Clinical Evaluation of Software

A regulator’s perspective.

Simon Lidgate
Technical Team Manager, AIMD, Software
2024-05

We will cover …

- Elements of a clinical evaluation
- Medical device, intended purpose, risk and Classification
- State of the art, Performance, Benefit
- Clinical Studies, Challenges & Examples
- Levels of Clinical Evidence
- Equivalence
- When Collection of clinical data is Inappropriate
- Considerations of AI
- Thoughts on PMCF

… in <45 minutes?

Radiotherapy Planning Software

Automated MR image segmentation software, where organs are delineated prior to planning. Clinician is forced to check and modify segmentations prior to use as part of the plan.

Claims:

Improves patient management via reducing planning time.

Clinician Programmer for a pacemaker

Windows - based software which connects to the implanted pacemaker via Bluetooth and is used to adjust the therapy, activate / deactivate features of the pacemaker and download activity and pacing history.

Claims:

Aligned with the claims of the pacemaker.

ICU Patient Monitoring Software

AI-powered software with patient's vital physiological parameters as an input, as well as a video feed of the patient. Provides early warnings of patient deterioration, along with suggestions of treatment.

Claims:

Improves patient outcome and reduces mortality.

(All similarities to real products are entirely accidental)

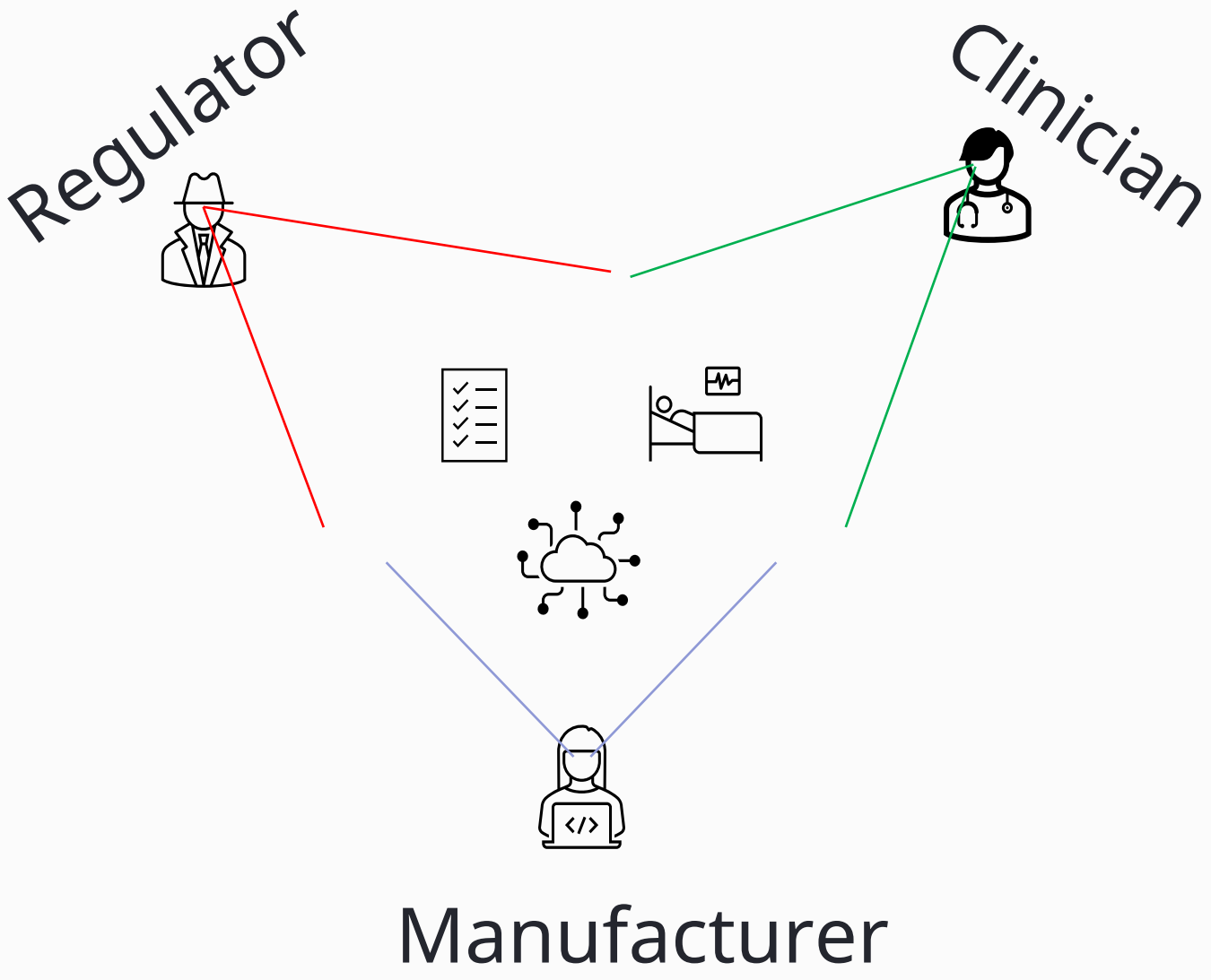
Think about:

- Risks/Performances/Benefits related to the device & claims?
- How would you classify it?
- What is the appropriate route for clinical assessment?
- What would any premarket study look like?
- What does the PMCF plan look like?

Perspective...

The Regulator:
- Performs assessments against Requirements.
- Communicates in terms of Gaps, or Non-Conformities.
- Is very time-limited.
- Will not make assumptions or perform parts of the manufacturer's assessment.

A good Clinical Evaluation:
- Is clear, complete, logically structured, and tells the story.
- Demonstrates compliance to requirements, alignment to guidance
- Is clear about the device & its intended purpose.
- Thoroughly researches the SotA, extracting claims, measurements and endpoints that define it.
- Quantifies side effects, risks, performances and benefits of the subject device.
- Identifies all relevant clinical data to support claims.
- Appraises quality of sources.
- Provides a clear, objective analysis
- Comes to supported conclusions
- Has a plan to collect data on real-world use. (PMCF)



True / False

The MDR has **no specific requirements** on the clinical evaluation of software.

TRUE

General Clinical Evaluation requirements apply. (specific requirements in other areas)

Scoping – Plan, Sota, Claims & Objectives

CEP

- **GSPRs** to address: What needs verification via clinical data?
- **Intended Purpose :** What the device does, and when it should be used.
- Expected clinical **Benefits + Risks :** Outline of clinical performances, risks and side effects associated with the device
- **Methods** of Examination: How these will be measured
- Indicative **Parameters :** The Measures to be used
- Clinical Development Plan : Timeline of collection of clinical evidence.

SotA

- **Using:** literature search & other sources
- **Establish:** disease type and state (e.g.), symptoms, and desired outcomes.
- **Identify:** alternative treatments, similar devices, relevant standards, guidance , etc.
- **Compare:** benefits, risks, features, side effects, etc.
- **Estimate:** Appropriate Level of Clinical Evidence

Claims / Objectives

- What does it **do**?
- What are the expected **benefits and performances**?
- What are the **risks** and the acceptable level of **side effects**?
- What are the **conditions and limitations** under which this can be achieved?
- Describe using quantitative measures and claims, aligned with the state of the art.

Identifying Data – Equivalence, Literature, Studies

Equivalence

Allows clinical data on equivalent device to support current device.
Proof of Technical, Biological, and Clinical similarity, such that there is no clinically significant difference
ALL differences identified; adequate scientific justifications provided.
Requires “sufficient access” to data on equivalent device.

Literature

Identify published literature on subject or equivalent.

Use multiple sources, Justify as appropriate

Identify all relevant literature. Reject poor quality articles and duplicates presenting same data.

Demonstrate integrity of search, eliminate possibility of bias. Provide transparency: access to literature and reasons for rejection.

Clinical Studies

Protocols: Study Design, ISO 14155 alignment, Devices, Objectives and Endpoints, Population, Observations, Statistical Power, Consideration of attrition, follow-up, etc.

Reports: Critical assessment against endpoints, claims and objectives, Deviations, Identification of gaps, Requirements for further study.

Evaluating Data – Appraisal, Analysis, Conclusions

Appraisal

Using Appropriate methods, defined and clear objectives, assess
Relevance to device, Intended Purpose, claims and objectives,
Disclosure of methods, products, population, outcomes, side effects, adverse events, deviations.
Power of conclusions / Sufficient number of observations
Declarations of Interest.

Analysis

Against GSPRs, data available on State of the Art, claims and objectives made on device,
Identification of Gaps / Weaknesses in Clinical data, Requirements for further study.
Confirmation & update of Risk analysis

Conclusions

Sufficiency of Clinical Data
Alignment against state of the art
Support of GSPRs, Claims and Objectives, Claims of State of the Art, lifetime of device and/or treatment.
Requirement of PMCF.

Clinical Evaluation : Plans for follow-up

Post Market Surveillance (Ann III)

Collect and analysis post market data on:

- Serious incidents and Field Safety Corrective Actions
- non-serious incidents and undesirable side-effects;
- trend reporting;
- relevant specialist or technical literature, databases and/or registers;
- feedbacks and complaints, provided by users, distributors and importers; and
- publicly available information about similar medical devices.

Post Market Clinical Follow-up

General and specific methods to :

- Demonstrate safety and performance of devices supported by equivalence.
- Address gaps in clinical data.
- confirm the safety / performance throughout lifetime,
- identify/monitor side-effects and contraindications
- identify and analyse emergent risks
- ensure continued acceptability of the benefit-risk ratio
- identify systematic misuse / off-label use

True/False (1)

All software used in a **clinical environment** is considered a Medical Device.

FALSE

Classification of Software as a MD is dependent on its Intended Purpose.

What is a medical device …

Article 2(1) and Recitals(19)

- Software must have a specific medical “Intended Purpose” to be considered a medical device. The manufacturer must make clinical performance claims related to that intended purpose.
- Software that is intended to be used in combination with a medical device, IS a medical device.
- Software designed for “lifestyle” and “wellbeing” purposes is NOT considered a medical device.
- It does not matter where the software is deployed or how it is connected, to consider it a medical device.
- Software that interacts with general purpose devices (e.g. PC/Windows/the internet) must be appropriately designed to mitigate risks associated with the.

Article 2 (1) ‘medical device’ any … software … to be used, **alone or in combination**, for human beings for one or more of the following specific medical purposes:

— diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease,

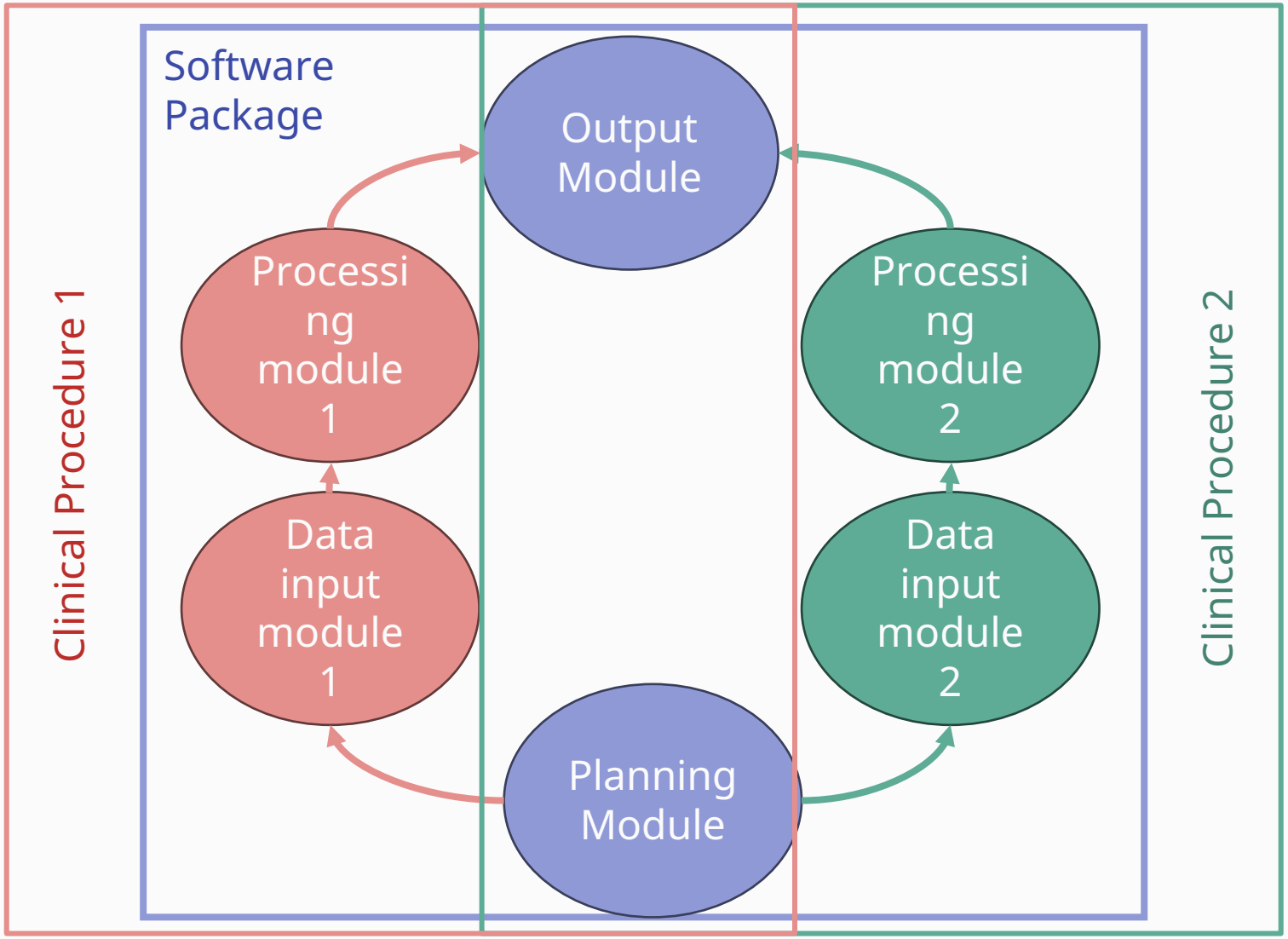
— diagnosis, monitoring, treatment, alleviation of, or compensation for, an injury or disability,

— investigation, replacement or modification of the anatomy or of a physiological or pathological process or state,

… examination of specimens derived from the human body, including organ, blood and tissue donations, … control or support of conception;

Recitals (19) It is necessary to clarify that software in its own right, **when specifically intended** by the manufacturer to be used for one or more of the medical purposes set out in the definition of a medical device, qualifies as a medical device**, while software for general purposes**, even when used in a healthcare setting, or software intended for **life-style and well-being purposes is not** a medical device. The qualification of software, either as a device or an accessory, is **independent of the software's location** or the type of interconnection between the software and a device.

How many medical devices?



Intended Purpose

MDR Article 2(12) and Meddev 2.7.1 Rev 4

- Intended purpose is defined by everything that is written about the device, in the labelling, CER, sales materials, website, etc.
- It includes everything about how the device may or may not be used, including indication, disease (incl. stage and severity), populations, user, contraindications, precautions, etc.
- The Intended Purpose must be fully considered within the clinical evaluation.
- Therefore, it must be specific and pass the “watertight” test… open-ended statements cannot be tested clinically.
- Software often has multiple specific intended purposes (e.g. involved in the treatment or detection of multiple diseases, multiple areas in the body, etc.
- Software often replaces or augments an established procedure, for which there is clinical data.
- How do you prove that the software, in each case, presents a benefit that outweighs the risk?

Requirements:

Article 2(12) Intended Purpose 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 and as specified by the manufacturer in the clinical evaluation;

Meddev 2.7.1 Rev 4: **intended purpose of the device**: exact medical indications (if applicable),- name of disease or condition/ clinical form, stage, severity/ symptoms or aspects to be,treated, managed or diagnosed, patient populations (adults / children / infants, other aspects), intended user (use by health care professional / lay person), …, contraindications, precautions required by the manufacturer, …, other aspects

Risks and Foreseeable Misuse:

GSPR (3) : … In carrying out the risk management manufacturers shall:

… (c) estimate and evaluate the risks associated with, and occurring during, the intended use and during reasonably foreseeable misuse;

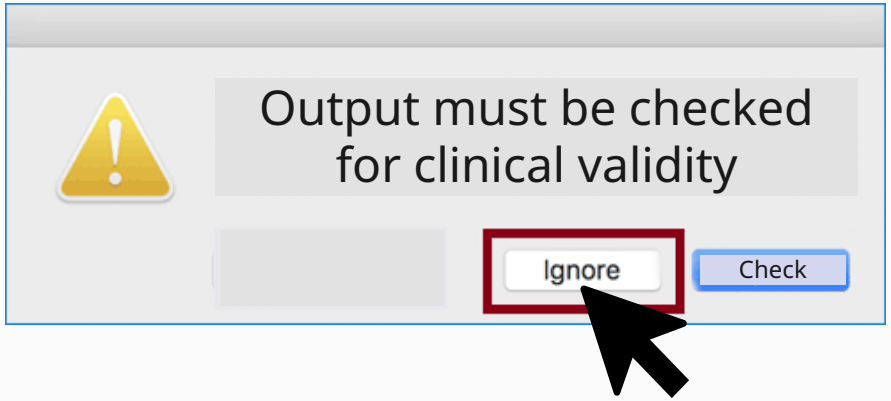
(d) eliminate or control the risks referred to in point (c) in accordance with the requirements of Section 4;

4) … in the following order

(4a) Eliminate or reduce risks AFAP through safe design and manufacture

(4b) … adequate protection measures, including alarms…

(4c) … provide information for safety …. training to users.

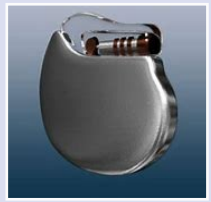


Medical Device Software and Clinical Evaluation Scope

MDCG 2020-1

Software that is a medical device in it’s own right a medical device (i.e. standalone SAMD) requires a clinical evaluation on it alone.

SAMD with its own intended purpose that works in combination with another medical device, requires a clinical evaluation on it combined with the other device.



Software that drives or influences another device, but has no intended purpose of its own, requires a clinical evaluation of the driven device incorporating that software.

True/False (2)

The **Manufacturer** determines and justifies the classification of their medical device.

TRUE

The NB evaluates the rationale as it is presented.

Classification MDD -> MDR

MDD Annex IX

2.3. Software, which drives a device or influences the use of a device, falls automatically in the same class….

3.2. Rule 10

Active devices intended for diagnosis are in Class IIa:

— if they are intended to allow direct diagnosis or monitoring of vital physiological processes, **unless they are specifically intended for monitoring of vital physiological parameters, where the nature of variations is such that it could result in immediate danger to the patient,** for instance variations in cardiac performance, respiration, activity of CNS in which case they are in Class IIb.

Active devices intended to emit ionizing radiation and intended for diagnostic and therapeutic interventional radiology including devices which control or monitor such devices, or which directly influence their performance, are in Class IIb.

3.3. Rule 12

All other active devices are in Class I.

MDR Annex VIII

6.3. Rule 11

Software **intended** to provide information which **is used to take decisions with diagnosis or therapeutic purposes** is classified as **class IIa**, except if such decisions have an impact that **may cause**:

— death or an **irreversible deterioration of a person's state of health, in which case it is in class III;** or

— a serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as class IIb.

Software **intended** to **monitor physiological processes** is classified as **class IIa**, except if it is intended for monitoring of **vital physiological parameters**, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is **classified as class IIb**.

All other software is classified as class I.

Classification

MDR Annex VIII, MDCG 2019-11,

- Obviously, Manufacturers are responsible for the correct classification of their device.
- Classification impacts decisions made during the clinical evaluation process.
- The key word is “intended” (e.g. in rule 11), i.e. Intended Purpose
- However, the manufacturer must realistically evaluate the patient impact, including in “reasonably foreseeable misuse”, and be able to justify their final decision.
- All rules must be considered, and the highest classification applied. Be wary that software may be subject to other rules e.g. as accessory of an AIMD (Rule 8)

6.3. Rule 11

Software **intended** to provide information which **is used to take decisions with diagnosis or therapeutic purposes** is classified as **class IIa**, except if such decisions have an impact that **may cause**:

— death or an **irreversible deterioration of a person's state of health, in which case it is in class III;** or

— a serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as class IIb.

Software **intended** to **monitor physiological processes** is classified as **class IIa**, except if it is intended for monitoring of **vital physiological parameters**, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is **classified as class IIb**.

All other software is classified as class I.

State of the Art (SOTA) and Claims

State of the Art

- State of the art is not defined within the MDR,
- MDCG 2020-1 says: *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"*

- State of the Art is defined by:
  - Your direct competitors
  - Similar devices
  - Alternative devices, treatments, or methods

SW devices often reach into new territory where there are few or no competitor devices, and the SW replaces a previously human driven methodology.

(a full history of Medical Devices is not required to define the State of the Art)

Claims, Objectives, Benefit, Risk, Side Effects

MUST be measureable

MUST be defined

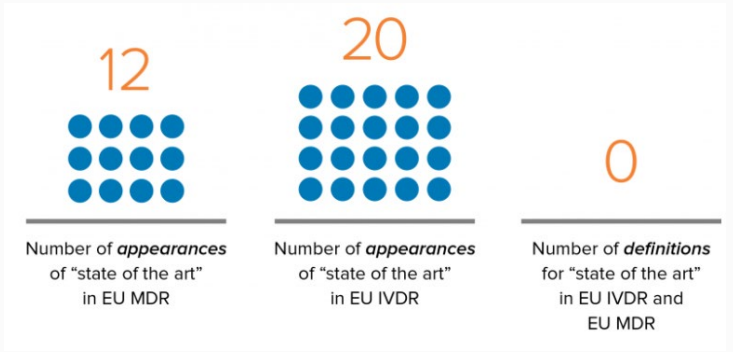
MUST align with state of the art.

MUST have an overall clinical benefit * that outweighs the risk

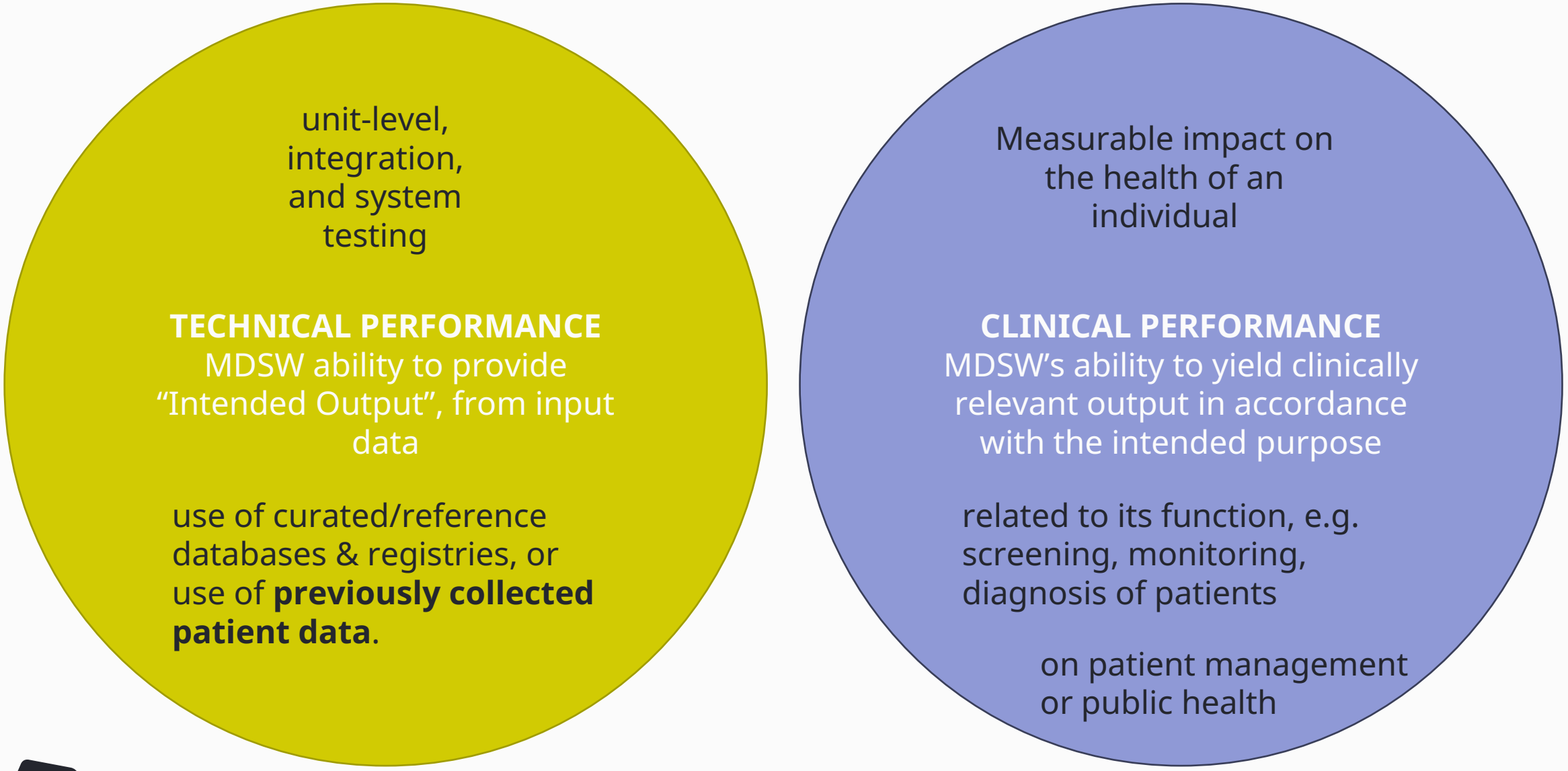
MUST consider the risks of "reasonably foreseeable misuse".

Disclaimers made to reduce risk are often untrue, or risk reduction is overestimated (disclaimers are likely to be ignored!)

* Benefit may be indirect.



MDCG 2020-1 Technical / Clinical Performance



Performance -> Benefit

Article 2(52):

> '**Clinical performance**' means the **ability of a device**, resulting from any direct or indirect medical effects which stem from its technical or functional characteristics, including diagnostic characteristics, **to achieve its intended purpose as claimed by the manufacturer, thereby leading to a clinical benefit** for patients, when used as intended by the manufacturer;

Article 2(53):

> "**Clinical benefit**' means the **positive impact of a device on the health of an individual**, **expressed in terms of a meaningful, measurable, patient-relevant clinical outcome**(s), including outcome(s) related to diagnosis, or a positive impact on patient management or public health;

Clinical Benefit may be considered direct or indirect. e.g.:
- The ability of a hip implant to allow a patient to walk -> direct benefit.
- The ability of a guide to facilitate correct implantation -> indirect benefit

Clinical Benefit may also be Broad or Generic, e.g.:
- the ability of a scalpel to resect, allowing a procedure to be performed.
- the ability of a scanner to image, facilitating other procedures.

Where possible, clinical benefit shall be directly measured. Elsewhere, it shall be clearly justified that a measured performance will deliver the expected benefit.

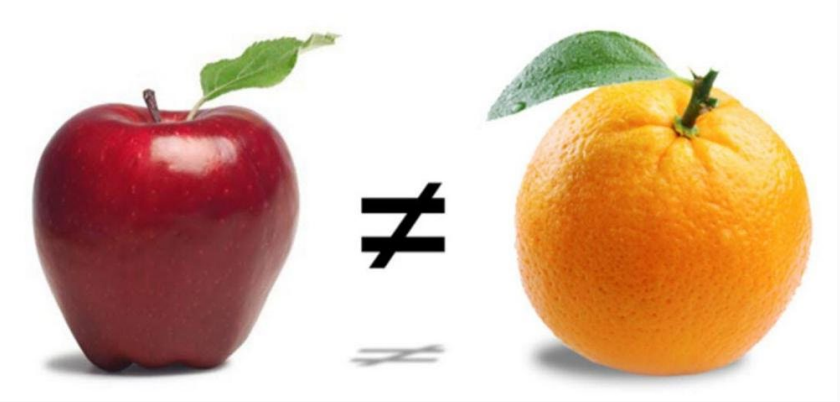
clinical-evaluation-under-eu-mdr.pdf (bsigroup.com)

Equivalence…

(MDR Annex XIV(3)3, Article 61(4), MDCG 2020-5,

MDCG 2023-7)

- Must be between the **current version** of the device and **the exact device which has clinical data**. Includes between different models or generations of the same device, were necessary.
- Must provide a thorough analysis of ANY technical, biological, or clinical differences between the devices.
- ANY differences must have scientific justifications not to impact the safety and performance profile.
- Also beware of the additional requirements for Class III or implantable devices, which restricts equivalence to devices based on modifications of a device from the same manufacturer



MDR Annex XIV 3. A clinical evaluation may be based on clinical data relating to a device for which equivalence to the device in question can be demonstrated. The following technical, biological and clinical characteristics shall be taken into consideration for the demonstration …

Article 61(4) the device has been designed by modifications of a device already marketed by the same manufacturer

Equivalence …

Technical

- **software algorithms** - different or modified algorithms will behave differently, such that safety and performance may not be the same as observed with the unmodified algorithm.

- **usability** – different interfaces, prompts, warnings, deployment, presentation, etc., may alter the perception of clinical output, and may alter the clinical safety / performance

Clinical

- **Intended Purpose -** Any differences in Intended Purpose may result in gaps or uncertainties in the clinical data when applied to the subject device.

MDR Annex XIV 3.

— Technical: the device is of similar design; is used under similar conditions of use; has similar specifications and properties including physicochemical properties such as intensity of energy, tensile strength, viscosity, surface characteristics, wavelength and software algorithms; uses similar deployment methods, where relevant; has similar principles of operation and critical performance requirements;

— Biological: the device uses the same materials or substances in contact with the same human tissues or body fluids for a similar kind and duration of contact and similar release characteristics of substances, including degradation products and leachables;

— Clinical: the device is used for the same clinical condition or purpose, … has similar relevant critical performance in view of the expected clinical effect for a specific intended purpose.

Equivalence …

Scientific Justification based on solid analysis (e.g. via bench testing, usability analysis, or clinical evaluation) shall verify that ANY difference will not have a clinical impact.

Sufficient Access to the data of the equivalent device is naturally required to perform this analysis. Justifications as to sufficient access are expected (and will be tested!)

PMCF Plan is generally expected to address follow-up on devices brought to market via equivalence.

PMCF studies are an explicit requirement for Class III or implantable devices.

MDR Annex XIV 3 The characteristics listed in the first paragraph shall be similar to the extent that there would be no clinically significant difference in the safety and clinical performance of the device. Considerations of equivalence shall be based on proper scientific justification. It shall be clearly demonstrated that manufacturers have sufficient levels of access to the data relating to devices with which they are claiming equivalence in order to justify their claims of equivalence.

MDR Article 61(4) In this case, the notified body shall check that the **PMCF plan is appropriate and includes post market studies** to demonstrate the safety and performance of the device.

True / False (4)

The offline processing of clinical data by a medical device will create clinical data, e.g. via comparing the automatic segmentation of features in an MRI scan, with previously collected manual segmentations collected as part of a clinical workflow.

FALSE

‘Clinical Data’ has a very specific meaning …

What is clinical data?

Article 2(48)

- Specifically, data collected from end-to-end clinical use of the device in its intended environment
- May come from clinical investigations, studies, surveys, registries, and other PMS and PMCF.
- Importantly the output of running previously collected clinical data through a (e.g.) new algorithm on a bench, is NOT clinical data, even if a previous version of the device was used to collect data.
- In this case, it may be considered scientific verification and validation, perhaps as part of an equivalence analysis …

- Article 2(48) 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;

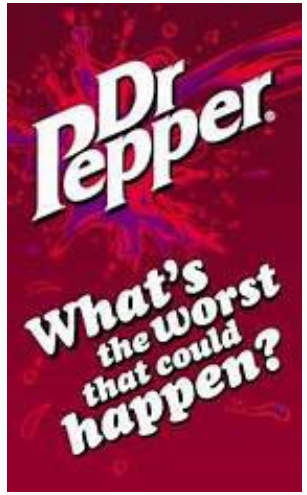
Article 61(1) Level of Clinical Evidence

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

Generally Clinical Evidence may be considered sufficient when:

- Claims are duly substantiated
- Risks are clearly understood and characterised.
- Clear comparisons can be made with the State of the Art

The higher the risk (including misuse), the greater the weight of the evidence.



MDCG 2020-6 Levels of Clinical Evidence

- The manufacturer must carefully consider the level of clinical evidence necessary to support device **intended purpose**, when considering **risk**, and evidence available on the **state of the art**.

- Although written specifically for legacy devices, guidance on quality of evidence is found in MDCG 2020-6.

- With a new device, given no historical data to support safety, new devices should aim to exceed expectations set in this guide.

Rank	Types of clinical data and evidence	Considerations / comments
1	Results of high quality[62] clinical investigations covering all device variants, indications, patient populations, duration of treatment effect, etc	This may not feasible or necessary for certain well-established devices with broad indications (eg Class IIb legacy sutures, which could be used in every conceivable patient population)
2	Results of high quality clinical investigations with some gaps	Gaps must be justified / addressed with other evidence in line with an appropriate risk assessment, and clinical safety, performance, benefit and device claims. Assuming the gaps can be justified, there should be an appropriate PMCF plan to address residual risks. Otherwise, manufacturers shall narrow the intended purpose of the device until sufficient clinical data has also been generated.
3	Outcomes from high quality clinical data collection systems such as registries[63]	Is there sufficient evidence of the quality of the data collected by the registry[64, 65]? Are the devices adequately represented? Are the data appropriately stratified? Are the endpoints appropriate to the safety, performances and endpoints identified in the clinical evaluation plan?
4	Outcomes from studies with potential methodological flaws but where data can still be quantified and acceptability justified[66]	Many literature sources fall into this category, due to limitations such as missing information, publication bias, time lag bias, etc. This applies equally to publications in the peer-reviewed scientific literature. However, for legacy devices

MDCG 2020-6 Appendix III

MDR Annex XV: Clinical Studies

- Recognised Ethical Principles (e.g. ISO 14155)
- An appropriate plan, reflecting the latest scientific and technical knowledge, designed to confirm or refute the manufacturer’s claims.
- Procedures appropriate to the device under investigation.
- Sufficient number of intended users, representative of normal conditions of use.
- All technical and functional features of the device involving safety and performance, and expected outcomes, shall be appropriately addressed.
- Endpoints that address benefit/performance claims and safety.
- The report shall contain a critical evaluation and include any negative findings.

Percieved quality of study…

Randomized controlled trials - Challenges

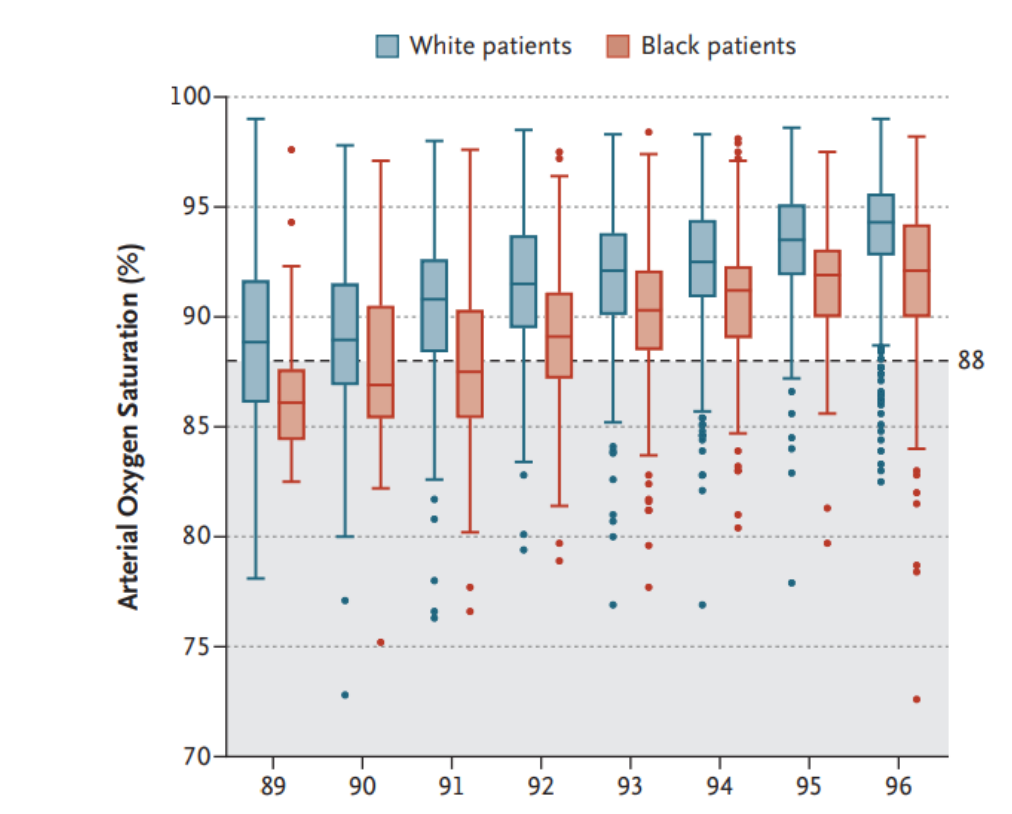
- Narrow trial populations or insufficient definition or representation of subgroups.
- Population may be unsuitable for randomization.
- Unintentional bias caused by study design.
- Controlled environment, vs “Real World” use.
- Unsuitable to answer many related questions.
- Ethical questions.

Sub-Populations and Bias

The manufacturer must carefully consider subpopulations and design a trial to fit.

NEJM letter 16 December 2020:

- Study performed during Covid-19 compared arterial saturation and pulse oximetry between black and white patients.
- Study found black population had nearly 3 times the frequency of occult hypoxemia, not detected by pulse oximetry.
- Highlighting “an ongoing need to understand and correct racial bias in pulse oximetry and other forms of medical technology.”
- Why was this not known before?



No. of Paired Measurements								
White patients	92	178	231	314	438	556	653	817
Black patients	20	52	59	83	127	126	188	225

[Racial Bias in Pulse Oximetry Measurement (nejm.org)]

The INFANT Trial – Computerised interpretation of fetal heart rate during labour (RCT)

Sometimes performances do not translate to benefit.

Study Design:

- Conducted between 2010 – 2013
- Unmasked Randomized Trial, 47062 women randomly assigned.
- Continuous Fetal Monitoring During Labour
- Comparing Guardian with Guardian + Infant
  - Guardian – Electronic Information Capture(Cardiac + other) System + display (local or remote)
  - INFANT – Interpretation of signals and Blue/Yellow/Red alert depending on severity. Does NOT provide treatment advice.
- Endpoints
  - at 24/48h Poor Neonatal Outcome
  - At 2 years, developmental assessment.

Outcome:

INFANT algorithm **worked**, but:

- no difference in the incidence of poor neonatal outcome between the groups
- At 2 years, no significant differences were noted in terms of developmental assessment

Conclusion:

Use of computerised interpretation of cardiotocographs in women who have continuous electronic fetal monitoring in labour **does not** improve clinical outcomes for mothers or babies.

The INFANT Collaborative Group, 2017: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)30568-8/fulltext

The INFANT trial, a flawed study?

Keith, 2017:
https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)30714-6/fulltext

- A lack of study arm separation may have lead to cross-over effects (where the INFANT algorithm may have reinforced learning in staff) -> Separate the arms?

- Use of Guardian acted as a “sophisticated surveillance system” -> Instead compared INFANT/Guardian with conventional monitoring.

Belfort & Clark, 2017:
https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)30762-6/fulltext

- actions taken when distress detected were not specified.

- random assignment limited demonstration of intrinsic value of cardiotocographic technology

- Apparent dismissal of value of the tool because the algorithm was not beneficial in the trial.

My thoughts:

Increased separation could lead to observations related to differences in practise, rather than with device

Would we then measure benefit from Infant or Guardian?

Specifying treatment decisions could significantly increase risk associated with tool

Randomization attempts to eliminate impact of other factors. Large study vs a more specific population?

Study was a specific test, for a specific intended purpose. Is the intended purpose wrong?

The response:

Brocklehurst at al, 2017 (in response to Keith, Belfort and Clark):

https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)31594-5/fulltext#%20

Study was a comparison of decision support with no decision support, with/without Guardian would be a different question.

Acknowledged that there may have been some impact to learning, but overall, the trial was testing whether tool could improve on appropriate training.

Specifying actions to take was not appropriate, given correct action is disputed.

Incorporating other variables may improve outcomes, but it could also be the case that fetal heart rate monitoring is not a good test of wellbeing.

When Collection of clinical data is **inappropriate**.

Sometimes it can be justified that clinical data is NOT required to support a medical device, but this should be viewed with caution.:

Not applicable to class III and IIb implantables

When is it not **appropriate**? At least:

- Indirect (not measurable) clinical benefit as per MDCG 2020-6
- No clinical-specific risks
- Nothing else to learn, or a study cannot be performed.

What is due **substantiation**? At least:

- The clear justification why it is not appropriate to collect clinical data
- Clear demonstration that all performances and risks fully addressed via preclinical testing.
- A clear plan for active collection in post-market phase, per Annex XIV Part B. *

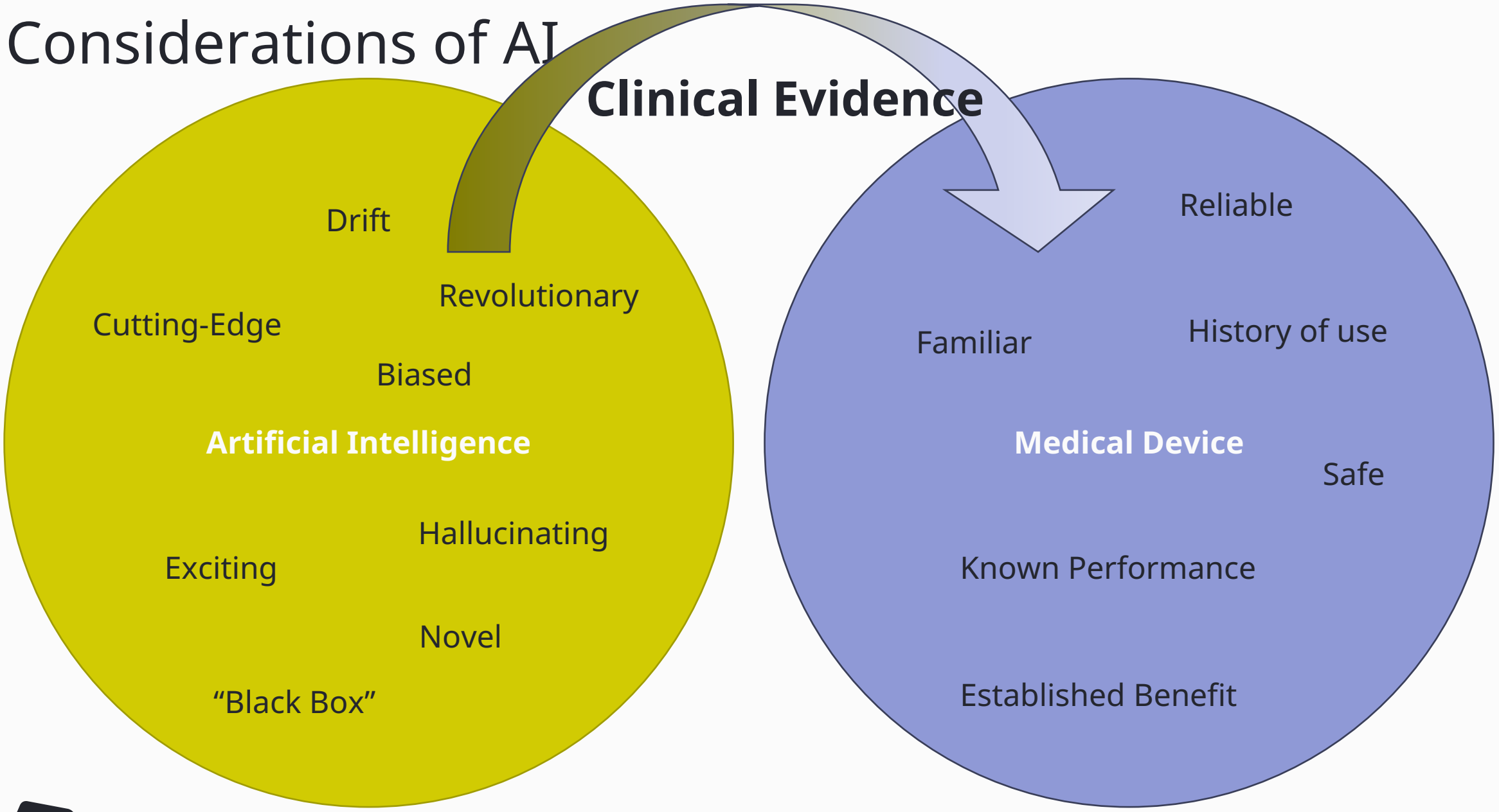
Note: A CER is required. SotA analysis should support (lack of) claims, and application via Article 61(10).

* PMCF should aim to understand the impact of the device in real world use and will generate clinical data. Therefore Article 61(10) is likely not to continue to be applicable after market release.

Article 61(10) *Without prejudice to paragraph 4, where the demonstration of conformity with general safety and performance requirements based on clinical data is not deemed appropriate, adequate justification for any such exception shall be given based on the results of the manufacturer's risk management and on consideration of the specifics of the interaction between the device and the human body, the clinical performance intended and the claims of the manufacturer.*

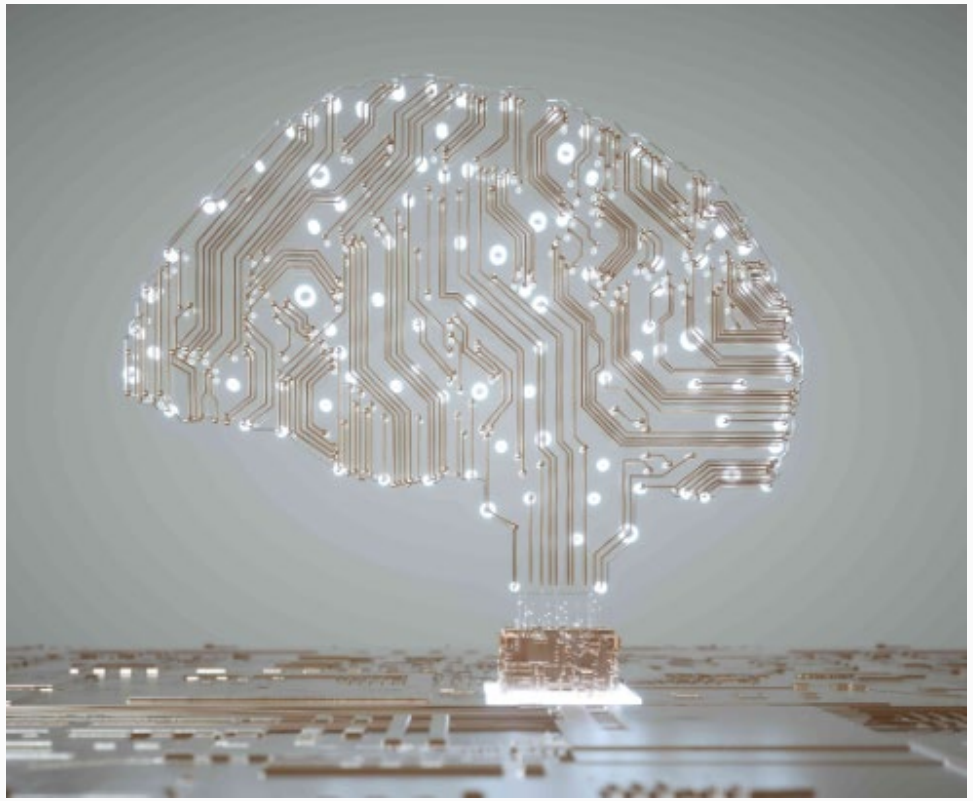
In such a case, the manufacturer shall duly substantiate in the technical documentation referred to in Annex II why it considers a demonstration of conformity with general safety and performance requirements that is based on the results of non-clinical testing methods alone, including performance evaluation, bench testing and pre- clinical evaluation, to be adequate.

Considerations of AI



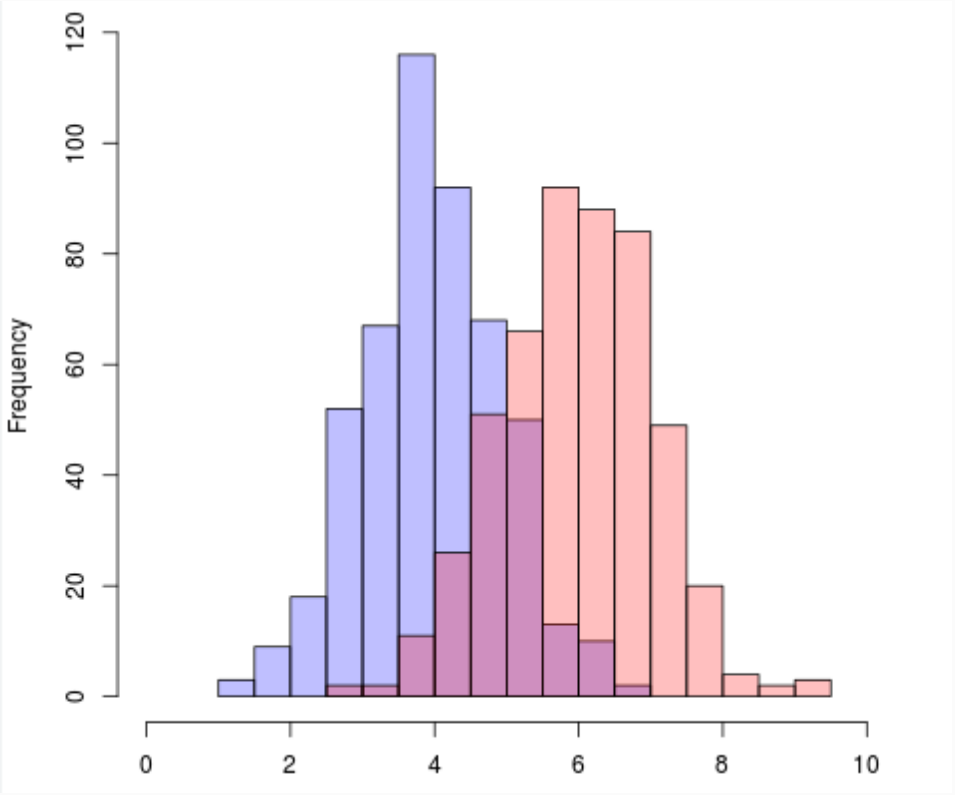
Considerations of AI

- Software is starting to make recommendations for prognosis, diagnosis, treatment, or discharge of patients.
- It is extremely important to consider misuse, or over-reliance of output, as well as its accuracy.
- Rates of algorithm failure or success are not sufficient to measure performance or safety of these devices.



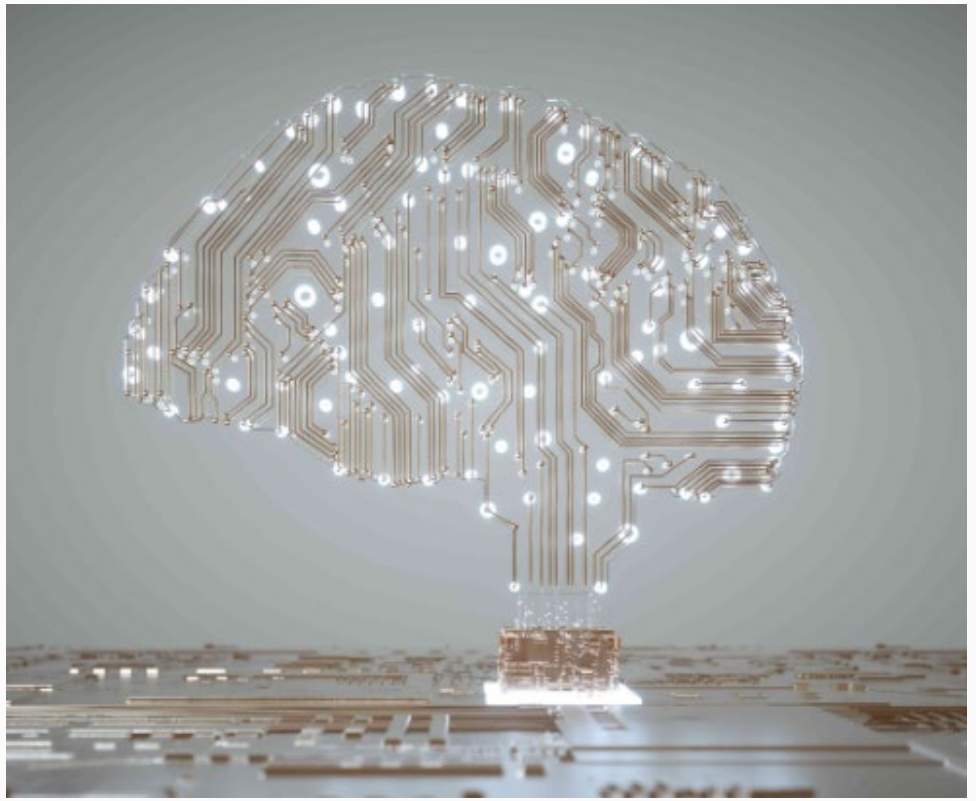
Considerations of AI

- Distributional shift - the mismatch between the training data distribution and real-world data distribution.
- Model Decay – where accuracy of models may decay due to Distributional Shift over time
- Real-world performance as the real-world inputs are of sufficient complexity that the behavior of the system is not simple to predict in advance, or the methods require comparison to human qualities WHICH MAY BE THEMSELVES HIGHLY VARIABLE.
- Whilst (obviously) appropriate premarket clinical evaluation is required, there **will** be many questions for the **post-market phase**.



Opportunities in PMCF? Final thoughts…

- There are clearly a number of challenges with AI and decision-making software, which will require unique solutions.
- Collection of “Real World” data via PMCF will undoubtedly be a big part of addressing these challenges
- SaMD is changing the type, amount, and quality of data that can be collected in the real world.
- This is starting to be recognised as a viable source of data by NBs and organisations, with caveats!



Radiotherapy Planning Software

Automated MR image segmentation software, where organs are delineated prior to planning. Clinician is forced to check and modify segmentations prior to use as part of the plan.

Claims:

Improves patient management via reducing planning time.

Clinician Programmer for a pacemaker

Windows - based software which connects to the implanted pacemaker via Bluetooth and is used to adjust the therapy, activate / deactivate features of the pacemaker and download activity and pacing history.

Claims:

Aligned with the claims of the pacemaker.

ICU Patient Monitoring Software

AI-powered software with patient's vital physiological parameters as an input, as well as a video feed of the patient. Provides early warnings of patient deterioration, along with suggestions of treatment.

Claims:

Improves patient outcome and reduces mortality.

(All similarities to real products are entirely accidental)

Think about:

- Risks/Performances/Benefits related to the device & claims?
- How would you classify it?
- What is the appropriate route for clinical assessment?
- What would any premarket study look like?
- What does the PMCF plan look like?

Questions?

BSI Group

389 Chiswick High Road

London, W4 4AL

+44 345 080 9000

bsigroup.com

Simon Lidgate, AIMD/SaMD, BSI
Simon.Lidgate@bsigroup.com