1 - Identifying Fast and Slow Growers: Moving Towards a Machine Learning Predictive Model Approach for Individualized Abdominal Aortic Aneurysm Surveillance

Authors: Ahmad Aljobeh, Praveen Parthasarathy, Jonathan Liao, William Kartsonis, Apostolos Tassiopoulos

Faculty Mentor: Apostolos Tassiopoulos

Authorship role: substantive intellectual contribution, study conception, study design, acquisition, analysis, and interpretation of data, writing the initial abstract, revising the submitted abstract, agreeing to be held publicly accountable for the accuracy/integrity of the project

Background: Current abdominal aortic aneurysm (AAA) surveillance guidelines rely on maximum aneurysm diameter to determine follow-up intervals. However, this approach may not capture individual variability in aneurysm growth. Machine learning (ML) models offer the potential for personalized surveillance protocols by predicting aneurysm growth using patient and imaging features. We performed a systematic review and critical appraisal of existing ML predictive models developed to classify AAAs as fast or slow growers.

Methods: A systematic review was conducted to identify studies using ML to predict AAA growth classification. PubMed/Medline, Embase, and Web of Sciences databases were queried with relevant keywords. Studies were included if they developed or validated an ML model to classify AAAs as fast or slow growing based on defined annual growth thresholds. Key model characteristics, input features, and performance metrics were extracted. Reporting quality was assessed using the TRIPOD-AI checklist, designed to evaluate transparent reporting in AI-based prediction model studies on 27 different domains.

Results: Eight studies met inclusion criteria after screening 3,129 articles. Models used a range of supervised learning algorithms including XGBoost, support vector machines (SVM), and an extra trees classifier. Fast/slow growth cutoffs varied between 2.5–5 mm/year. Input features included geometric, computational fluid dynamics (CFD), radiomic, and patient health data. Sample sizes ranged from 36 to 195 patients. Reported model performance varied, with AUCs ranging from 0.79 to 0.93 and RMSE as low as 1.83. TRIPOD-AI scores ranged from 17 to 18 out of 27 applicable items. Results are summarized in **Table 1**. All studies appropriately reported objectives, outcomes, and model performance. Common reporting deficiencies included limited discussion of data representativeness, lack of attention to health disparities, insufficient sample size justification, and inadequate transparency for reproducibility (e.g., lack of code availability).

Conclusion: ML models show promise in predicting AAA growth and supporting individualized surveillance strategies. However, variability in model inputs, performance, and definitions of fast growth pose challenges for clinical adoption. Future research should focus on external validation, improved reporting transparency, and addressing sociodemographic representativeness to enhance clinical applicability. Adoption of TRIPOD-AI reporting standards may facilitate the development of robust, generalizable ML tools for AAA surveillance.