Machine Learning Algorithms to Predict Intracranial Aneurysm Rupture Risk

Authors: Karan Daga, Siddharth Agarwal, Zaeem Moti, Matthew BK Lee, Munaib Din, David Wood, Marc Modat, Thomas C Booth.

Purpose

- Intracranial aneurysms affect 3.2% of the population.[1]
- Interventional management of these aneurysms are not without considerable risk.
- Currently no universally agreed reference standard exists to determine rupture risk.
- Examples of scoring criteria in use: PHASES, UIATS, ISUIA and ELAPSS.
- Recent studies have shown these have poor predictive performance in practice.[2]
- Current clinical practice combines validated score prediction with multi-disciplinary expert clinician decision making.
- Machine learning (ML) poses the ability to rapidly process large multi-modal datasets.
- This includes regression models, classical ML and deep learning
- Existing literature demonstrates several ML models have been developed for prediction of aneurysm rupture risk, but there is a lack of clinical validation.

Methods

- PROSPERO Registered: CRD 42023452509.
- Review followed PRISMA guidelines.
- Informed by CLAIM and Cochrane review guidance.

- Databases searched: EMBASE, MEDLINE, Web of Science, Cochrane Library
- Search performed by 3 clinicians (KD,SA,ZM).

- **Inclusion Criteria:**
  - Any ML model predicting unruptured intracranial aneurysm (UIA) rupture risk.
  - Studies involving UIA stability prediction were included if stability was a composite outcome including rupture risk.

- **Exclusion Criteria:**
  - Studies that solely investigated singular risk factors, not developing a predictive model
  - Studies predicting rupture status (binary classification of if the aneurysm is ruptured or unruptured)
  - Animal studies and studies in pediatric populations (Age <18 years)
  - Conference abstracts, pre-prints, non-peer reviewed studies were not included

Results

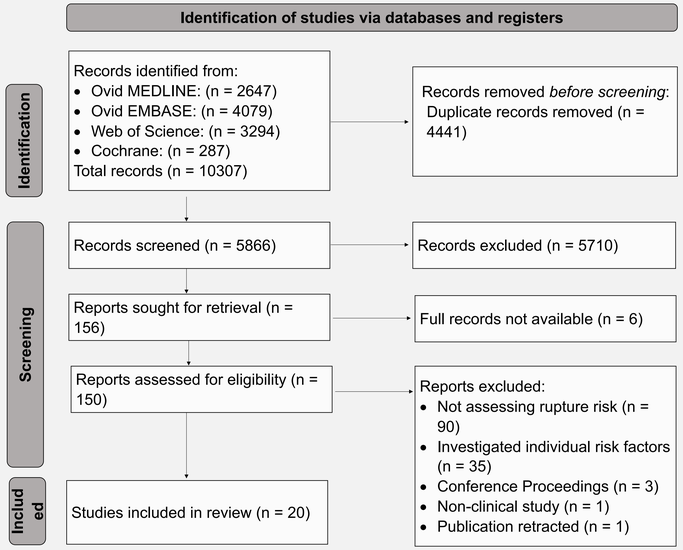
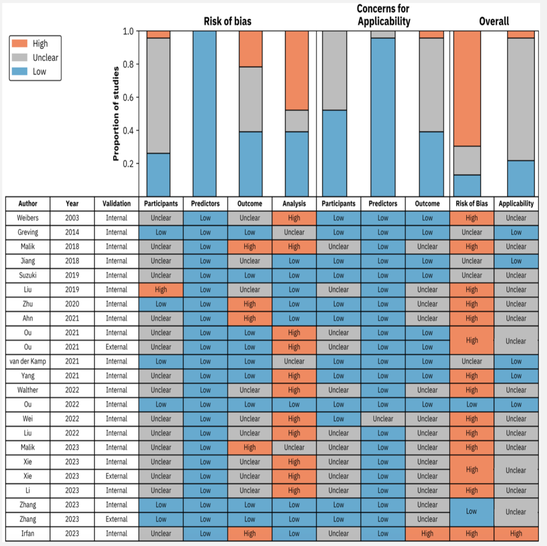


Figure 1: PRISMA Flow Diagram

- **Reference Standards**
  - 16/20 studies: Occurrence of stated end point; rupture or composite outcome of stability
  - 3/20 Studies: Expert neurosurgeon/neuroradiologist opinion
  - 1/20 Studies: Comparison to the PHASES Score

- **Index tests:**
  - Best performing index tests across studies:
  - 10/20 studies: regression model or classical ML model
  - 5/20 studies: deep Learning model
  - 5/20 studies: Combination model (regression/classical ML/deep learning)

- **Study Design:**
  - 4/20 Prospectively validated
  - 8/20 Retrospectively validated
  - 8/20 Performed no separate validation



Author	Year	Validation	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of Bias	Applicability
Weibers	2003	Internal	Unclear	Low	Unclear	High	Low	Low	Low	High	Unclear
Greving	2014	Internal	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Malik	2018	Internal	Unclear	Low	High	High	Unclear	Low	Unclear	High	Unclear
Jiang	2018	Internal	Unclear	Low	Unclear	Low	Low	Low	Low	Unclear	Low
Suzuki	2019	Internal	Unclear	Low	Low	Low	Low	Low	Low	Unclear	Unclear
Liu	2019	Internal	High	Low	Unclear	Low	Unclear	Low	Unclear	High	Unclear
Zhu	2020	Internal	Low	Low	High	Low	Low	Low	Unclear	High	Unclear
Ahn	2021	Internal	Unclear	Low	High	Low	Low	Low	Unclear	High	Unclear
Ou	2021	Internal	Unclear	Low	Low	High	Unclear	Low	Low	High	Unclear
Ou	2021	External	Unclear	Low	Low	High	Unclear	Low	Low		
van der Kamp	2021	Internal	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low
Yang	2021	Internal	Unclear	Low	Low	High	Low	Low	Low	High	Low
Walther	2022	Internal	Unclear	Low	Unclear	High	Unclear	Low	Unclear	High	Unclear
Ou	2022	Internal	Low	Low	Low	Low	Low	Low	Low	Low	Low
Wei	2022	Internal	Unclear	Low	Unclear	High	Low	Unclear	Unclear	High	Unclear
Liu	2022	Internal	Unclear	Low	Unclear	High	Unclear	Low	Unclear	High	Unclear
Malik	2023	Internal	Unclear	Low	High	Unclear	Unclear	Low	Unclear	High	Unclear
Xie	2023	Internal	Unclear	Low	Unclear	High	Unclear	Low	Unclear	High	Unclear
Xie	2023	External	Unclear	Low	Unclear	High	Unclear	Low	Unclear		
Li	2023	Internal	Unclear	Low	Unclear	High	Unclear	Low	Unclear	High	Unclear
Zhang	2023	Internal	Low	Low	Low	Low	Low	Low	Unclear	Low	Unclear
Zhang	2023	External	Low	Low	Low	Low	Low	Low	Unclear		
Irfan	2023	Internal	Unclear	Low	High	Low	Unclear	Low	High	High	High

Figure 2: PROBAST risk of bias diagram

- **Results:**
  - 14/20 (70%) of studies assessed UIA rupture risk.
    Accuracies ranged 0.66-0.99.
  - 6/20 (30%) of studies assessed UIA stability (defined as a composite outcome of rupture or growth)
    Accuracies ranged 0.83-0.94

  - 6/20 studies (30%) compared ML performance to validated scoring criteria used in practice (eg. PHASES).
    ML performed better
    3/20 studies (15%) compared ML performance to expert clinician risk assessment.

- 3/20 (15%) studies had a geographically separate test set to their training set. (ie External hold-out data set)
- 13/20 studies identified a 'high' risk of bias or concern for applicability in at least one domain.

Discussion

- Current evidence:

  - Current literature demonstrates AI rupture risk prediction might have superior predictive value than validated scoring criteria.
  - Neurosurgeon/neuroradiologist opinion remains superior to AI prediction.
  - AI + reader was shown to be the most effective prediction.
  - For example, one study investigated AI alone, expert neurosurgeon prediction alone (reader alone) and expert opinion aided by the neural network's assistive diagnosis system (reader + AI), recording an AUC of 0.85, 0.88 and 0.95 (p=0.037) respectively.[3]
  - The systematic review findings suggest a role for AI as an adjunct in clinical decision making

- Role in Clinical Practice:

  - Aneurysm rupture prediction serves to guide clinicians and patients towards management options.
  - Risks of treatment (thromboembolic ischemia, peri-operative hemorrhage, inadequate occlusion, need for re-intervention and death)must be balanced against that of potential rupture.
  - Developments in endovascular treatments have substantially reduced peri-operative morbidity and mortality compared to previous studies in the literature.
  - Consequently, clinical practice on whether an UIA should be treated or not will continue to evolve.

- Limitations

  - Selection bias – Institutions prospectively investigating UIA have a duty to offer high risk candidates prophylactic treatment, excluding them from analysis
  - Validation – Only three studies were found to perform external validation of their ML model
  - Reference standard ambiguity – 5 studies did not specify time periods for risk assessment. High risk of observer bias where reference standards were expert opinion

- **References:**

[1]Vlak MH, Algra A, Brandenburg R, Rinkel GJ. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. The Lancet Neurology. 2011 Jul 1;10(7):626-36.
[2]Juvela S. PHASES score and treatment scoring with cigarette smoking in the long-term prediction of rupturing of unruptured intracranial aneurysms. Journal of Neurosurgery. 2021 Jul 9;136(1):156-62.
[3]Ou C, Li C, Qian Y, Duan CZ, Si W, Zhang X, Li X, Morgan M, Dou Q, Heng PA. Morphology-aware multi-source fusion–based intracranial aneurysms rupture prediction. European Radiology. 2022 Aug;32(8):5633-41.