Supplementary materials

Supplementary A. PROSPERO protocol

Review title: Machine learning for predicting violent behaviour: a systematic review.

Condition or domain being studied: Violent behaviour; Violence; Violence Prevention; Crime; Machine learning

Rationale for the review: There is a growing number of machine learning prediction tools for violence that are being developed, and little is known about the quality of such tools.

Review objectives: What is the quality and clinical applicability of machine learning models that predict violent behaviour?

Keywords: Machine learning; Violence; Prediction

Country: United Kingdom

Eligibility criteria:

- Population: General population, prison population, psychiatric populations. No exclusion criteria was set for the population.
- Intervention(s) or exposure(s): This review will be exploring the predictive modelling tools designed using machine learning methods, such as neural networks, kNN, random forest etc. Excluded: white box and non-machine learning models.
- Comparator(s) or control(s): People who do not exhibit violent behaviour. No exclusion criteria was set.
- Study design: Only nonrandomized study types will be included. Observational (retrospective and prospective cohort) studies that resulted in development and/or validation of the prediction model will be included.

Searching and screening:

- Both published and unpublished studies will be sought.
- The main databases to be searched are Embase - Embase via Ovid, MEDLINE, PsycInfo, PubMed, SCI - Science Citation Index and Scopus.

- Other important or specialist databases that will be searched: Dissertations, arxiv.org, Google Scholar.
- There are no language restrictions.
- There are no search date restrictions.
- Other studies will be identified by: contacting authors or experts, reference list checking, searching conference proceedings and searching dissertation and thesis databases.
- Studies will be screened independently by at least two people (or person/machine combination) with a process to resolve differences.

Data extraction from published articles and reports:

- Data will be extracted independently by at least two people (or person/machine combination) with a process to resolve differences.
- Authors will be asked to provide any required data not available in published reports.
- Risk of bias will be assessed using: *PROBAST*
- Data will be assessed by one person (or a machine) and checked by at least one other person (or machine).
- Additional information will be sought from study investigators if required information is unclear or unavailable in the study publications/reports.

Outcomes to be analysed: Any violent behaviour: violent recidivism, terrorism, murder, domestic violence etc. There are no additional outcomes.

Strategy for data synthesis: Narrative synthesis is planned for this review. We are planning to extract and synthesise all data related to model performance, internal or external validation and calibration of the models analysed.

Supplementary B. Search strategy

((violen or recidiv* or crim* or reoffen*) and predict* and (machine learning or deep learning or AI)).mp. not (systematic review or meta-analysis or meta analysis or narrative review or literature review).ab*