EU MDR – Regulation (EU) 2017/745 - Level of clinical evidence and what sufficient clinical evidence means

Observation 1 - This is the second installment of my series on medical device clinical evaluation. I suggest reading the first part (How to perform a clinical evaluation of medical devices – Part 1 – Overview and sample of activities - http://www.medicaldevice.expert/europe/european-commission/medical-device-regulation/how-to-perform-a-clinical-evaluation-of-medical-devices-part-1-overview-and-sample-of-activities/0 to have a better understanding of the concepts so the understanding of this part is easier).

Introduction

The EU MDR – Regulation (EU) 2017/745 was created with a strong basis on the need for clinical evidence, and formal requirements are detailed in Article 61.

Article 61 - Clinical evaluation

1. Confirmation of conformity with relevant general safety and performance requirements set out in Annex I under the normal conditions of the intended use of the device, and the evaluation of the undesirable side-effects and of the acceptability of the benefit-risk- ratio referred to in Sections 1 and 8 of Annex I, shall be based on **clinical data providing sufficient clinical evidence**, including where applicable relevant data as referred to in Annex III.

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.

EU MDR – Regulation (EU) 2017/745

CHAPTER VI. CLINICAL EVALUATION AND CLINICAL INVESTIGATIONS

Some questions arise from the requirements, in particular:
- How do I specify and justify the level of clinical evidence necessary to demonstrate conformity with the relevant GSPRs?

- Based on the specification and justification of the level of clinical evidence, what is sufficient clinical evidence?

To understand and answer these questions, it's necessary to understand what evidence-based practice (EBP) is, and how it applies to the medical device clinical evaluation process (and possible other processes).

What is evidence-based practice (EBP)?

Evidence-based practice has been formally introduced some decades ago (Evidence-based medicine - A new approach to teaching the practice of medicine - 1992) as the term "evidence-based medicine", although it was originally championed, among others, by Archie Cochrane in his influential 1972 book "Effectiveness and Efficiency: Random Reflections on Health Services". Even before that, some medical practitioners were already using the concepts of what turned into evidence-based medicine later (for example, by evaluating the effectiveness of bloodletting).

Evidence-based practice can be seen, in a high level view, as a philosophical approach that is in opposition of the "traditional" approach to decision-making, which usually focus on intuitive, unsystematic, "the way it was always done" activities which usually lack several of the characteristics required or expected by a more scientific approach (such as replication). Evidence-based practice, on the other hand, requires a formal approach to decision-making that relies both on the expertise of the practitioner and the strength of the research already performed on the subject, between other components.

One often cited definition of evidence-based practice (as evidence-based medicine) is from an Editorial in BMJ from *Sackett et al*:

> **Evidence based medicine**
>
> Evidence based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients. The practice of evidence based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research.
>
> *Sackett et al*, BMJ, 1996
>
> Evidence based medicine: what it is and what it isn't

In 2005, *Dawes et al* proposed some definitions for evidence-based practice (,):

> **Evidence-based practice**
>
> Evidence-Based Practice (EBP) requires that decisions about health care are based on the best available, current, valid and relevant evidence. These decisions should be made by those receiving care, informed by the tacit and explicit knowledge of those providing care, within the context of available resources.
>
> *Dawes et al*, BMC Medical Education, 2005
>
> Sicily statement on evidence-based practice

And in 2009, *Satterfield et al* proposed a revised EBP model (Toward a Transdisciplinary Model of Evidence-Based Practice, The Milbank Quarterly), as can be seen in Figure 1 below:

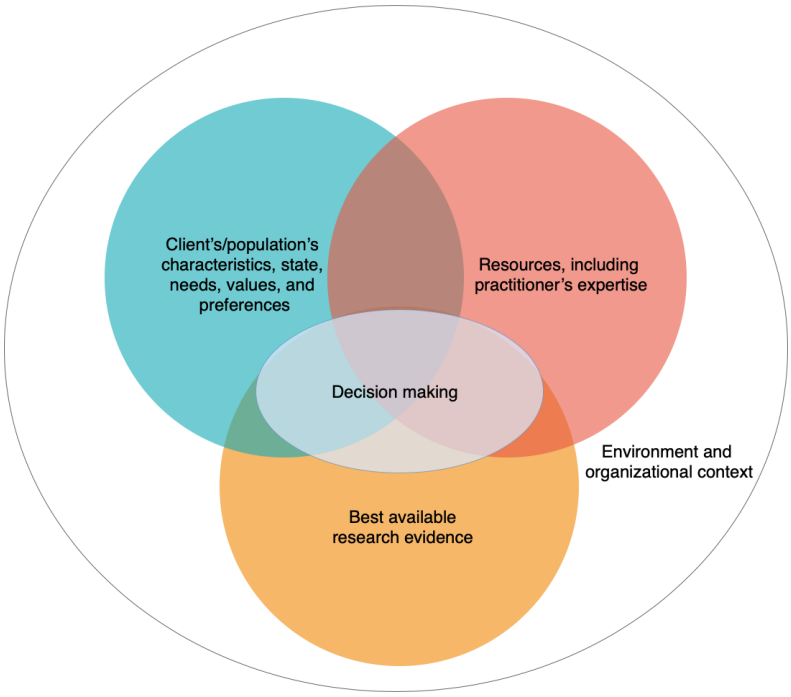

Figure 1 - Adapted revised EBP model (from *Satterfield et al, 2009)*

The evidence-based practice process is usually described as a 5-step process:

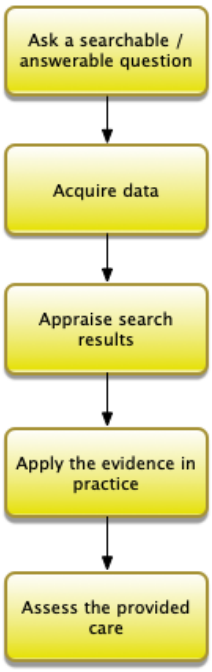

Figure 2 - 5-step evidence-based practice process

This process is (with some modifications) basically what is shown as the medical device clinical evaluation process in MEDDEV 2.7.1 Rev 4, as can be seen in Figure 3 below:

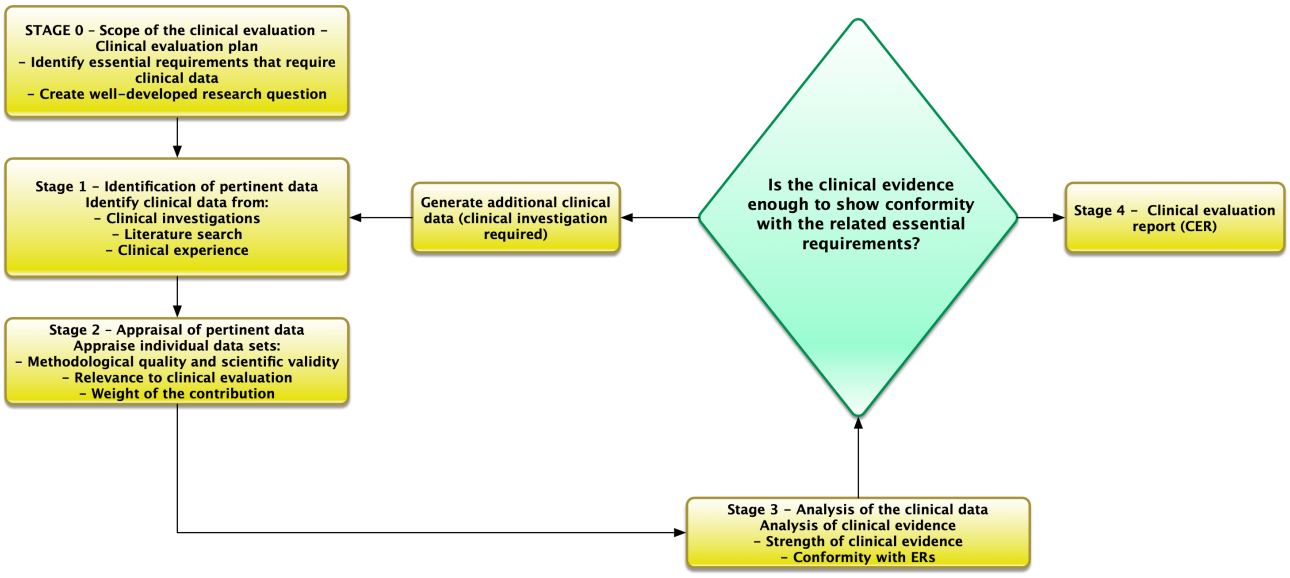

Figure 3 - Overview of medical device clinical evaluation, from MEDDEV 2.7.1 Rev 4

Question 1 - How do I specify and justify the level of clinical evidence necessary to demonstrate conformity with the relevant GSPRs?

With the general understanding of what EBP is, and it's relationship to medical device clinical evaluation, we can then answer the first question.

Level of evidence, in a general sense, "is a heuristic used to rank the relative strength of results obtained from scientific research (Wikipedia). The concept of levels of evidence were originally introduced, in healthcare, in a report by the Canadian Task Force on the Periodic Health Examination in 1979 (The Levels of Evidence and their role in Evidence-Based Medicine, *Burns et al*, Plast Reconstr Surg. 2011 and Canadian Task Force on the Periodic Health Examination, Can Med Assoc J 1979).

Level of evidence is, thus, a system of rating evidence - the higher the rating, the best the evidence could be used as one of the aspects of grading recommendations.

Some definitions (from GRADE)

Quality of evidence indicates the extent to which we can be confident that an estimate of effect is correct.

Strength of a recommendation indicates the extent to which we can be confident that adherence to the recommendation will do more good than harm.

Oxman et al, BMJ, 2004

Grading quality of evidence and strength of recommendations

In practice, the level of clinical evidence (and the related strength of recommendation) is part of the appraisal step in EBP (which is the same as Step 2 - Appraisal of pertinent data, of a medical device clinical evaluation).

The appraisal step "is a systematic method of evaluating the strengths and limitations of a research study, as well as its applicability to practice (Making sense of the quality of evidence (Wilkins, Canadian Oncology NursIng Journal, 2016). Thus, it's specifically related to the evaluation of methodological quality and scientific validity under the appraisal of individual data sets under Step 2 of a medical device clinical evaluation (see Figure 3).

(It's also important to remember that the critical appraisal should use critical appraisal worksheets for each type of study being appraised).

There are more than 100 frameworks (grading systems) that can be used to rate the quality of evidence and strength of recommendations. Two of these frameworks, which are usually cited as the most used: the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) and Strength of Recommendation Taxonomy (SORT) (*Maymone et al*, Evaluating the Strength of Clinical Recommendations in the Medical Literature: GRADE, SORT, and AGREE, Journal of Investigative Dermatology, 2014). A quick comparison of the two can be seen in figure 4 below.

	Strength of recommendation	Quality of the evidence
GRADE	Strong for = benefits outweigh risks of the intervention	High quality = further research is very unlikely to change our confidence in the estimate of effect
	Strong against = risks outweigh benefits of the intervention	Moderate quality = further research is likely to have an important impact on our con- fidence in the estimate of effect and may change the estimate
	Weak = most informed people would choose this recommendation but a substantial number would not (risks and burdens finely balanced)	Low quality = further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
		Very low quality = any estimate is very uncertain
SORT	A = based on consistent and good quality patient-oriented evidence	Level 1 = good quality, patient-oriented
	B = based on inconsistent or limited quality patient-oriented evidence	Level 2 = limited quality, patient-oriented
	C = based on consensus, usual practice, opinion, disease-oriented evidence or case series	Level 3 = other evidence (usual practice, opinion, disease oriented evidence)

Table 4 - Comparison between GRADE and SORT with regard to the strength of recommendation and the quality of evidence (adapted from *Maymone et al*, 2014)

Some additional examples from GRADE and SORT follows.

Factors influencing the quality of evidence	
Study design (experimental vs observational)	
Factors that can decrease the quality	Limitations in study design and/or execution Inconsistency of results Indirectness of evidence Imprecision of results Publication bias
Factors that can increase the quality of evidence	Large magnitude of effect All plausible confounding may be working to reduce the demonstrated effect or increase the effect if no effect was observed Dose-response gradient

Table 5 - Factors influencing the quality of evidence (adapted from Schünemann *et al*, 2014)

Criteria for assigning grade of evidence (GRADE)	
Type of evidence	Randomized trial = high Observational study = low Any other evidence = very low
Decrease grade if:	• Serious (– 1) or very serious (– 2) limitation to study quality • Important inconsistency (– 1) • Some (– 1) or major (– 2) uncertainty about directness • Imprecise or sparse data (– 1) • High probability of reporting bias (– 1)
Increase grade if:	• Strong evidence of association—significant relative risk of > 2 (< 0.5) based on consistent evidence from two or more observational studies, with no plausible confounders (+1) • Very strong evidence of association—significant relative risk of > 5 (< 0.2) based on direct evidence with no major threats to validity (+2) • Evidence of a dose response gradient (+1) • All plausible confounders would have reduced the effect (+1)

Table 6 - Criteria for assigning grade of evidence (adapted from *Oxman et al,* BMJ, 2004)

	Type of study		
Study quality	Diagnosis	Treatment/prevention/ screening	Prognosis
Level 1: good-quality, patient-oriented evidence	- Validated clinical decision rule - Systematic review/ meta-analysis of high-quality studies - High-quality diagnostic cohort study	- Systematic review/ meta-analysis or RCTs with consistent findings - High-quality individual RCT - All-or-none study	- Systematic review/ meta-analysis of good-quality cohort studies - Prospective cohort study with good follow-up
Level 2: limited-quality patient-oriented evidence	- Unvalidated clinical decision rule - Systematic review/ meta-analysis of lower quality studies or studies with inconsistent findings - Lower quality diagnostic cohort study or diagnostic case-control study	- Systematic review/ meta-analysis of lower quality clinical trials or of studies with inconsistent findings - Lower quality clinical trial - Cohort study - Case-control study	- Systematic review/ meta-analysis of lower quality cohort studies or with inconsistent results - Retrospective cohort study or prospective cohort study with poor follow-up - Case-control study - Case series
Level 3: other evidence	Consensus guidelines, extrapolations from bench research, usual practice, opinion, disease-oriented evidence (intermediate or physiologic outcomes only), or case series for studies of diagnosis, treatment, prevention, or screening		

Table 7 - Determining quality of Study (SORT) (adapted from Strength of recommendation taxonomy (SORT): a patient-centered approach to grading evidence in the medical literature, *Ebell et al,* Am Fam Physician*., 2004*)

	Consistency across studies
Consistent	Most studies found similar or at least coherent conclusions (coherence means that differences are explainable) *or* If high-quality and up-to-date systematic reviews or meta- analyses exist, they support the recommendation
Inconsistent	Considerable variation among study findings and lack of coherence *or* If high-quality and up-to-date systematic reviews or meta- analyses exist, they do not find consistent evidence in favor of the recommendation

Table 8 - Determining consistency across studies (SORT) (adapted from *Ebell et al,* Am Fam Physician*., 2004*)

Question 2 - Based on the specification and justification of the level of clinical evidence, what is sufficient clinical evidence?

With the understanding of what level of clinical evidence means, we can answer the second question.

As mentioned previously, the level of evidence is related to the quality of evidence and strength of recommendation, but it is also related to the type of studies under evaluation. Moreover, the types of studies are directly related to the well-developed research question that is the bases of the clinical evaluation, as can be seen in Table 9.

Types of study to consider depending on research question type	
Type of research question	Types of studies
Harm	RCT > cohort > case control > case series
Causal / Risk factors	RCT > cohort > case control > case series
Screening / Diagnosis	RCT > cohort > case control > case series
Prognosis	Retrospective, blind comparison with gold standard
Prevention	Cohort > case control > case series
Patient / consumer / participant experience or perceptions	RCT > cohort / case control / case series
Service Delivery	Qualitative studies, the most common being phenomenological, ethnographic and grounded theory.
Cost effectiveness	RCT

Table 9 - Types of study to consider depending on research question type (adapted from *Foster,* Assembling the Pieces of a Systematic Review: A Guide for Librarians (Medical Library Association Books Series), 2017)

Also, it's important to remember that clinical evidence is not clinical data only - it's clinical data that has been evaluated under the clinical evaluation process and that the appraisal stage identified as having methodological quality and scientific validity.

Finally, the "quantity" of clinical evidence depends on the type of results (quantitative or qualitative) that are achieved, and is represented by the evaluation of whether the clinical evidence (which is a sample) can be generalized. Table 10 defined the generic aspects of each approach (and these have to be defined in the Clinical Evaluation plan).

Quantitative versus qualitative research design		
	Quantitative approach	Qualitative approach
Number of Observations	Many	Few or one
Research Question	Who, what, where, when	How, Why
Variables	Specified earlier, based on theoretical concepts	Emerges from the study, based on grounded research
Data collect	One variable at a time	One case at a time
Analyis	Level of variables and relationships between them; statistical analysis	Event or process pattern discovery
Objective	Generalizable for observations or contexts beyond the sample	Generalizable for theoretical concepts

Table 10 - Aspects of quantitative versus qualitative research design

Thus, what is a sufficient level of clinical evidence is the result of the following :
- Analysis of research question;
- Analysis of types of studies that turned into evidence, having methodological quality and scientific validity after being appraised using a critical appraisal worksheet;
- Evaluation of match between research question and types of studies that turned into evidence;
- Evaluation of any additional, quality evidence;
- Evaluation os quantity of clinical evidence, based on type of evidence (quantitative versus qualitative) and the conclusion if the clinical evidence is generalizable;
- Conclusion if the final clinical evidence can be understood as having enough quality (as defined by the level of clinical evidence from the appraisal step) and quantity (as defined by the analysis if the clinical evidence can be generalized) to show compliance with the related ERs.

Final observation: This discussion only provided the basic scientific backgrounds to the related topics. For a medical device clinical evaluation, there's need to be some adaptation of the concepts, which are not discussed here. Maybe in the future I will write a discussion on how to perform this adaptation.