (like ECG, CT scans, x-rays, etc.), which can also include clinical annotation[36].

Note: **Published literature** (including meta-analyses) *may* be considered a source of RWE if RWD was used to generate the clinical evidence presented in the publication. Additionally, literature searches can be helpful in identification of potential new RWD sources.

3) What is the definition of real-world evidence?

Multiple definitions of RWE also coexist across different jurisdictions. They are, however, consistent in stating that it comes from the analysis of RWD.

European Medicines Agency: "Evidence derived from the analysis of RWD."[37]

UK NICE: "Evidence generated from the analysis of real-world data. This includes studies using real-world data to form an external control to a clinical trial."[38]

Swissmedic: "Information derived from analyses of RWD."[39]

US FDA: "Clinical evidence about the usage and potential benefits and risks of a medical product derived from analysis of RWD."[40]

4) Does the IVDR mention RWE?

The IVDR does not explicitly mention RWE, but it refers to "published experience gained by routine diagnostic testing" in the context of demonstrating clinical performance in Annex XIII Part A[41] - which can reasonably be interpreted to include RWE. The IVDR also imposes the obligation to systematically and proactively collect and review device-generated data through PMS to monitor device quality, performance, and safety throughout its lifespan. This requires the use of data and evidence collected during regular device use in the real world – which fits in very well with the definition of RWD/RWE.

5) What does the supporting EU guidance say about RWD/RWE?

The existing EU guidance sometimes specifically refers to RWD and RWE, though mostly in the MDR context. Below are some examples:

Although **MEDDEV 2.12/2 rev 2**[42] predates the EU MDR and the IVDR, it provides valuable considerations on study design, methodologies, and analysis of data obtained after the device is placed on the market to generate clinical evidence. The guidance underscores the importance of gathering real-world post-market clinical data to ensure long-term device safety and performance, as well as the acceptability of residual risks of devices in a real-world clinical setting. It emphasises investigating device issues in a larger and more varied patient population, arguably encouraging the use of RWD. One of the Post Market Clinical Follow-Up (PMCF) study methods proposed in the guidance is reviewing "relevant retrospective data from patients previously exposed to the device" in a variety of settings, as opposed to a narrow group of controlled study participants.

MDCG 2020-6[43] focuses on sufficient clinical evidence for legacy medical devices under the EU MDR. It proposes a hierarchy of evidence, and names "Outcomes from high quality clinical data collection systems such as registries" as acceptable to support market access applications for legacy devices.

MDCG 2020-7 Section C[44] specifically names device registries, as well as “planned” RWD analyses as one of the options for PMCF under the EU MDR. It emphasises the need for high-quality and reliable RWD sources to ensure meaningful results.

MDCG 2022-2[45] uses identical phrasing to the IVDR itself ("data from published experience gained by routine diagnostic testing") when discussing both analytical and clinical performance of IVDs, and it also makes a reference to "other sources of clinical performance data" – both of which could be interpreted to include RWD. The most specific reference can be found in Section 6.5 on analytical performance of IVD medical device software: its generalisability should be demonstrated with “real-life datasets”.

6) How can RWD and RWE be utilised throughout the IVD product lifecycle?

There are many potential ways in which RWE can be utilised during the lifecycle of an IVD. See Figure 15 below for a high-level overview.

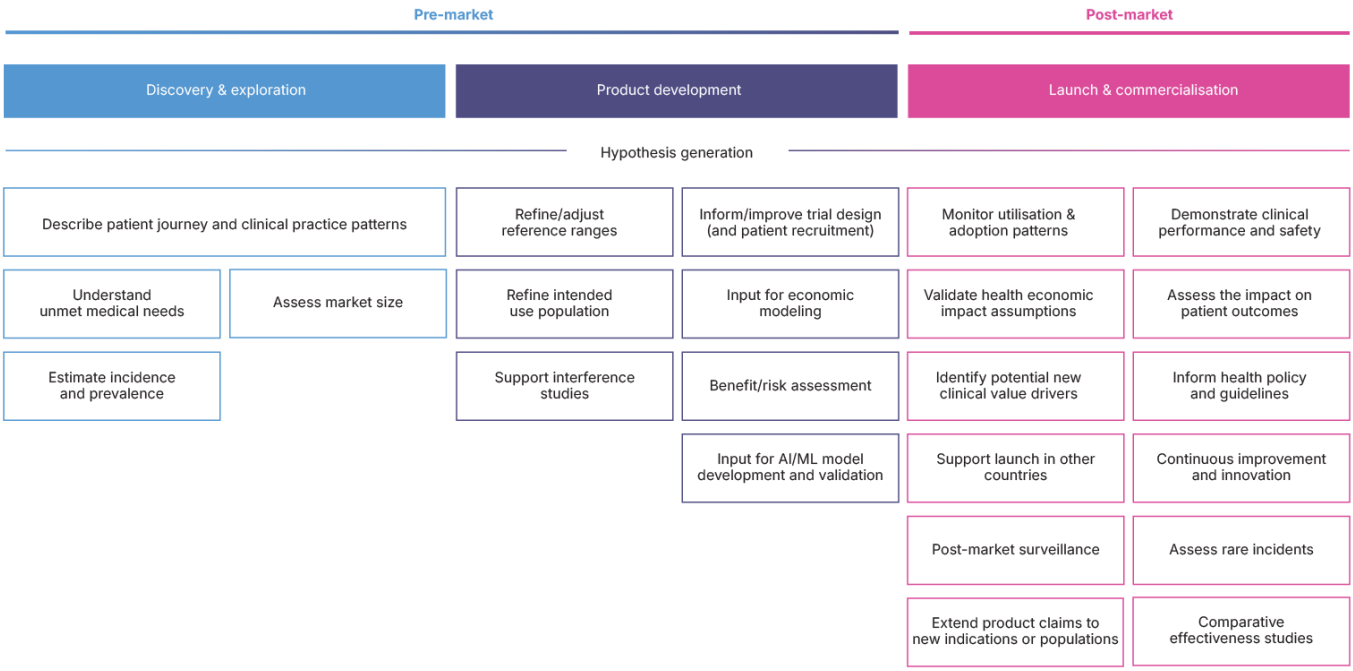


Figure 15: RWD/RWE opportunities throughout IVD product lifecycle

Note: Specific use case examples are highlighted elsewhere in this chapter.

7) How can RWD/RWE be used for IVDs in the pre-market phase?

Throughout the entire IVD lifecycle

Hypothesis generation: RWD analysis can help generate novel questions and hypotheses to be tested in clinical studies – both traditional and real-world data studies.

During early discovery and exploration

Describing patient journey and clinical practice patterns: RWD can be a useful source of information for designing effective clinical strategy by describing the current standard of care, which includes the patient journey, clinical practice patterns, and diagnostic pathways within different clinical settings, both at the individual patient level and in aggregate. It can also be used to identify patients who would benefit most from a given diagnostic test or intervention.

Understanding unmet **medical needs:** Through RWD analysis, gaps in current care pathways and unmet patient needs can be identified and quantified; understanding them is an essential pre-market activity to ensure that devices that are brought to market fulfil a specific need and have a clear placement.

Assessing market size: RWD can be used to estimate the potential patient population size, testing frequency, product utilisation rates, geographic distribution, and other market-wide trends and patterns. This type of analysis can also continue as a post-market activity.

Estimating incidence and prevalence: RWD can be used to calculate incidence and prevalence of conditions of interest in different real-world populations.

During product development

Refining/adjusting reference ranges: There are multiple established indirect methods for using retrospective RWD to calculate reference intervals. These approaches require a larger sample size than conventional direct methods, but usually don't need in-depth clinical information, making it a viable alternative in many cases. Some examples are the Bhattacharya method, the Hoffman method, kosmic, and RefineR[46,47].

Informing/improving clinical study design: RWD can be used to inform IVD clinical study design and conduct them in a variety of ways. It can be used to validate dependent variables or study endpoints, to ensure they accurately reflect clinical practice. RWD can also support patient/site selection and recruitment, as well as diversity and inclusion planning. Further, RWD can be used to support the validity, interpretability, or generalisability of clinical study results, even using linked patient medical records to supplement or complement them. Finally, RWD and RWE infrastructure can serve as a framework for data collection for both pre- and post-market clinical studies.

Refining intended use population: RWD can help identify patient subgroups where a device is most (or least) effective, or where its use may be associated with specific risks. It can also provide information on

the device's performance in diverse patient populations, including vulnerable groups, those with comorbidities, or other factors that may affect its accuracy or reliability.

Providing input for economic modelling: RWD can be used to generate insights and evidence to support and enhance health economic models, such as estimating healthcare utilisation and costs, describing patient demographics, validating assumptions, etc.

Supporting interference studies: RWD can be used to generate hypotheses or evidence to help determine factors impacting the accuracy or reliability of diagnostic tests, for example, identifying potentially interfering endogenous or exogenous substances, or any patient-specific factors.

Benefit/risk assessment: RWD can help generate evidence of an IVD's real-world effectiveness and impact on patient care, both positive and negative.

Providing input for AI/ML models: RWD can be used as input data for development, training, and validation of medical device software such as diagnostic AI/ML models or predictive algorithms (which can be classified as medical device software under either the IVDR or the MDR).

8) How can RWD/RWE be used for IVDs in the post-market phase?

Monitoring utilisation and adoption patterns: After product launch, RWD can show the patterns of its adoption and utilisation in everyday clinical care. This information can subsequently guide marketing strategies, allowing broader patient access.

Demonstrating clinical performance and safety: RWD enables the evaluation of IVD clinical performance and safety by capturing real-world device use after launch, allowing for the assessment of diagnostic accuracy, clinical utility, and the identification of potential adverse events or performance issues in diverse patient populations and settings. Longitudinal RWD can enable the monitoring of potential safety and effectiveness issues which are unforeseen during early adoption or may only appear after long-term use.

Validating health economic impact assumptions: RWD can be used to derive patient outcomes, resource utilisation, and healthcare costs to assess the real-world cost-effectiveness of IVDs after placement on the market.

Assessing the impact on patient outcomes: RWD can be used to assess patient outcomes, such as clinical outcomes, biomarker values, patient-reported outcomes including quality of life, behaviour change, user satisfaction and engagement. RWD may be used to estimate effects on final outcomes of interest rather than on surrogate outcomes, and may also correlate these outcomes, as well as measure effects over an extended period.

Identifying potential new clinical value drivers: RWD can potentially help uncover patterns, correlations, and trends pointing to previously unrecognised clinical factors or outcomes, especially if large-scale data is available.

Informing health policy and guidelines: RWD can be used to assess the large-scale impact of interventions, including IVDs, on service delivery and care decisions. Through RWD, the utility of medical interventions in broad, general patient populations in routine care settings can be investigated. Health and economic outcomes can be used to aid in policy decision-making and guideline development.

Supporting launch in other countries: As described in Chapter 13 of this eBook, the IVDR permits the use of clinical data collected outside of the EU; this could also apply to real-world data. RWD coming from different jurisdictions, if shown to be representative of the intended use population, can potentially be used to generate evidence for submissions to multiple authorities.

Product improvement and innovation: Throughout the product lifecycle, insights from RWD can be used to drive continuous improvement and innovation. By analysing real-world performance data, manufacturers can identify areas for product enhancement, optimise testing algorithms, and develop new applications or indications based on real-world evidence.

Post-market surveillance: RWD can be used to monitor real-world device performance and to detect signals suggesting quality or safety issues. RWD can also be used to determine and refine appropriate corrective actions. Additionally, analysis of RWD collected to fulfil PMS obligations can generate RWE of product performance in real-world care settings.

Assessing rare incidents: The usually large sample size of RWD can enable detection of rare adverse events and safety incidents after the IVD enters the market; this can also be achieved through combining multiple RWD sources, e.g., across different countries.

Extending product claims: RWD can be analysed to evaluate product performance across various populations and settings, to provide evidence of effectiveness and safety in new intended populations and apply for an expanded label.

Comparative effectiveness studies: RWD can be used to compare the benefits, risks, and effectiveness of different diagnostic tests already on the market.

9) What are the limitations of RWD/RWE?

RWE offers significant potential to support IVD registration and fulfilling of post-market regulatory obligations. However, it is important to acknowledge certain limitations associated with its use.

The use of RWE in the IVD space is still relatively new across all geographies. This limited experience may lead to hesitation among authorities to consider applications containing RWE; at the same time, the recent

efforts to highlight the potential of RWE (e.g. the EMA report[48] or the EHDS), and the proliferation of guidelines and frameworks on the responsible use of RWD to generate clinical evidence for IVDs, suggest an ongoing interest and willingness to engage on the topic.

The most important limitation of RWD can be its provenance – it is usually not specifically collected for research purposes, but simply represents how healthcare is delivered and recorded in real-world conditions. As a result, it may contain inherent biases and confounders that need to be carefully identified and addressed[49]. Appropriate analytical and study design methodologies are necessary to mitigate these biases, in the absence of randomisation. Fortunately, many useful resources on RWD study design already exist and are referenced in this chapter.

RWD can also face reliability and accuracy issues. Incomplete documentation, errors or missingness, and variation in data collection practices across healthcare settings can impact its quality. In some cases, there may be a significant time lag between data collection and availability for analysis, potentially leading to a delay before the data could be used in a regulatory submission. Another issue is the frequent lack of clear device and manufacturer identification in RWD (especially a standardised device identifier[50,51]), which is particularly challenging for the use of RWD for PMS purposes.

Addressing these limitations requires collaboration between the regulators, the IVD industry, healthcare professionals, and other stakeholders. Efforts are underway to establish robust RWD analysis methodologies[52,53,54,55], to catalogue and validate available RWD sources[56], and to develop consensus on best practices for utilising RWE for regulatory purposes in Europe[57,58]and beyond[59,60,61]. This collaborative approach will help create a better understanding of the strengths and limitations of RWD, ultimately enabling its more effective and reliable utilisation in generating clinical evidence for IVDs.[18]

10) **How is the quality of RWD defined?**

The suitability of a particular RWD source to generate evidence will depend on various factors like the device type, its risk classification, the depth and detail of data needed, the number of exposed patients, and finally the research question and its aim (regulatory, policymaking, health economics, etc.)[62]. The EMA Data Quality Framework (DQF) emphasises a "fit for purpose" definition of quality, focusing on how correct, sufficient, and analysable the data is for specific regulatory purposes within the EU medicines regulation. Even though it doesn't directly impact the IVDR or MDR, it's still a useful resource for understanding how European regulators think about data quality.

According to the EMA DQF, data quality can be contextualised within the following three categories, called determinants (see also Figure 16):

1. **Foundational**: pertaining to the processes and systems through which data are generated, collected, processed, and made accessible.

[18] See https://www.fda.gov/regulatory-information/search-fda-guidance-documents/draft-use-real-world-evidence-support-regulatory-decision-making-medical-devices, which includes a section on what documentation is recommended when submitting RWE to the FDA (Section VII) and gives very detailed expectations of what information is necessary.

2. **Intrinsic**: aspects of the data specific to a given dataset, like its composition, plausibility or correctness.
3. **Question specific**: must be defined with a specific scientific question in mind.

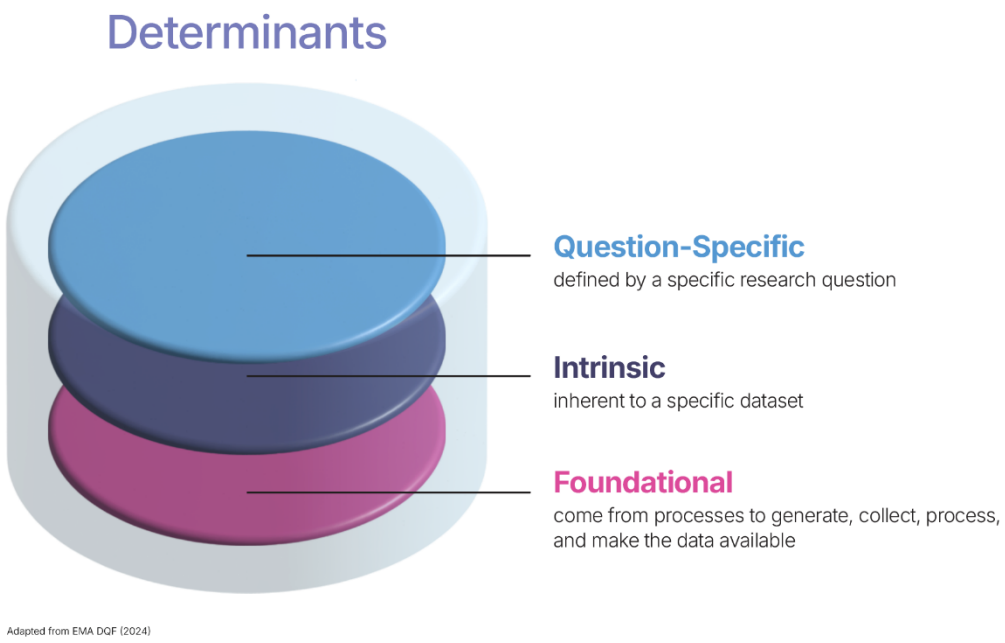


Figure 16: Key data quality determinants

Adapted from https://www.ema.europa.eu/system/files/documents/regulatory-procedural-guideline/data-quality-framework-eu-medicines-regulation_en_1.pdf

The EMA DQF defines data quality along these five key dimensions (see Figure 17):

1. Reliability
2. Extensiveness
3. Coherence
4. Timeliness
5. Relevance

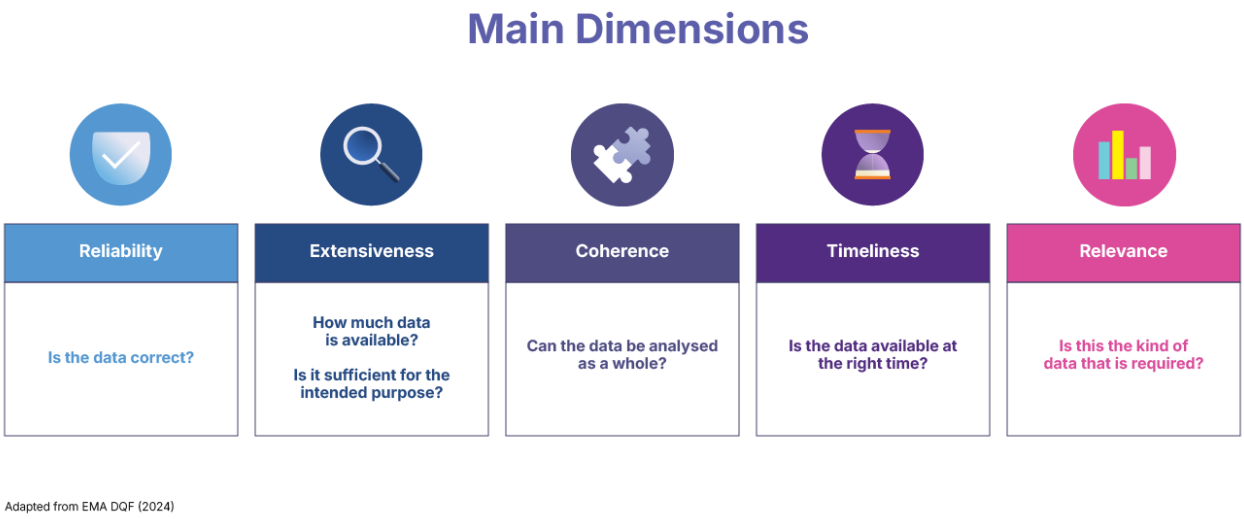


Figure 17: Key dimensions of data quality

1. **Reliability** is defined by how closely the data reflect what they are directly describing or measuring.
Key sub-dimensions:

- Accuracy: How much discrepancy exists between data and reality (e.g. transcription errors)
- Precision: Degree of approximation by which data represents reality (e.g. age in years vs. months)
- Plausibility: Likelihood of data being true, used as a proxy to detect errors
- Traceability: Knowledge of data origin and processing history

2. **Extensiveness** captures the amount of data available.
Key sub-dimensions:

- Completeness: Amount of information available relative to the total possible (e.g. non-missing values)
- Coverage: Amount of information available relative to what exists in the real world (e.g. population percentage in a dataset).
- Representativeness: Data having the same characteristics as the whole it represents
- Missingness: Impact of incomplete data on dataset coverage

3. **Coherence** (or consistency) describes how the different parts of a dataset are consistent in representation and meaning.
Key sub-dimensions:

- Format Coherence: Consistent data expression throughout the dataset (e.g. date formats)
- Structural/Relational Coherence: Consistent entity identification throughout the dataset (e.g. patient or HCP identifiers)
- Semantic Coherence: Consistent meaning of values throughout the dataset (e.g. harmonised lab result units)
- Uniqueness: Non-duplication of information within the dataset

4. **Timeliness** refers to the availability of data at the right time for regulatory decision-making.
Key sub-dimensions:

- Currency: Freshness of data (e.g. current and immediately useful)
- Lateness: Data captured later than asserted, affecting reliability

5. **Relevance** is the extent to which a dataset presents the data elements useful to answer a given research question.
Key sub-dimensions:

- Context-dependence: Characterised in relation to a specific research question and data analysis strategy
- Fit for purpose: How well the data cover the aspects of reality intended to be measured
- Domain relevance: Frequently required research questions characterised for a specific type of data source

Other recommended sources on data quality considerations include:

- The EMA's Real-world evidence framework to support EU regulatory decision-making[63]
- The FDA's Draft Guidance Document: Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices[64]

11) How is the quality of RWD ascertained?

RWD quality assurance is a critical aspect of ensuring the reliability and validity of the data used for RWE generation, making it adequate for the intended purpose. Below are some of the best-practice approaches, which may be used to ensure the quality of RWD when part of regulatory submissions.

In addition, the National Evaluation System for health Technology Coordinating Center (NESTcc) proposed a framework focused on ensuring the overall quality of EHR data used in clinical care and for generating RWE, especially in relation to medical devices[65]. It provides guiding principles across the data lifecycle, not just for a specific research question. Similarly, NICE created a practical tool for assessing the fitness-for-purpose of a specific dataset for a defined research question[66].

Data governance:

- Establish policies and procedures for data access, storage, usage, and sharing
- Implement access controls to protect data confidentiality and integrity

Audit trails:

- Establish processes to ensure complete regulatory-grade documentation
- Maintain logs of data entry and changes to track the history of the data
- Ensure traceability of data transformations and analyses

Ethical and privacy considerations:

- Ensure ethical standards are upheld in the collection and use of RWD, including patient consent where required
- Protect patient privacy and adhere to data protection laws in relevant jurisdictions

Data source/third-party validation:

- Evaluate the credibility and reputation of the data source
- Assess the purpose for which the data was originally collected
- Verify the legal and ethical compliance of the data collection process, including consent and re-use clauses

Data traceability and transparency:

- Maintain clear documentation of the data provenance, including collection methods, data sources, and any transformations applied
- Provide transparency into the methods and processes used in analysing the data

Data accuracy and completeness checks:

- Implement procedures to check for missing or incomplete data
- Analyse the data for outliers or anomalies that might indicate errors
- Validate the data against known benchmarks or other data sources including literature

Standardisation and harmonisation:

- Use standardised coding systems (like International Classification of Diseases (ICD) codes for diagnoses, or Observational Medical Outcomes Partnership (OMOP)) to ensure consistency
- Harmonise data from different sources to allow for comparability, particularly in case of linkage

Data cleaning and preprocessing:

- Perform data cleaning to correct errors and inconsistencies
- Normalise data to ensure consistent formatting and representation

Quality metrics, reporting and monitoring:

- Define quality metrics that are relevant to the specific dataset and its intended use
- Regularly generate quality reports and review them for continuous improvement
- Implement a feedback loop to improve data quality processes over time

External audits and certifications:

- Engage third-party auditors to independently assess the quality of the data

- Seek certifications (e.g. ISO) to demonstrate adherence to quality standards

Scientific validity:

- Use RWD-appropriate and validated methods and study designs

Regulatory guideline compliance:

- Ensure that the use of RWD and the generation of RWE comply with the relevant regulatory guidelines
- Engage with regulatory authorities early and often to understand their expectations and receive feedback

Risk management:

- Identify and manage risks associated with the use of RWD, including privacy concerns and potential biases
- Address the limitations of the data and the analyses in the submission documents

12) How are robustness and transparency of RWD and RWE methodologies ensured?

When using RWD to generate evidence, transparency of the approach (particularly around its limitations) and using appropriate and robust methodologies is key. Limitations related to data sources and data quality need to be clearly highlighted in the study protocol and the analysis plan – which should be completed and consulted with the relevant stakeholders (especially regulatory authorities) before data analysis is conducted. Likewise, it is critical to explain and mitigate bias inherent in the data, including the aspects which are concealed or difficult to control, such as data provenance and the real-world patient population it represents, or missing data elements. A separate section of the protocol should be devoted to discussing the limitations of the specific dataset and the approaches utilised in the study.

Organisations such as the European Network of Centres for Pharmacoepidemiology & Pharmacovigilance (ENCePP)[67] provide best practice guides for RWD analysis methodologies, such as its latest Guide on Methodological Standards in Pharmacoepidemiology (Revision 11)[68].

Another essential resource is the HARPER protocol template published in January 2023 by the joint International Society for Pharmacoepidemiology (ISPE)/ International Society for Pharmacoeconomics and Outcomes Research (ISPOR) task force[69]. It aims to provide a unified approach to enable transparency, reproducibility and harmonisation of non-interventional/RWD study protocols.

In April 2024 the EMA Committee for Medicinal Products for Human Use (CHMP) published their draft reflection paper on use of real-world data in non-interventional studies to generate real-world evidence, which also contains some helpful guidance on RWD study design, data quality, and statistical analysis[70].

Note: there are two promising developments for IVD manufacturers in the US:

1. The NESTcc[71] is currently working on developing the NEST Mark[72], which will provide confirmation that "for the specific study question, the covered Real-World Evidence has followed the NEST Frameworks" according to Medical Device User Fee Amendments (MDUFA) V[73]; and
2. The Medical Device Innovation Consortium's (MDIC) Open Hand[74]project aims to accelerate RWE generation through collaborative, open-source development of standardised methodologies, data sharing frameworks, and analytical tools, aiming to improve patient safety and foster device innovation through high-quality evidence[75].

13) What is the current availability of RWD in Europe and what challenges are impeding RWD collection?

The availability of RWD for both research and regulatory purposes in Europe is currently relatively limited, especially in comparison to the US, which has a well-developed commercial market for RWD. The main challenges are the lack of robust infrastructure to support EU-wide RWD collection, limited interoperability between the many coexisting healthcare systems, uneven adoption of digital healthcare technologies, along with limitations on data sharing and privacy imposed by the GDPR and other national regulations[77]. Despite all this, healthcare digitalisation is progressing across Europe, and many initiatives are making RWD more readily available throughout the continent[78] - though much more effort from the regulatory authorities will be required in the coming years to realize RWD's potential.

DARWIN EU is a research network established by the EMA to generate real-world evidence from healthcare data across Europe to support regulatory decision-making on medicines. It aims to provide a source of high-quality, validated real-world data on diseases, populations, and the uses, safety, and efficacy of medicines, and enable conducting non-interventional studies and interrogating relevant data sources from its network of data partners[79]. This data is, however, only available for regulators and national health authorities. The network had announced 55 completed or ongoing studies as of May 2025.

Other examples of European RWD research networks and consortia include More-EUROPA[80], the Swiss Personalized Health Network[81], and Nordic RWE[82].

To foster broader awareness and access to RWD across Europe for more stakeholders (including medicine and device manufacturers), in April 2024 the EMA and Heads of Medicines Agencies (HMA) announced the launch of **RWD and Study Catalogues**[83]. As of 30 May 2024, the Catalogues contained 217 data sources and 2,828 studies. As the information is collected through voluntary submissions, the available data sources vary widely and their availability for research or RWE generation needs to be established on an individual basis. The list is expected to grow significantly over time, as new submissions are added.

14) What are the latest developments that could enable broader access and use of RWD in Europe and worldwide?

The European Health Data Space (EHDS) is an initiative by the European Union to create a harmonised framework for securely sharing and accessing electronic health data across member states. Following adoption by the EU Parliament and Council, the EHDS Regulation (EU) 2025/327 was published in the Official Journal on March 5, 2025, and entered into force on March 26, 2025. It aims to empower individuals with control over their health data while fostering a single market for digital health services and enabling cross-border access to data for research, innovation, and regulatory purposes[84]. It should also enable better portability of individual healthcare data across the entire EU territory, improving the quality and continuity of care for the patients[85]. The EHDS is expected to foster many potential uses of real-world data for secondary purposes, including clinical evidence generation and supporting regulatory submissions, its key provisions will apply gradually. Notably, the framework facilitating the secondary use of data for research, innovation, and regulatory activities is scheduled to become applicable only from March 26, 2029 (or March 26, 2031 for certain data types like genomic data). Therefore, while a significant development, the practical impact of EHDS on RWD access will unfold over several years as implementing acts are developed and these timelines approach.

Federated networks are emerging as an essential technology for unifying distributed RWD sources, allowing researchers to query and analyse data across multiple institutions without changing the location or ownership of the data[86]. This approach addresses critical data privacy concerns and facilitates collaborative research by breaking down data silos, enabling simultaneous utilisation of multiple data sources. Examples of mature federated RWD networks include European Health Data & Evidence Network[87] (EHDEN), DARWIN EU[88], and the Australian Cancer Data Network[89].

Data linkage and tokenisation techniques enable secure integration of disparate healthcare datasets, significantly improving the comprehensiveness of RWD analyses[90]. These technologies allow for the linking of patient data from various sources (e.g. EMR with claims, or EMR with clinical study data) while minimising the risk of patient reidentification, creating integrated and longitudinal patient profiles that are essential for studying complex health conditions.

15) Would data from outside of the EU also be accepted in regulatory submissions?

As described in Chapter 13 of this eBook, the IVDR permits the use of clinical data collected outside of the EU; by extension, this would also apply to RWD coming from different jurisdictions. The same RWD can potentially be used to generate evidence for submissions to multiple authorities, if shown to be adequately representative of the intended population and accepted by the authority in question.

16) How does RWE compare to traditional clinical evidence?

RWE differs from traditional clinical evidence in several key aspects, which may allow it to complement, supplement, and sometimes even replace traditional evidence generation approaches – however, as

discussed previously, its suitability for a given purpose needs to be appropriately established. Its main differentiating factors are typically large and diverse patient populations, and its potential to reflect real-world effectiveness of interventions[91,92,93]. A concise summary can be found in Table 24 below.

	Traditional clinical evidence	Real-world evidence
Main purpose	Efficacy	Effectiveness
Study design	Experimental, highly controlled	Observational
Clinical practice	As per protocol	Reflects the real world
Population	Homogenous, carefully selected	Heterogenous, depends on data provenance
Sample size	Relatively small	From small to very large
Intervention	Assigned	Observed
Data collection	Specified by the protocol	Reflecting real-world patterns
Follow-up	Short-term	Short- or long-term
Validity	Maximises internal validity	Balances internal/external validity
Causal inference	Usually by design (randomisation)	Requires appropriate design & analysis
Resources required	Costly and time-consuming	Lower cost, fewer resources, and less time required (usually)

	Traditional clinical evidence	Real-world evidence
Access to special/vulnerable populations	Difficult or even impossible	Possible with retrospective medical records or registries

Table 24: Traditional clinical evidence vs. Real-World Evidence

***Disclaimer**: The examples in questions 18-24 below are only indicative of previous approvals and do not constitute advice for future submissions. This list is non-exhaustive; readers are encouraged to explore the following references for more examples:*

- Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions (FDA)[94]
- Real-world evidence framework to support EU regulatory decision-making (HMA/EMA)[95]
- Introduction (Data Saves Lives)[96]

17) How can RWE be used for an IVD claim expansion?

During the COVID-19 pandemic, the FDA authorised emergency use of multiple SARS-CoV-2 serology tests. A sponsor used RWD from a patient registry to support an expansion of intended use of their test for asymptomatic individuals, initially approved only for symptomatic testing. This RWD, along with data from a traditional clinical study and peer-reviewed literature, supported a dual 510(k) and CLIA Waiver by Application[97]. The FDA considered the RWE valid, as the RWD was deemed fit-for-purpose, minimised biases, and included necessary data elements like Logical Observation Identifiers Names and Codes (LOINC) or Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT) codes. The sponsor ensured that device identifiers, patient demographics, clinical information, and rationale for exclusions were appropriately documented. They also addressed the challenges of using different instruments and controls, which could affect performance comparability[98].

18) How can RWE be used to support treatment safety and effectiveness monitoring in rare diseases?

Evidence generation for rare diseases can be particularly challenging, especially due to the very small number of available patients, and lack of in-depth clinical information. Case studies or reports are frequently used as lower-tier evidence in absence of well-designed studies. The following example showcases the potential of linked disease registries.

Spinal muscular atrophy (SMA) is a rare neuromuscular disease, with some of its forms particularly severe in paediatric patients. The first approved treatment for SMA, nusinersen, received regulatory clearance based on clinical trials that focused on a narrow paediatric population. As the therapy usage broadened and other targeted treatments gained traction, there was a need to closely monitor potential rare adverse events and long-term effectiveness of SMA treatments in a larger patient population.

The Canadian Neuromuscular Disease Registry[99] is a longitudinal, prospective, observational study designed to evaluate the safety and effectiveness of SMA therapies in Canadian patients, who are enrolled regardless of their treatment status. The expanded dataset aligns with other existing databases to facilitate international collaboration and aims to standardise SMA-relevant outcome measures. It will provide valuable insights into the safety and effectiveness of SMA therapies in the post-approval stage, which are essential to inform improvements in care and access to therapy for all SMA patients.

In Europe, RWE generated after the launch of nusinersen was used to enhance product information with clinical effectiveness in adult SMA Type II and III patients, confirming the treatment's safety profile in the adult population[100].

19) **How can RWE be used to support IVD registration in paediatric populations or other vulnerable patient populations?**

Vulnerable patient populations, like newborns and infants, the elderly, or those suffering from particularly sensitive or life-threatening conditions, can be very challenging to study directly or recruit for clinical studies. In such cases, the use of RWD may result in larger sample sizes and significantly shorten the time to conduct a study.

Example 1: Lysosomal storage diseases (LSDs) are a group of rare conditions which cause neurodegeneration, as well as other organ pathogenesis. They usually present in infancy and childhood, making them a potential target of newborn screening. An IVD device for measuring the activity of enzymes associated with LSDs in dried blood spots used premarket evidence from a pivotal trial embedded in a public newborn screening program to support their *de novo* FDA application[101].

Example 2: Serum interleukin-6 (IL-6) is a highly sensitive marker of immune response widely used in sepsis diagnosis, including in newborns. However, the diagnostic cut-off values were not available for this population. To determine those values, researchers at the University of Vienna conducted a retrospective cohort study using electronic medical records from a neonatal intensive care ward, based on 8,488 measurements in 1,695 patients. The study determined that IL-6 has distinct diagnostic cut-off values in newborns, which can be precisely stratified by age in days[102].

20) **How can RWE be used in Post Market Surveillance/Follow-Up for IVDS?**

A manufacturer of an IVD device which detects genomic DNA variants in whole blood was required to create a database of all test results obtained during real-world use of their product as a condition of FDA approval. They were required to monitor all gene variants detected by the assay, as well as its performance in respect to different sample collection tubes[103].

21) How can RWE be used for digital health device registration?

There are multiple examples of using RWE to support registration of digital health products, mainly thanks to the easier access to the data flowing through or generated by the use of the device. Here, we will highlight one from a European manufacturer who utilised primarily European user data for a successful FDA submission.

A manufacturer of a patient-facing fertility tracking and contraceptive digital health application used a retrospective analysis of data collected from 15,000 non-US users (majority located in Europe) as a primary source of clinical evidence to support their *de novo* classification request to the FDA[104].

22) **How can RWE be used for medical device registration?**

Many examples of RWE use for medical devices have been described in the CDRH report cited above[105]. Here we highlight another interesting use case from the China National Medical Products Administration (NMPA):

The 2020 NMPA approval of a glaucoma drainage device marked the first use of domestic RWE from the Boao Lecheng pilot zone for medical device registration in China. The application integrated international clinical data with locally generated RWE and was enabled by the pilot zone's accelerated regulatory pathway, aimed at increasing patient access to innovative interventions. The approval time was reduced from the usual 3-5 years to just 5 months[106].

23) **How can RWE be used in the pharmaceutical sector?**

The use of RWE for registration and safety monitoring of pharmaceuticals is already quite well-established, in part thanks to much clearer guidance[107,108], the FDA's Sentinel Initiative, and easier identification of medical products in RWD sources (e.g., using NDCs[109].

Below is an example of medication label expansion into a new patient population:

Palbociclib was approved for treatment of female patients with hormone receptor-positive (HR+), human epidermal growth factor receptor 2-negative (HER2-) advanced or metastatic breast cancer; however, it was not approved for the very rare male patients. Real-world evidence from three sources (claims, electronic health records, and post-marketing safety reports) was used to evaluate the safety and effectiveness of palbociclib in male breast cancer patients without conducting a clinical trial. RWE showed that the treatment was both effective and safe in male patients, with outcomes comparable to female patients. Consequently, palbociclib was officially approved for male patients, demonstrating the value of real-world data in expanding treatment options for underrepresented patient groups[110]

For an in-depth analysis of RWE use in the pharmaceutical industry, as well as further examples, please refer to the recently published Council for International Organizations of Medical Sciences (CIOMS) WG XIII report: Real-world data and real-world evidence in regulatory decision making[111].

References:

1. Health Insurance Portability and Accountability Act of 1996 (HIPAA) | Public Health Law | CDC https://www.cdc.gov/phlp/php/resources/health-insurance-portability-and-accountability-act-of-1996-hipaa.html
2. THE HITECH ACT: An Overview | Journal of Ethics https://journalofethics.ama-assn.org/article/hitech-act-overview/2011-03
3. Adaptive pathways | European Medicines Agency (EMA) https://www.ema.europa.eu/en/human-regulatory-overview/research-development/adaptive-pathways
4. 21st Century Cures Act | FDA https://www.fda.gov/regulatory-information/selected-amendments-fdc-act/21st-century-cures-act
5. Boao Lecheng pilot zone explores fast track into China's medical market https://www.chinadaily.com.cn/a/202402/08/WS65c49669a3104efcbdaea7cb.html
6. Next Generation Medical Infrastructure Law https://www.ldi.or.jp/en/law
7. Real World Evidence (RWE) and patient reported outcomes (PROs) | Therapeutic Goods Administration (TGA) https://www.tga.gov.au/real-world-evidence-rwe-and-patient-reported-outcomes-pros
8. Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions | FDA https://www.fda.gov/media/146258/download
9. Real-world evidence framework to support EU regulatory decision-making | EMA https://www.ema.europa.eu/en/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained-regulator-led-studies-september-2021-february-2023_en.pdf
10. European Health Data Space Regulation (EHDS) https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en
11. Data Analysis and Real World Interrogation Network (DARWIN EU) | EMA https://www.ema.europa.eu/en/about-us/how-we-work/data-regulation-big-data-other-sources/real-world-evidence/data-analysis-real-world-interrogation-network-darwin-eu
12. FDA's Sentinel Initiative https://www.fda.gov/safety/fdas-sentinel-initiative
13. FDA In Brief: FDA announces a new Sentinel System contract, affirming its commitment to harnessing Real-World Data to improve the safety and effectiveness of drugs https://www.fda.gov/news-events/fda-brief/fda-brief-fda-announces-new-sentinel-system-contract-affirming-its-commitment-harnessing-real-world
14. International Harmonization of Real World Evidence Standards Dashboard | Duke Margolis Institute for Health Policy https://healthpolicy.duke.edu/projects/international-harmonization-real-world-evidence-standards-dashboard
15. Real-world evidence provided by EMA Support for regulatory decision-making https://www.ema.europa.eu/system/files/documents/other/guidance-real-world-evidence-provided-ema_en.pdf
16. Glossary | NICE https://www.nice.org.uk/Glossary?letter=R

17. NICE real-world evidence framework https://www.nice.org.uk/corporate/ecd9/chapter/overview
18. Swissmedic position paper on the use of real-world evidence https://www.swissmedic.ch/swissmedic/en/home/news/mitteilungen/positionspapier-verwendung-real-world-evidence.html
19. Real-World Evidence | FDA https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
20. Proposal for a Regulation of the European Parliamant and of the Council on the European Health Data Space, Document 52022PC0197 | EUR-Lex https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0197
21. Real-World Data: Assessing Electronic Health Records and Medical Claims Data To Support Regulatory Decision-Making for Drug and Biological Products | FDA https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-electronic-health-records-and-medical-claims-data-support-regulatory
22. Claims Data | NIH https://www.nlm.nih.gov/oet/ed/stats/03-300.html
23. Patient registries | EMA https://www.ema.europa.eu/en/human-regulatory-overview/post-authorisation/patient-registries
24. Real-World Data: Assessing Registries To Support Regulatory Decision-Making for Drug and Biological Products | FDA https://www.fda.gov/regulatory-information/search-fda-guidance-documents/real-world-data-assessing-registries-support-regulatory-decision-making-drug-and-biological-products
25. EudraVigilance | EMA https://pmc.ncbi.nlm.nih.gov/articles/PMC9790425/ Baumfeld Andre E, Carrington N, Siami FS, Hiatt JC, McWilliams C, Hiller C, Surinach A, Zamorano A, Pashos CL, Schulz WL. The Current Landscape and Emerging Applications for Real-World Data in Diagnostics and Clinical Decision Support and its Impact on Regulatory Decision Making. Clin Pharmacol Ther. 2022 Dec;112(6):1172-1182. doi: 10.1002/cpt.2565. Epub 2022 Apr 29. PMID: 35213741; PMCID: PMC9790425. https://ascpt.onlinelibrary.wiley.com/doi/10.1002/cpt.2565
26. EudraVigilance | EMA https://www.ema.europa.eu/en/human-regulatory-overview/research-development/pharmacovigilance-research-development/eudravigilance
27. Yellow Card | MHRA https://yellowcard.mhra.gov.uk/
28. List of Field Safety Corrective Actions (FSCAs) | Swissmedic https://www.swissmedic.ch/swissmedic/en/home/medical-devices/fsca.html
29. Focus Area: Product Safety Surveillance | FDA https://www.fda.gov/science-research/focus-areas-regulatory-science-report/focus-area-product-safety-surveillance
30. About Manufacturer and User Facility Device Experience (MAUDE) Database | FDA https://www.fda.gov/medical-devices/mandatory-reporting-requirements-manufacturers-importers-and-device-user-facilities/about-manufacturer-and-user-facility-device-experience-maude-database
31. Use of Real-World Evidence to Drive Drug Development Strategy and Inform Clinical Trial Design | NIH https://pmc.ncbi.nlm.nih.gov/articles/PMC9299990/
32. Patient-Generated Health Data | HealthIT.gov https://www.healthit.gov/topic/scientific-initiatives/pcor/patient-generated-health-data-pghd
33. Reflection paper on the regulatory guidance for the use of health-related quality of life (HRQL) measures in the evaluation of medicinal products | EMA https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-regulatory-guidance-use-health-related-quality-life-hrql-measures-evaluation-medicinal-products_en.pdf
34. Mandatory reporting system | Swiss Federal Office of Public Health (FOPH) https://www.idd.bag.admin.ch/survey-systems/oblig
35. About us | UK Biobank https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/about-us
36. Unlocking the Potential of Medical Imaging as Real-World Data | American Pharmaceutical Review https://www.americanpharmaceuticalreview.com/Featured-Articles/597636-Unlocking-the-Potential-of-Medical-Imaging-as-Real-World-Data/
37. Real-world evidence provided by EMA https://www.ema.europa.eu/system/files/documents/other/guidance-real-world-evidence-provided-ema_en.pdf
38. Glossary | NICE https://www.nice.org.uk/Glossary?letter=R

39. Swissmedic position paper on the use of real-world evidence https://www.swissmedic.ch/swissmedic/en/home/news/mitteilungen/positionspapier-verwendung-real-world-evidence.html
40. Real-World Evidence | FDA https://www.fda.gov/science-research/science-and-research-special-topics/real-world-evidence
41. Regulation 2017/746/ EU of the European parliament and of the council of April 5, 2017 on in vitro diagnostic medical devices
42. Guidance document Medical devices - Market surveillance - Post Market Clinical Follow-up studies, MEDDEV 2.12-2 Rev. 2 | European Commission https://ec.europa.eu/docsroom/documents/10334/attachments/1/translations
43. MDCG 2020-6 Regulation (EU) 2017/745 | European Commission https://ec.europa.eu/docsroom/documents/40904
44. MDCG 2020-7 Post-market clinical follow-up (PMCF) Plan Template. A guide for manufacturers and notified bodies | European Commission https://ec.europa.eu/docsroom/documents/40905
45. MDCG 2022-2 - Guidance on general principles of clinical evidence for In Vitro Diagnostic medical devices (IVDs) | European Commission https://health.ec.europa.eu/latest-updates/mdcg-2022-2-guidance-general-principles-clinical-evidence-vitro-diagnostic-medical-devices-ivds-2022-01-27_en
46. Indirect methods for reference interval determination – review and recommendations | De Gruyter Brill https://www.degruyter.com/document/doi/10.1515/cclm-2018-0073/html
47. Comparison of three indirect methods for verification and validation of reference intervals at eight medical laboratories: a European multicenter study | De Gruyter Brill https://www.degruyter.com/document/doi/10.1515/labmed-2023-0042/html?lang=en]
48. Real-world evidence provided by EMA https://www.ema.europa.eu/system/files/documents/other/guidance-real-world-evidence-provided-ema_en.pdf
49. Assessing and Interpreting Real-World Evidence Studies: Introductory Points for New Reviewers | PMC https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8771197/
50. Unique Device Identifier | European Commission https://health.ec.europa.eu/medical-devices-topics-interest/unique-device-identifier-udi_en
51. What We Do | GMDN https://www.gmdnagency.org/what-we-do/
52. HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force | PubMed https://pubmed.ncbi.nlm.nih.gov/36215113/
53. STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies | The BMJ https://www.bmj.com/content/372/bmj.m4856
54. Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence - Scientific guideline | EMA https://www.ema.europa.eu/en/reflection-paper-use-real-world-data-non-interventional-studies-generate-real-world-evidence-scientific-guideline
55. Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence - Scientific guideline | EMA https://www.fda.gov/regulatory-information/search-fda-guidance-documents/draft-use-real-world-evidence-support-regulatory-decision-making-medical-devices
56. HMA-EMA Catalogues https://catalogues.ema.europa.eu/
57. DARWIN EU https://www.darwin-eu.org/index.php
58. Innovative Health Initiative https://www.ihi.europa.eu/
59. Draft: Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices | FDA https://www.fda.gov/regulatory-information/search-fda-guidance-documents/draft-use-real-world-evidence-support-regulatory-decision-making-medical-devices
60. NESTcc Data Quality Framework https://nestcc.org/data-quality-and-methods/
61. Real-World Evidence | ISPOR https://www.ispor.org/strategic-initiatives/real-world-evidence
62. Data Quality Framework for EU medicines regulation | HMA/EMA https://www.ema.europa.eu/system/files/documents/regulatory-procedural-guideline/data-quality-framework-eu-medicines-regulation_en_1.pdf

63. Real-world evidence framework to support EU regulatory decision-making | EMA https://www.ema.europa.eu/en/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained-regulator-led-studies-september-2021-february-2023_en.pdf
64. Draft: Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices | FDA https://www.fda.gov/regulatory-information/search-fda-guidance-documents/draft-use-real-world-evidence-support-regulatory-decision-making-medical-devices
65. NESTcc Data Quality Framework https://nestcc.org/nestcc-data-quality-framework/
66. Appendix 1: Data Suitability Assessment Tool (DataSAT) | NICE real-world evidence framework https://www.nice.org.uk/corporate/ecd9/chapter/appendix-1-data-suitability-assessment-tool-datasat
67. ENCePP https://encepp.europa.eu/index_en
68. Methodological Guide - European Union | ENCePP https://encepp.europa.eu/encepp-toolkit/methodological-guide_en
69. HARmonized Protocol Template to Enhance Reproducibility of hypothesis evaluating real-world evidence studies on treatment effects: A good practices report of a joint ISPE/ISPOR task force | PubMed https://pubmed.ncbi.nlm.nih.gov/36215113/
70. Reflection paper on use of real-world data in non5 interventional studies to generate real-world evidence (Draft) | EMA https://www.ema.europa.eu/en/reflection-paper-use-real-world-data-non-interventional-studies-generate-real-world-evidence-scientific-guideline
71. NESTcc | https://nestcc.org/about/about-us/
72. MDIC News | Medical Device Innovation Consortium https://mdic.org/news/latest-clearance-success-for-intuitive-powered-by-nest-mark/
73. Medical Device User Fee Amendments (MDUFA) | FDA https://www.fda.gov/industry/fda-user-fee-programs/medical-device-user-fee-amendments-mdufa
74. Open Hand | Medical Device Innovation Consortium https://mdic.org/our-work/open-hand/
75. The Open Hand Initiative: Facilitating the Use of Real-World Evidence in Regulatory Submissions Through Collaboration and Transparency | PMC https://pmc.ncbi.nlm.nih.gov/articles/PMC11924146/
76. White Paper on Data Protection, Privacy and Global Health Data within the Medical Technology Industry | GMTA http://www.globalmedicaltechnologyalliance.org/papers/20221017_privacy_white-paper.pdf
77. Special report 25/2024: Digitalisation of healthcare – EU support for member states effective overall, but difficulties in using EU funds | European Court of Auditors https://www.eca.europa.eu/en/publications/SR-2024-25
78. Accelerating digital health transformation in Europe: a two-year progress report | World Health Organization https://www.who.int/europe/news/item/28-10-2024-accelerating-digital-health-transformation-in-europe--a-two-year-progress-report
79. https://www.darwin-eu.org/index.php/studies
80. More-EUROPA – Research https://umcgresearch.org/more-europa
81. Swiss Personalized Health Network (SPHN) https://sphn.ch/
82. Insights | NordicRWE https://www.nordicrwe.com/insights
83. Launch of new HMA-EMA catalogues of real-world data sources and studies https://www.ema.europa.eu/en/news/launch-new-hma-ema-catalogues-real-world-data-sources-studies, https://catalogues.ema.europa.eu/
84. European Health Data Space Regulation | European Commission https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en, accessed 30 May 2024
85. Q&A on the European Health Data Space | European Commission https://ec.europa.eu/commission/presscorner/detail/en/qanda_24_2251
86. Federated Networks for Distributed Analysis of Health Data | PMC https://pmc.ncbi.nlm.nih.gov/articles/PMC8514765/
87. Ehden.eu https://www.ehden.eu/
88. DARWIN EU | EMA https://www.ema.europa.eu/en/about-us/how-we-work/big-data/data-analysis-real-world-interrogation-network-darwin-eu

89. Australian Cancer Data Network (ACDN) | ARDC https://ardc.edu.au/project/australian-cancer-data-network/
90. Linking clinical trial participants to their U.S. real-world data through tokenization: A practical guide | PMC https://pmc.ncbi.nlm.nih.gov/articles/PMC11399707/
91. The importance of collecting structured clinical information on multiple sclerosis | BMC Medicine https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-016-0627-1
92. Real-World Evidence and Randomized Studies in the Precision Oncology Era: The Right Balance https://ascopubs.org/doi/10.1200/PO.17.00132
93. Conducting real-world evidence studies in India | PMC https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6463503/
94. Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions | FDA https://www.fda.gov/media/146258/download
95. Real-world evidence framework to support EU regulatory decision-making | HMA/EMA https://www.ema.europa.eu/en/documents/report/real-world-evidence-framework-support-eu-regulatory-decision-making-report-experience-gained-regulator-led-studies-september-2021-february-2023_en.pdf
96. Introduction — Data Saves Lives https://datasaveslives.eu/introductioncase
97. CLIA Categorizations | FDA https://www.fda.gov/medical-devices/ivd-regulatory-assistance/clia-categorizations
98. Draft: Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices | FDA https://www.fda.gov/regulatory-information/search-fda-guidance-documents/draft-use-real-world-evidence-support-regulatory-decision-making-medical-devices
99. A National Spinal Muscular Atrophy Registry for Real-World Evidence | PMC https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7656664/
100. Patient registries are a valuable source of real-world data (RWD) | Data Saves Lives https://datasaveslives.eu/patient-registries-are-a-valuable-source-of-real-world-data-rwd
101. Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions, Example 85 | FDA https://www.fda.gov/media/146258/download
102. Cut-off values of serum interleukin-6 for culture-confirmed sepsis in neonates | Pediatric Research https://www.nature.com/articles/s41390-022-02329-9
103. Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions, Example 89 | FDA https://www.fda.gov/media/146258/download
104. Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions, Example 72 | FDA https://www.fda.gov/media/146258/download
105. Examples of Real-World Evidence (RWE) Used in Medical Device Regulatory Decisions | FDA https://www.fda.gov/media/146258/download
106. Use of real-world evidence to support regulatory decisions on medical devices in China and a unique opportunity to gain accelerated approval in “Boao Lecheng Pilot Zone” | BMC | Cost Effectiveness and Resource Allocation https://resource-allocation.biomedcentral.com/articles/10.1186/s12962-022-00412-w
107. Considerations for the Use of Real-World Data and Real-World Evidence To Support Regulatory Decision-Making for Drug and Biological Products | FDA https://www.fda.gov/regulatory-information/search-fda-guidance-documents/considerations-use-real-world-data-and-real-world-evidence-support-regulatory-decision-making-drug
108. Reflection paper on use of real-world data in non-interventional studies to generate real-world evidence - Scientific guideline | EMA https://www.ema.europa.eu/en/reflection-paper-use-real-world-data-non-interventional-studies-generate-real-world-evidence-scientific-guideline
109. National Drug Code Directory | FDA https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory
110. Using real-world evidence for label expansion: Use of palbociclib in male breast cancer patients | Flatiron https://resources.flatiron.com/publications/real-world-evidence-label-expansion
111. Real-world data and real-world evidence in regulatory decision making | CIOMS https://cioms.ch/publications/product/real-world-data-and-real-world-evidence-in-regulatory-decision-making/

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 that 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:

Iana Slobodeaniuc
Senior Manager IVDs, Industrial Policies
MedTech Europe
regulatory@medtecheurope.org

Reference: 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 (EU) Regulation 2017/746, it is ultimately for the courts to interpret legislation.