



CLINICAL EVIDENCE

How to Choose the Right Evidence
Strategy for Your Medical Device

PART I – CLINICAL USE

4BetterDevices GmbH

VERSION1

Clinical Evidence

How to Choose the Right Evidence Strategy
for Your Medical Device

PART I – CLINICAL USE

4BetterDevices GmbH

Version 1

Copyright © 2026 4BetterDevices GmbH

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the author.

Disclaimer

This document has been created by 4BetterDevices GmbH without the endorsement, support, or approval of the Medical Device Coordination Group (MDCG) nor the European Commission. The comments, corrections, and examples provided in this document reflect solely the opinions of the 4BetterDevices GmbH members. 4BetterDevices GmbH makes no warranties or representations, express or implied, about the completeness, accuracy, reliability, suitability, or availability with respect to the document or the information, products, services, or related graphics contained within the document for any purpose. Any reliance placed on such information is therefore strictly at your own risk.

In no event will 4BetterDevices GmbH be liable for any loss or damage including without limitation, indirect or consequential loss or damage, or any loss or damage whatsoever arising from loss of data or profits arising out of, or in connection with, the use of this document. Users are encouraged to consult with professional advisers for advice concerning specific matters before making any decision based on the information in this document. This disclaimer includes any possible liability for errors, omissions, or inaccuracies in the document.

Version History

Version 1	First version released.
-----------	-------------------------

Contents

Part I: Clinical Use	1
1 Introduction	1
2 What is clinical and what is pre-clinical?	2
2.1 The framework	4
2.2 Clinical decision loop	5
2.3 Clinical action loop	7
2.4 One device, several loops	8
3 “Using” a device clinically	8
Conclusion to Part I	11
Meet 4BetterDevices	12

Acronyms

AI Artificial Intelligence

ECG Electrocardiogram

IVD In Vitro Diagnostic

IVDR In-Vitro Diagnostic Regulation

MDR Medical Device Regulation

1 Introduction

“For our device, pre-clinical data is enough,” the regulatory lead said. She sounded confident. On my screen, I could see a few heads nodding.

“Okay,” I said. “Why?”

A short pause.

“Well, first of all, it’s software. It doesn’t touch the patient. It just provides supporting information.”

“And in the end, it’s the clinician who makes the final decision anyway,” someone added. “We don’t give a result.”

I waited to see if anyone would offer something more concrete. They did not look particularly convinced.

It turned out that, for their device, pre-clinical data was indeed enough. At least for demonstrating performance.

It also turned out that the pre-clinical data they needed were not the pre-clinical data they had in mind. When they spoke of pre-clinical data, they meant data generated from device testing in-house. When I spoke of pre-clinical data, I meant data from *patients*.

“Wait, wait a second, Cesare!” they stopped me. “We don’t have the time or money now to run a randomized controlled trial, you know?”

What followed was a long discussion about pre-clinical and clinical data, about performance and safety, about acceptance criteria and state of the art.

And, in case you are wondering: no, they did not ultimately need to run a clinical trial for certification.

But that discussion led to a workshop, which was followed by many others.

Over time, those workshops turned into a framework we kept using again and again.

Rather than keeping that framework inside individual projects and workshops, I decided to put it together and make it publicly available.

That turned out to be harder than expected. It took us months to structure the material, challenge our own assumptions, and turn years of project work into something that could actually be used by others.

This guide is the result.

It will take you through a way of thinking about medical devices that is rarely presented explicitly. Much of it comes from working across hundreds of devices, resolving countless

deficiencies, and spending many long hours discussing clinical evidence, performance, safety, and regulatory strategy with manufacturers.

The objective is not to give you another summary of the MDR or IVDR. It is to give you a framework for reasoning about the questions behind the regulations.

And the best place to start is with the question that started those workshops in the first place:

2 What is clinical and what is pre-clinical?

Neither the MDR nor the IVDR provide a definition of non-clinical or pre-clinical data. The MDR characterizes non-clinical data by listing examples: results of tests, such as engineering, laboratory, simulated use and animal tests, and published literature applicable to the device, biocompatibility, physical, chemical and microbiological characterisation, electrical safety and electromagnetic compatibility, software verification and validation, stability, including shelf life (MDR, Annex II, Section 6.1).¹

In practice, non-clinical data can be understood as any evidence that does not meet the MDR definition of clinical data. The MDR defines clinical data in Article 2(48) as follows:

‘clinical data’ means information concerning safety or performance that is generated from the use of a device and is sourced from the following:

- *clinical investigation(s) of the device concerned,*
- *clinical investigation(s) or other studies reported in scientific literature, of a device for which equivalence to the device in question can be demonstrated,*
- *reports published in peer-reviewed scientific literature on other clinical experience of either the device in question or a device for which equivalence to the device in question can be demonstrated,*
- *clinically relevant information coming from post-market surveillance, in particular the post-market clinical follow-up.*

Clinical data is therefore *not* defined by the setting (e.g. “in a clinic”), nor by the subject (e.g. “data from patients”). For example, a manufacturer of an AI-based medical device may collect annotated datasets from the target patient population in a clinical environment for the purpose of training the model. Although such data may originate

¹The European Commission’s proposal for a targeted revision of the MDR and IVDR, published on 16 December 2025 expressly adds in vitro, ex vivo and in silico testing and computational modeling to the examples of testing that may provide evidence regarding the pre-clinical safety of a device. At the time of writing, however, this amendment remains a legislative proposal and has not yet become applicable law.

from patients and clinical workflows, it does not involve the *use of the device* itself and is therefore not clinical data under the MDR.

Instead, three factors are critical in determining whether data qualifies as clinical data:

1. **How the data were generated:** whether the data were generated “*through the use of a device*”;
2. **Which device was used:** whether the data concern “*the device in question*” or a “*device for which equivalence can be demonstrated*”; and
3. **The activity that generated the data:** whether the data arose from “*a clinical investigation*”, or clinical experience documented in “*peer-reviewed scientific literature*”, or “*post-market surveillance*”.

Of the three points, the first—clinical “*use of a device*”—sounds the most intuitive. That is perhaps why it also causes the most confusion. In practice, this concept is far less clear than it appears.

Consider a surgical knife. If a surgeon uses it during an operation, that clearly constitutes use of the device. No ambiguity there.

Now consider a manufacturer that applies an AI-based software to identify cancer lesions in X-ray images. Does this constitute “use” of the device in the regulatory sense? The answer depends on what is done with the information coming from the software.

If a doctor uses that information to decide on subsequent patient management, then yes, the device is used as intended. But what if the same software is run retrospectively on a set of X-ray images, purely to assess how accurately it identifies cancer lesions, and its output is never shown to the treating physician and never influences patient management?

The software is certainly being *operated*. But is it being *used* as a medical device?

This is where the apparently simple concept of “use of a device” becomes difficult. Operating a device is not necessarily the same as using it for its medical purpose. To determine whether a device is being used as a medical device, we need to understand what role it plays in the clinical process: what information enters the device, what the device does with that information, what comes out, and what happens as a consequence.

When dealing with medical devices, you will encounter thousands of variations of this problem. Fortunately, they can be analysed systematically.

The key is to stop looking at the device in isolation and instead look at the clinical process in which it operates.

💡 Tip: Clinical data vs. clinical evidence

“Clinical data” and “clinical evidence”. In everyday conversation, these two terms are often used interchangeably. In the MDR, they are not. Clinical data is only one component of clinical evidence. Clinical evidence combines clinical data with the conclusions of the clinical evaluation. In other words, clinical data is what you collect; clinical evidence is what you demonstrate.

2.1 The framework

Clinical practice, and medical devices as part of it, can be described in terms of nested *clinical decision* and *clinical action* loops (Figure 1).

In a *clinical decision loop*, an input is obtained from the patient, processed, and turned into information that can support a decision. A medical device may participate at any point in this loop: it may obtain the input, process it, interpret it, communicate the resulting output, or, in some cases, take the decision itself.

The clinical decision closes the loop, and it can go three ways. Do nothing. Get more information, which starts another clinical decision loop. Or act on the patient. Only

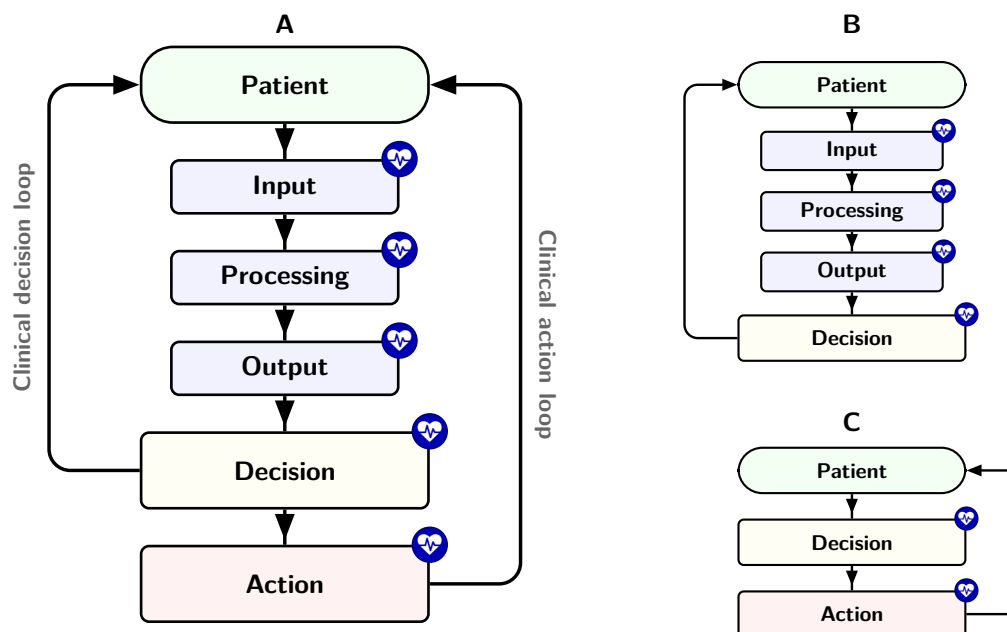


Figure 1: Clinical workflow. **A** Complete workflow combining both clinical decision and clinical action loops. **B** Clinical decision loop: input (information or biological material) obtained from the patient is processed and the output presented to support a clinical decision. **C** Clinical action loop: a clinical decision is turned into a clinical action performed on the patient. 🩺 Step in which a medical device can be used.

the third option starts a *clinical action loop*: the action is carried out on the patient, it changes the patient's state, and the changed state is the input of the next clinical decision loop.

The two loops are therefore closely connected. A clinical decision loop produces the basis for clinical action, while clinical action changes the state from which new information is obtained.

This distinction gives us a practical way to examine what it actually means to *use* a medical device. We will look at the two loops separately. We start with the clinical decision loop, because that is where clinical practice starts: before anything is done to a patient, somebody has to find out what is wrong.

💡 Tip: Patient and user

“Patient” and “user” are roles, and the same person can have both. The patient is the source of the input or the recipient of an action or decision. The user is anyone who assesses the device's suitability, selects or prepares it, prepares the patient, sets or adjusts parameters, interprets outputs, stops or removes the device, or disposes of it. A nurse who applies ECG electrodes and a physician who reads the trace are both users of the same device.

2.2 Clinical decision loop

A clinical decision loop has four stages: input, processing, output and decision. A device may contribute to one stage or several. We examine each contribution separately.

A clinical decision loop is fundamentally informational. Its immediate result is information about the patient, rather than a change in the patient's state. This remains true when the input is physical material. In laboratory testing, for example, a specimen is obtained from the patient and processed to produce a result. The specimen does not return to the patient; the result enters the clinical decision. Similarly, an image, measurement or physiological signal is processed to produce information that can guide what happens next. Any benefit to the patient arises through the resulting decision and, where applicable, the action that follows.

Input

The input originates from the patient, even when a user selects, records, collects or enters it into the device. Here, *input* means the information or biological material that enters the clinical decision loop, not the act of entering data into the device. It may be information, such as symptoms, history, a measurement, a signal or an image; or biological material, such as blood, urine, tissue or a swab. Devices acquire these inputs, but they can also shape their *content*, *timing* and *quality*.

A device shapes the *content* when it determines what is obtained: which question is

asked, which parameter is measured, which structure is imaged, or which specimen is taken and from where. An anamnesis system that selects questions from earlier answers decides part of what the physician will receive.

A device shapes the *timing* when it decides when, how often or for how long to acquire the input. A monitor that shortens the measurement interval when readings drift takes over a decision from the healthcare professional. Timing matters because a measurement or specimen obtained too early or too late may miss the condition being investigated. A reminder alone does not transfer this decision: if the physician set the time, the decision remains the physician's.

A device affects *quality* through how faithfully the input represents the patient. This may depend directly on technical performance, such as sensor accuracy, image resolution or whether a collection tube haemolyses a sample. It may also depend on guidance: an app helps position the camera, or a meter rejects a strip with too little blood. Unlike content and timing, quality cannot be left entirely to the professional. Every device that acquires an input affects its quality. The question is which aspects the device controls and which depend on the user.

Processing

Processing turns the input into information: a physical quantity becomes a value, detector data an image, a signal a finding, several inputs a score, or a specimen an assay result. This is the main mode of action of diagnostic devices, including imaging systems, measuring devices, in-vitro diagnostics and non-therapeutic software. Some also acquire and present information; others only process it.

Two things matter here. First, processing can transfer responsibility from the professional to the device. A scoring algorithm contains the rule; an image-classification model contains the judgement. The user receives the result without necessarily being able to see how that rule or judgement was applied.

Second, processing can lose information irreversibly. Filtering a signal, compressing an image, rounding a value or consuming a specimen may remove information that cannot be recovered from the result. That loss carries through the rest of the loop.

Output

The output is what the device makes available for the next decision. It may represent the patient's condition directly, as an image or waveform; express a derived result, such as a value, classification or score; or communicate an alert, recommendation or decision made by the device. When the output is presented to a person, this is where the device hands responsibility back. But a physician can only correct what they can see. Information lost during processing or omitted from the output cannot be checked simply because a physician is at the end of the chain.

Displays, alarms, reports and image-viewers shape the output through the same three characteristics as the input. *Content* is what is shown, prioritised or omitted. More information is not necessarily more useful: selection influences the decision. *Timing* is when the information arrives, including display refresh rates, alarm timing and delays. *Quality* is how faithfully and clearly the result is presented, through choices such as image windows, ECG scales, colour maps, units and decimal places.

The device also influences the decision through information supplied *with* the output. Instructions for use, labelling and claims tell the decision maker how to interpret it and what to expect. A claim of clinical benefit may affect whether a physician prescribes the device or how much they trust its output. This material is all part of the device's contribution to the loop.

Decision

The decision closes the loop. There are three options: obtain more information, starting another clinical decision loop; act on the patient, starting a clinical action loop; or do nothing. The decision maker may be a healthcare professional, a patient, a caregiver or the device itself. An automated external defibrillator, for example, decides whether a shock is indicated. When a device takes the decision, it takes over the final stage and the responsibility that comes with it.

Although the loop produces information, it carries risk at both ends. Acquiring the input may itself be an intervention: an X-ray delivers radiation, a biopsy removes tissue, and a blood draw removes blood. At the other end, a wrong output may lead to a wrong action or a wrong decision not to act. Information is not harmless merely because it does not act directly on the body.

2.3 Clinical action loop

A clinical action loop begins with a decision to act. Somebody, or something, carries out the action, and the patient's state changes: *decision, action, effect*. The changed state becomes the input of the next clinical decision loop.

Usually, the action is physical. Pumps and syringes deliver substances; implants and dressings are placed into or onto the body; surgical instruments act on tissue; and other devices apply energy or radiation. Dialysis removes blood, changes it and returns it. Although it processes blood, what it delivers to the patient is a change in composition, so this contribution belongs to the action loop.

Actions can also work through information. A guided psychotherapy programme or a physiotherapy plan delivered through videos gives instructions whose effect comes from the patient following them. The distinction is what the information does in the loop: here, it guides the therapeutic action.

The purpose of the action loop is to identify the device’s contribution to carrying out the action. A user may perform the action through the device, as when a surgeon acts on tissue with an instrument. The device may perform the action after a user has selected, configured or started the treatment, as when an infusion pump meters a drug or a radiotherapy system delivers radiation. Or the device may remain on or in the patient and produce its effect through its presence or properties, as with a dressing or an implant. In each case, the device participates in the action, although the degree of user involvement differs.

A device may also decide whether or how to act and then perform the action itself. An automated insulin-delivery system, for example, may calculate a dose and deliver it without a separate user decision. These remain two distinct contributions: selecting the action closes a clinical decision loop; delivering the insulin completes the action in a clinical action loop. Separating them makes it possible to ask whether the evidence demonstrates that the device chose the right action, performed it as intended, produced the intended effect, or caused harm.

2.4 One device, several loops

In practice, the loops alternate: a decision leads to an action, the action changes the patient, and the patient is assessed again. Both loops may sit inside the same device.

An automated external defibrillator records the ECG, analyses the rhythm and decides whether a shock is indicated. That is a decision loop, closed by the device. If indicated, the shock is then delivered: an action loop. One device, two loops, and a different question for each: does it recognise a shockable rhythm, and does the shock work?

The loops can also run continuously, be nested, or occur in sequence. A dialysis machine monitors pressures and conductivity while treating the blood and adjusts treatment accordingly. A radiotherapy system verifies position before delivering a dose; an infusion pump with dose-error reduction software checks the programmed dose before infusion.

This is why the framework examines loops rather than devices as a whole. Each loop needs to be considered separately. In the next section, we use this analysis to determine when a device is being used clinically.

3 “Using” a device clinically

We can now return to the question that led us here: when is a device actually being used as a medical device?

In this guide, we bring forward the idea that “using” a device clinically means completing a loop: a clinical action loop or a clinical decision loop in which the device takes part.

The two loops provide a procedure for determining whether a device is being used clinically. First, identify every clinical decision loop and clinical action loop in which the device takes part. Then, for each loop, identify the device's contribution and follow it through the clinical process. Finally, determine whether the loop is completed, as follows.

A clinical *action* loop is completed when the action in which the device participates is carried out on the patient and produces, or is intended to produce, an effect. A clinical *decision* loop is completed when the information to which the device contributes is used to decide what happens next: obtain more information, act on the patient, or do nothing.

If the process stops before that point, the device may have been operated and its performance may have been evaluated, but the loop has not been completed. Under the framework proposed in this guide, the device has therefore not been used clinically.

Assessing whether a device is used clinically thus involves three steps:

1. Identify all the clinical loops in which the device takes part.
2. Identify the device's contribution to each loop.
3. Determine whether each loop is completed.

We will apply this procedure first to a clinical action loop and then to a clinical decision loop.

Applying the procedure is usually straightforward for clinical action loops. For clinical decision loops, determining whether the loop is completed can be more difficult. Let's see why.

Completion of a clinical action loop is easy to recognise because the device's function requires an interaction with the body, or at least with the patient. Take that interaction away, Figure 2B, and there is no clinical effect. Manufacturers understand this instinc-

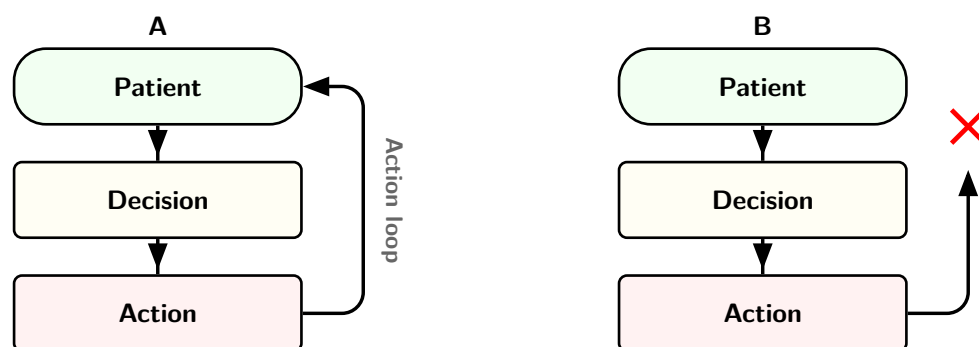


Figure 2: Clinical action loop. **A** Complete loop: the decision is turned into an action on the patient. **B** Incomplete loop: no action is performed on the patient.

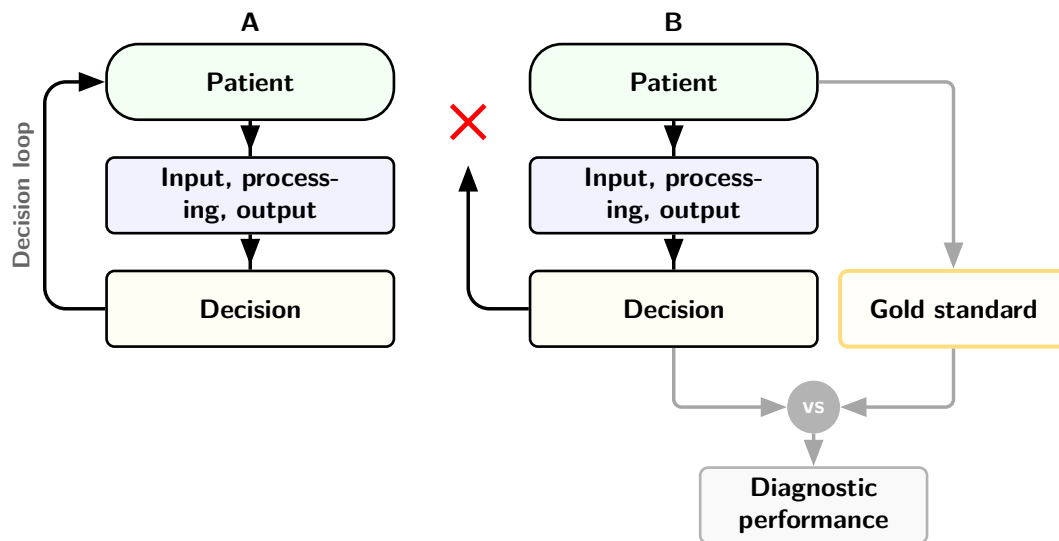


Figure 3: Clinical decision loop. **A** Complete loop. **B** Incomplete loop with gold standard comparison: the decision taken on the output of the device does not reach the patient. The loop does not close, but comparing the decision with a gold standard obtained from the same patient yields the diagnostic performance of the device.

tively. If the cannula is not inserted, no fluid is given. If the radiation is not delivered, there is no therapy. If the dressing is not applied, the wound is not protected. If the physiotherapy plan never reaches the patient, nothing is treated. So the loop is easy to see, and so is the moment it closes: when the device acts on the patient, clinical data is being generated; when it does not, it is not.

Establishing when a device completes a clinical decision loop proves more challenging in practice. Here is why.

Consider the manufacturer of a diagnostic AI that detects acute myocardial infarction from ECG recordings. This manufacturer collects ECG traces from patients, processes them with their device, and provides a myocardial infarction probability score to clinicians. The manufacturer then compares the AI scores against troponin blood test results (the biochemical gold standard for myocardial damage). From this comparison, the manufacturer calculates clinical sensitivity and specificity. The study stops there. In this scenario, the clinical decision that the user made using the device's output is not applied to patient management—that is, it does not inform the next diagnostic or therapeutic steps. This corresponds to the workflow shown in Figure 3.

As in the incomplete clinical action loop, the arrow from the decision to the patient is missing. The loop does not close. And yet the study is not empty. Comparing the output of the AI with the gold standard gives the sensitivity and the specificity of the device, and these are clinical parameters: they say how the device performs on real patients. This is where manufacturers get stuck. The data come from patients, they describe the

clinical performance of the device, but the device has not been used on anybody. Are they clinical data or not?

Under the framework proposed here, they are not clinical data. They are pre-clinical data because the device's output was not used in patient management and the clinical decision loop was not completed. Nevertheless, they can support a clinical performance claim. The fact that a claim concerns clinical performance does not make the evidence used to demonstrate it clinical data.

Conclusion to Part I

Part I has established a procedure for determining whether a device is being used clinically: identify every clinical decision and clinical action loop in which the device takes part, identify the device's contribution to each loop, and determine whether each loop is completed.

This procedure applies to medical devices governed by the MDR and to IVDs governed by the IVDR. Most IVDs participate in a clinical decision loop: a specimen originates from the patient, the device processes it into a result, and that result informs what happens next. In both regulatory frameworks, the questions are the same: what does the device contribute to the clinical process, and is the loop completed?

Loop completion tells us whether the device has been used clinically and therefore whether the resulting information can constitute clinical data. It does not, by itself, determine whether a study is useful or which claim its results can support. As the diagnostic example shows, an incomplete loop may still generate evidence of clinical performance.

In Part II, we develop the framework further, using the loops to distinguish performance, analytical performance for IVDs, clinical performance, clinical benefit, and safety claims, and to determine the evidence needed to support each type of claim.

Meet 4BetterDevices



The principles presented in this guide are embedded in the 4BetterDevices software platform: a suite of connected applications supporting regulatory and evidence work throughout the medical-device lifecycle. The platform turns methods developed through hundreds of submissions, across diverse device classes and in collaboration with manufacturers and notified bodies, into structured, repeatable workflows. It helps teams reduce uncertainty, identify gaps earlier, maintain consistency and prepare review-ready documentation. Training and specialist support complement the software when regulatory judgement is required.