Validation of ML Algorithms for Medical Applications

Andreu Badal, DIDSR/OSEL/CDRH/**Food and Drug Administration**

- **So what?** The FDA is responsible for protecting the public health by ensuring the **safety, efficacy, and security** of human and veterinary drugs, biological products, and **medical devices**.
  - Software *in* Medical Devices (**SiMD**) and Software *as* Medical Devices (**SaMD**) are an essential, and fast growing, segment of the medical devices market.
  - **Machine learning** (AI/ML) methods are being implemented in all kinds of medical applications, and with particular strength in **medical imaging**.
    - Unlike older algorithms, ML algorithms learn from examples and their performance might drop with new imaging devices or sub-populations not included in the training dataset.

- **Who cares?** **Patients** trust FDA-approved or cleared products to be safe and effective. Device **manufacturers** expect a clear and fair path to market (least burdensome).

ALERT AWARENESS AND LOCALIZATION OF EXPLOSIVES-RELATED THREATS
ADSA-24
October 6, 2021

- **Example SaMD in radiology**:
  - Image enhancement software, such as denoising.
  - Automatic image segmentation (radiation treatment planning, 3D printing).
  - Computer-Aided Diagnosis (CAD) software:
    - CADe (detection): Aids in localizing/marking abnormal regions.
    - CADx (diagnosis): Aids in assessing disease type, severity, stage, progression.
    - CADe/x (detection/diagnosis): Aid in localizing and characterizing conditions.
    - CADt (triage): Aids in prioritizing/triaging time sensitive patient diagnosis.
    - CADa/o (acquisition/optimization): Aid in image or signal acquisition (DEN190040).

- Example CADe/x: **MammoScreen** by Therapixel (www.mammoscreen.com)

- The regulatory evaluation of ML software has two components:
  - Standalone performance: assess performance, robustness and generalizability of the algorithm with a testing dataset independent of training dataset.
  - Clinical performance: required clinical evidence depending on risk level.
    - For image enhancement software, evaluating sample clinical images might be sufficient.
    - For CAD software, ROC analysis (sensitivity vs specificity) and a MRMC reader study might be used to assess clinician's performance with/without the device.
      - Each CAD type has special controls or guidance documents defining the required testing: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-performance-assessment-considerations-computer-assisted-detection-devices-applied-radiology
- For more information on FDA ML regulations: www.fda.gov/digitalhealth
  - Whitepaper: Artificial Intelligence and Machine Learning Software as a Medical Device Action Plan (www.fda.gov/media/145022/download)
  - Upcoming Virtual Public Workshop (October 14, 2021): Transparency of Artificial Intelligence/Machine Learning-enabled Medical Devices