AIaMD – Sailing the boat while we build it

AI as a medical device (AIaMD)

AIaMD as a subset of SaMD

Bulk of the UK AIaMD market is orientated toward diagnosis or triaging (~ 80%)

Avoid 'AI exceptionalism'

Requires the best of medical device regulation and the best of data science

IMDRF AI Medical Device Working Group

IMDRF International Medical Device Regulators Forum

Final Document

IMDRF/AIMD WG/N67

Machine Learning-enabled Medical Devices: Key Terms and Definitions

Good Machine Learning Practice

Good Machine Learning Practice for Medical Device Development: Guiding Principles

October 2021

Medicines & Healthcare products Regulatory Agency

Guidance

Good Machine Learning Practice for Medical Device Development: Guiding Principles

Published 27 October 2021

Multi-Disciplinary Expertise Is Leveraged Throughout the Total Product Life Cycle	Good Software Engineering and Security Practices Are Implemented	Clinical Study Participants and Data Sets Are Representative of the Intended Patient Population
Training Data Sets Are Independent of Test Sets	Selected Reference Datasets Are Based Upon Best Available Methods	Model Design Is Tailored to the Available Data and Reflects the Intended Use of the Device
Focus Is Placed on the Performance of the Human-AI Team	Testing Demonstrates Device Performance During Clinically Relevant Conditions	Users Are Provided Clear, Essential Information
	Deployed Models Are Monitored for Performance and Re-training Risks Are Managed	

AIaMD Challenges
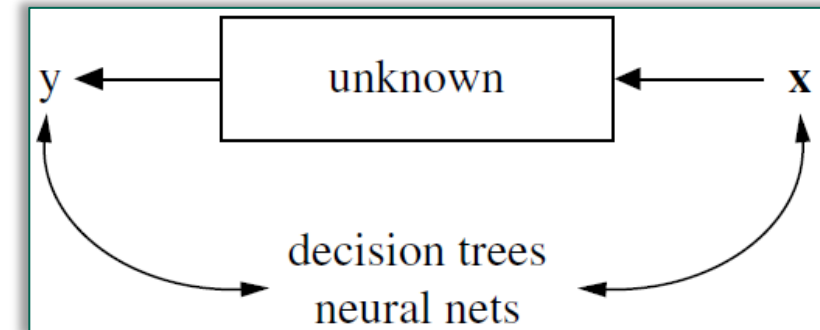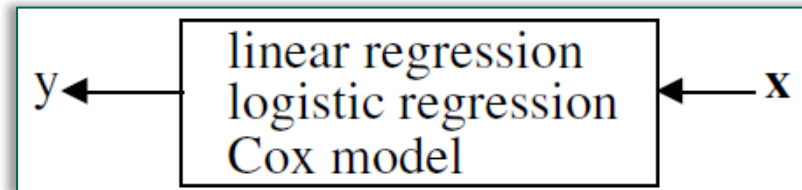
AIaMD *can* (but need not always) provide challenges over and above SaMD, namely:

A. Interpretability of AIaMD
B. Evidencing AIaMD
C. Adaptivity of AIaMD

AIaMD Interpretability (1)

Data modeling culture v algorithmic modeling culture ([Breiman 2001](#))

Two primary challenges of uninterpretable AIaMD:

1. Linking to clinical / scientific evidence or otherwise validating the model

2. Human factors

      Performance of the Human-AI team
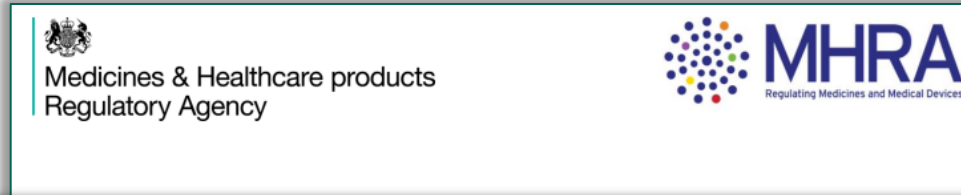
AIaMD Interpretability (2) Human Factors

Medicines & Healthcare products Regulatory Agency

MHRA — Regulating Medicines and Medical Devices

IEEE P7001 - TRANSPARENCY OF AUTONOMOUS SYSTEMS

Scope: This standard describes measurable, testable levels of transparency, so that autonomous systems can be objectively assessed and levels of compliance determined.

Guidance on applying human factors and usability engineering to medical devices including drug-device combination products in Great Britain

The Lancet Digital Health
Volume 3, Issue 11, November 2021, Pages e745-e750

ELSEVIER

Viewpoint

The false hope of current approaches to explainable artificial intelligence in health care

Marzyeh Ghassemi PhD [a, b], Luke Oakden-Rayner [c], Andrew L Beam PhD [d, e]

Chartered Institute of Ergonomics & Human Factors

WHITE PAPER

Human Factors and Ergonomics in Healthcare AI

Consensus Statement | Published: 18 May 2022

Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI

Baptiste Vasey, Myura Nagendran, Bruce Campbell, David A. Clifton, Gary S. Collins, Spiros Denaxas, Alastair K. Denniston, Livia Faes, Bart Geerts, Mudathir Ibrahim, Xiaoxuan Liu, Bilal A. Mateen, Piyush Mathur, Melissa D. McCradden, Lauren Morgan, Johan Ordish, Campbell Rogers, Suchi Saria, Daniel S. W. Ting, Peter Watkinson, Wim Weber, Peter Wheatstone, Peter McCulloch & the DECIDE-AI expert group

Nature Medicine 28, 924–933 (2022) | Cite this article
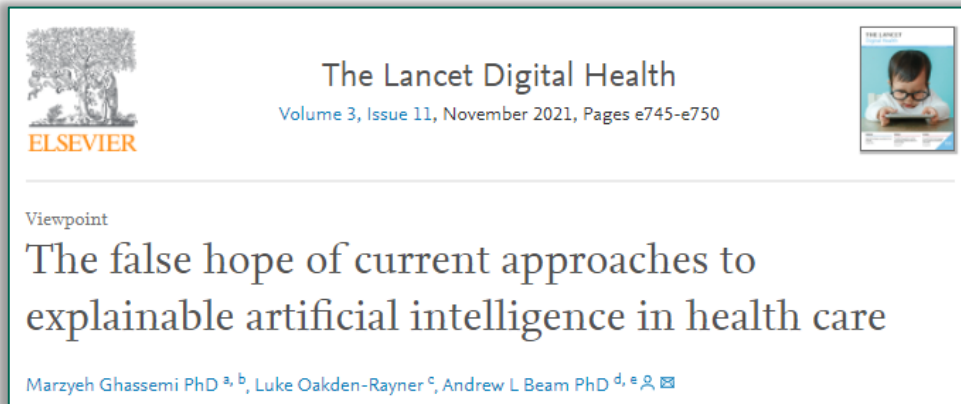
BS EN 62366-1:2015+A1:2020

Medical devices. Application of usability engineering to medical devices

Evidencing AIaMD (1)

AIaMD *can* provide the following challenges:

- Performance of the human-AI team
- Linking to clinical or scientific evidence
- Reference standard ≠ gold standard
- Difficult to claim equivalence between models for clinical evidence (Annex A, <u>MEDDEV 2.7/1 rev 4</u>)
- Representativeness of training data / generalizability / applicability of models
- Reproducibility
- Calibration of models
- Bias

Article | Published: 31 May 2021

AI for radiographic COVID-19 detection selects shortcuts over signal

Alex J. DeGrave, Joseph D. Janizek & Su-In Lee ✉

Nature Machine Intelligence (2021) | Cite this article

On the Reproducibility of Neural Network Predictions

Srinadh Bhojanapalli,* Kimberly Wilber, Andreas Veit, Ankit Singh Rawat, Seungyeon Kim, Aditya Menon, Sanjiv Kumar

Google Research, New York

Opinion | Open Access | Published: 16 December 2019

Calibration: the Achilles heel of predictive analytics

Ben Van Calster ✉, David J. McLernon, Maarten van Smeden, Laure Wynants & Ewout W. Steyerberg On behalf of Topic Group 'Evaluating diagnostic tests and prediction models' of the STRATOS initiative

BMC Medicine **17**, Article number: 230 (2019) | Cite this article

31k Accesses | **177** Citations | **190** Altmetric | Metrics

Evidencing AIaMD (2) Bias

Three primary issues from a medical device POV:

1. Unrepresentative or skewed data may lead to lower performance in subpopulations
2. Representative data but without context may lead to poorer outcomes
3. AIaMD may not serve the needs of communities in which it is deployed if those communities' needs are not understood

Reading Race: AI Recognizes Patient's Racial Identity In Medical Images

[*] Banerjee I PHD[1], Bhimireddy AR MS[2], Burns JL MS[2], Celi LA MD[3,4], Chen L[5], Correa R[6], Dullerud N[7], Ghassemi M PHD[3,8], Gichoya JW MD[9], Huang S[10], Kuo P PHD[5], Lungren MP MD[10], Price BJ[11], Purkayastha S PHD[2], Pyrros AA MD[15], Oakden-Rayner L MD[12], Okechukwu C[13], Seyyed-Kalantari L PHD[14], Trivedi H MD[9], Wang R[5], Zaiman Z[6], Zhang H[7]

Evidencing AIaMD (3) Bias

Proceedings of Machine Learning Research 154:49–62, 2021   LIDTA 2021

BayesBoost: Identifying and Handling Bias Using Synthetic Data Generators

Barbara Draghi
Zhenchen Wang
Puja Myles
Medicine and Healthcare products Regulatory

Allan Tucker
Brunel University London, UK

Press release

Review launched into the health impact of potential bias in medical devices

Independent review will look at potential bias in items like oxygen measuring devices and the impact on patients from different ethnic groups.

STANDING Together
Developing STANdards for data Diversity, INclusivity and Generalisability

PD ISO/IEC TR 24027:2021

TECHNICAL REPORT   **ISO/IEC TR 24027**

First edition
2021-11

Information technology — Artificial intelligence (AI) — Bias in AI systems and AI aided decision making
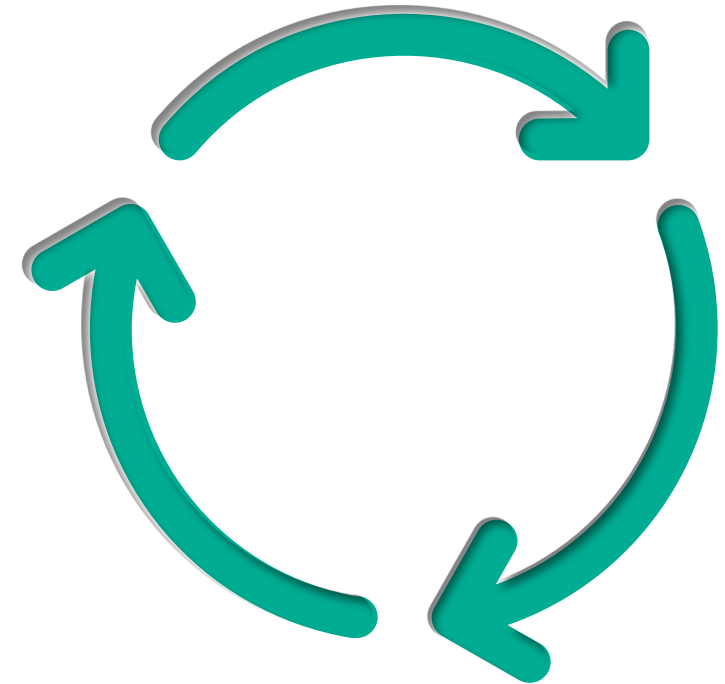
Technologie de l'information — Intelligence artificielle (IA) — Tendance dans les systèmes de l'IA et dans la prise de décision assistée par l'IA

AIaMD Adaptivity (1) Distinguishing Concepts

Distinguish between at least four different issues:

1. Static devices
2. Batch training
3. 'Individualised' models
4. Continuous learning on streaming data

AIaMD Adaptivity (2) Change Management

Different aspects of change that need consideration:

- Assuring intentional change made by the manufacturer

  Non-linearity of change

  Bugs

- Changes in deployment

  Generalisability

  Localisation

- Changes to the environment

AIaMD Adaptivity (3)

MDIC
MEDICAL DEVICE
INNOVATION CONSORTIUM

Artificial Intelligence and Machine Learning for In Vitro Diagnostics

Developing metrics that could signal significant changes in adaptive learning AI algorithms

Project led by

The Medicines and Healthcare products Regulatory Agency

AIaMD: What's needed?

- The critical need for robust standards for AIaMD – legislation and guidance only get us so far

- Harmonisation of regulation at an international level

- The state of the art for AIaMD is still settling

- Different core AI challenges are at different levels of maturity

- Complexity of the AI standardisation landscape internationally