CORE-MD

Coordinating Research and Evidence for Medical Devices

Regulation of AI Medical Devices

Task 2.3

Frank E. Rademakers

Membership of CORE–MD Task 2.3

Frank Rademakers (KU Leuven)

Elisabetta Biasin (KU Leuven)

Bart Bijnens (Pompeu Fabra University)

Nico Bruining (Erasmus MC)

Enrico Caiani (Politecnico de Milano)

Koen Cobbaert (Philips / ISO & IEC)

Rhodri Davies (University College London)

Job Doornberg (University Medical Center Groningen)

Stephen Gilbert (Technische Universität Dresden)

Leo Hovestadt (Elekta)

Erik Kamenjasevic (KU Leuven)

Zuzanna Kwade (Dedalus)

Gearoid McGauran (HPRA)

Gearoid O'Connor (HPRA)

Baptiste Vasey (Oxford University)

Alan Fraser (ESC, Cardiff University)

Some EU & global organisations preparing guidance for AIMDs

- Definitions
- Existing guidance
- Reporting standards
- Clinical utility
- Regulatory initiatives

Fraser AG et al, Exp Rev Med Dev. 2023; 20: 467–491

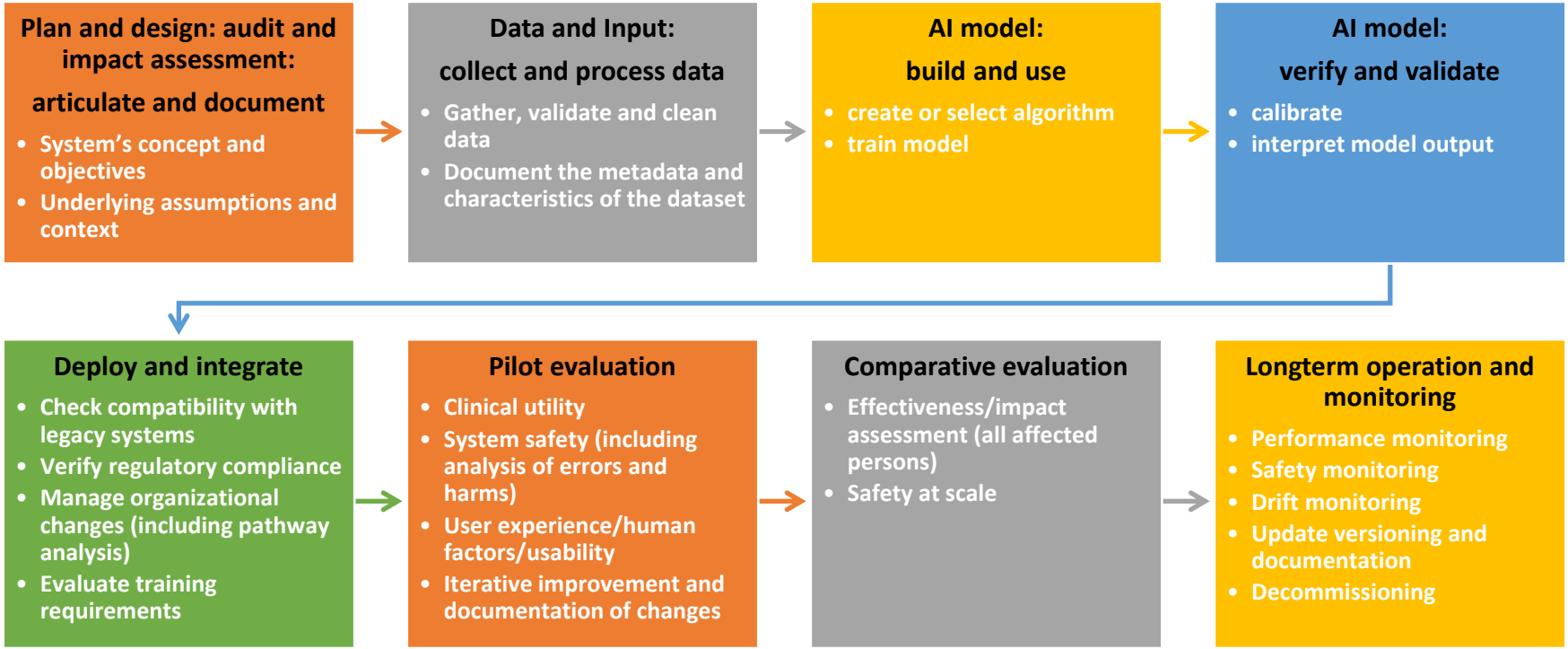
Recommendations – principles

- Risk–based approach with scoring system
  - Type of disease, condition, healthcare situation
  - Significance of information
  - Quality and transparency of data used for training, validation, testing
  - Human interpretability & usability in clinical workflow
- Depending on risk score
  - Use MDCG 2020-1 document on Guidance on Clinical Evaluation
  - Matrix of requirements for clinical evaluation: the ability of the AI tool to yield clinically meaningful output, in accordance with the intended purpose
    - ‘Certificates with conditions’ for lower-risk AI devices with proportional level of clinical evaluation
    - Pre-release / Post-release balance

Risk–benefit assessment : factors to consider

1. **Medical purpose**
2. **The intended medical conditions**
3. **The intended population**
4. **The AI tool's operation**
    1. Inputs: data
    2. Algorithm
    3. Outputs
    4. Presentation of results
    5. Integration in workflow
    6. Outcomes impacted by AI tool
5. **The intended user in their respective environment and scope**
6. **The potential for shift/drift ('off-label use')**

Requirements throughout the AI life-cycle



(adapted after NIST / National Institute of Standards and Technology)

Clinical performance score [relevant and beneficial]		**CPS**
Type of disease, condition, disability, healthcare situation: risk for patient		
Non-serious / serious / critical	**1 / 2 / 3**	
Significance of information: use in clinical flow		
Inform / drive / diagnose or treat	**1 / 2 / 3**	
Technical performance score [accurate and reliable]		**TPS**
Extent of validation and testing		
Broad, external / narrow, external / internal		
Strong / moderate / weak	**1 / 2 / 3**	
Valid clinical association score [indication]		**VCAS**
Transparency and oversight		
Easy / difficult / impossible	**1 / 2 / 3**	

Need for extended clinical evaluation

CPS ≥ 5

CPS + TPS ≥ 6

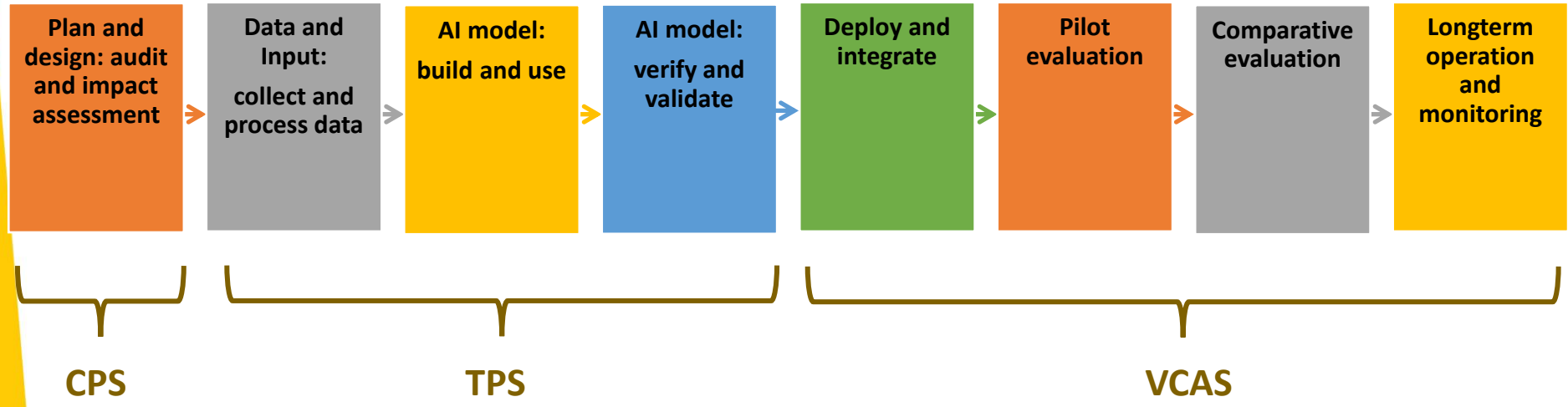
CPS + TPS + VCAS ≥ 8

Proportional level of *pre-market* clinical evaluation if CPS + TPS + VCAS Total score ≤ 7

CORE-MD

Coordinating Research and Evidence for Medical Devices

Different Scores relative to AI Life-Cycle



Recommendations for *pre-release* evaluation & approval

Evaluation proportionate to lesser level of risk

Evaluation proportionate to greater level of risk

Plan and design: audit and impact assessment	System's concept and objectives	+	+
	Underlying assumptions and context	+	+
Data and Input: collect and process data	Gather, validate and clean data	+	+
	Document the metadata and characteristics of the datasets	+	+
AI model: build and use	Create or select algorithm	+	+
	Train model	+	+
AI model: verify and validate	Calibrate	+	+
	Interpret model output	+	+
Deploy and integrate	Check compatibility with legacy systems	+	+
	Verify regulatory compliance	+	+
	Manage organizational changes (including pathway analysis)	-	+
	Evaluate training requirements	-	+
Pilot evaluation	Clinical utility	+	+
	System safety (including analysis of errors and harms)	+	+
	User experience/human factors/usability	-	+
	Iterative improvement and documentation of changes	-	+
	Effectiveness/impact assessment (all affected persons)	-	+
Comparative evaluation	Safety at scale	-	+
Long term operation and monitoring	Performance monitoring	-	-
	Safety monitoring	-	-
	Drift monitoring	-	-
	Update versioning and documentation	-	-
	Decommissioning	-	+

Recommendations for *post-release* evaluation & approval

Evaluation proportionate to lesser level of risk

Evaluation proportionate to greater level of risk

Plan and design: audit and impact assessment	System's concept and objectives	+	+
	Underlying assumptions and context	+	+
Data and Input: collect and process data	Gather, validate and clean data	+	+
	Document the metadata and characteristics of the datasets	-	-
AI model: build and use	Create or select algorithm	-	-
	Train model	-	-
AI model: verify and validate	Calibrate	-	-
	Interpret model output	-	-
Deploy and integrate	Check compatibility with legacy systems	+	+
	Verify regulatory compliance	+	+
	Manage organizational changes (including pathway analysis)	+	+
	Evaluate training requirements	+	+
Pilot evaluation	Clinical utility	+	+
	System safety (including analysis of errors and harms)	+	+
	User experience/human factors/usability	+	+
	Iterative improvement and documentation of changes	+	+
	Effectiveness/impact assessment (all affected persons)	+	+
Comparative evaluation	Safety at scale	+	+
Long term operation and monitoring	Performance monitoring	+	+
	Safety monitoring	+	+
	Drift monitoring	+	+
	Update versioning and documentation	+	+
	Decommissioning	+	+

Conclusion

- Use of Simplified Risk Score
- Proportional Pre- and Post- Release Clinical Evaluation
- More emphasis on post –release in view of AI specific self-learning, drift and personalized use

In order to

- Create maximal potential Benefit
- Avoid Harm and Injustice
- Provide complete Transparency and Autonomy

For all End-Users

Future Directions

- The aim is to collaborate further with EU regulators and all stakeholders, within the framework of the MDCG, to develop a specific European guidance document on the clinical evaluation of AI MDSW.
- There is a pressing need for regulatory collaborations to avoid over-regulation.

CORE-MD, Coordinating Research and Evidence for Medical Devices, aims to translate expert scientific and clinical evidence on study designs for evaluating high-risk medical devices into advice for EU regulators.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 945260

For more information, visit: www.core-md.eu

ESC
EFORT
Biomedical Alliance in Europe
European Academy of Paediatrics
EPF European Patients Forum
Leiden University Medical Center
University of Oxford
Region Uppsala
RCSI University of Medicine and Health Sciences
Inselspital Universitätsspital Bern Hôpital universitaire de Berne
KU Leuven
UMIT Tirol
University of Gothenburg
Politecnico Milano 1863
HPRA
Lægemiddelstyrelsen Danish Medicines Agency
National Institute for Public Health and the Environment Ministry of Health, Welfare and Sport
Istituto Superiore di Sanità
ainta HTA Austria Austrian Institute for Health Technology Assessment GmbH
AETSA
TEAM NB

CORE-MD
Coordinating Research and Evidence for Medical Devices