Conduct and Reporting of studies evaluating Artificial Intelligence/Machine Learning algorithms used for diagnosis and prognosis

Review Protocol

Version of March 15, 2019

Alexandre Vivot, Mauricia Davidson, Raphaël Porcher, Isabelle Boutron, Philipe Ravaud

Centre de Recherche Épidémiologie et Statistiques Université de Paris (CRESS-UMR1153)

Inserm / Université Paris Descartes

Centre d'épidémiologie clinique,

Hôpital Hôtel-Dieu, APHP,

1 place du Parvis Notre-Dame

75004 Paris, France

1. INTRODUCTION

Artificial Intelligence (AI) and machine learning (ML) has the potential to completely transform the management of patients in medicine. Through algorithms, AI can learn, diagnose, predict and advise patients and care providers. For instance, AI can now diagnose breast cancer with radiologist-level accuracy, skin cancer like dermatologists and even predict lung cancer patient survival classification (1–3). These successes are driven by two factors: the increasing availability of medical data and the emergence of new powerful technologies of machine learning (e.g. gradient boosting, reinforcement learning, deep learning, support vector machine, artificial neural networks, random forests etc.)

However, there is an important gap between the extremely rapid development and implementation of AI in medicine and the lack of appropriate assessment of the properties of these algorithms. Furthermore, prediction based on AI is not immune to the limitations and bias of classical prediction models such as missing data, imperfect gold standard etc. (4). Some concerns are specific to AI, whereby these algorithms act as black boxes - there is no clear rationale behind the AI algorithms' decisions on a particular course of action. This is cause for concern, especially for matters that affect public health (5).

All in all, advances in methodological research for diagnosis and prognosis need to keep the pace with advances in AI technology. Otherwise, sub-optimal validation of AI algorithms will limit our ability to capitalize on AI technologies and ultimately, to improve patient care and outcomes. In this project, we will focus on two major areas of applications of AI in medicine: "diagnosis" (including screening and early diagnosis) and "predictive models" (either for the prognosis of patients, the response to treatment or the prediction of side effects). In addition, we will develop a current mapping of the limitations of these algorithms in order to describe the framework for improvement in this field.

1.1 Objectives

1. To examine the conduct and reporting of studies assessing diagnostic and prognostic AI algorithms used in medicine
2. To review how limitations are currently acknowledged and reported in medical literature regarding Artificial Intelligence algorithms used for diagnosis and prognosis.
3. To map the current limitations of studies evaluating medical AI algorithms.

2. METHODS

We will perform a methodological systematic review of clinical studies evaluating AI-based models for diagnosis and prognosis in medicine. The protocol will be registered prospectively on the Prospero registry (https://www.crd.york.ac.uk/prospero/).

2.1 Eligibility Criteria

We will include all original research articles, regardless of the design fulfilling the following inclusion criteria:

1) reporting on an AI/ML-based medical device, algorithm or prediction model;
2) that developed, validated, updated or tested/implemented an AI-based medical device or algorithm;
3) for prediction of a medical/clinical outcome at present (diagnosis of a disease) or in the future (prognosis, prediction of response, of side effects); and
4) in living human patients.

The distinction between AI, machine learning and classical statistical models is not so easy (6). In this review, we do not define what is or what is not AI between all available techniques in the *spectrum* of AI and consider all studies self-reporting as using AI, ML or any classifier prediction model as eligible in this review.

We will exclude systematic and narrative reviews, protocols, expert opinions, guideline reports and studies where prediction is not the main focus. For instance, we will exclude studies in radiation oncology developing AI-based model for contouring, studies of natural language processing without predictions, imaging studies of AI-based segmentation without prediction, studies where AI is used for clustering without associated predictions, studies where AI is only used for pre-processing of image, studies using AI to identify cases in electronic health record (e.g. to include them in an epidemiological study). We will also exclude studies where the predicted outcome is not a disease or a clinical outcome such as sleep stage classification, personality trait prediction. Finally, we will exclude animal studies, studies on healthy people and studies where the prediction concern dead people (e.g. studies in forensic medicine).

2.2 Information Sources

We will search PUBMED/MEDLINE restricting our search to the top 10 first journals for general internal medicine category and the top 5 for each category of medicine using data of Journal Impact Factor from Journal Citation Reports 2018 (Clavirate Analytics). We will exclude journals publishing only reviews (e.g., Nature Reviews series). The list of journals is provided with the search equation in Appendix 1.

2.3 Search Strategy

The search equation is constructed using controlled vocabulary (MeSH) and free-text terms, referring to artificial intelligence, cancer, diagnosis, prognosis and prediction. The search equation is available in Appendix 1. Moreover, because AI is a rapidly evolving field we will restrict the search to articles published after January 1, 2018.

2.4 Study Records

2.4.1 Data Management

The Rayyan QCRI software will be used to manage the records and data obtained for screening and full-text consideration.

2.4.2 Selection Process

The titles and abstracts of retrieved records will be screened independently by two reviewers (MD, AV) according to the pre-specified eligibility criteria in section 2.1. Disagreements will be resolved by consensus. Eligible studies will be included for full-text consideration. Full-text consideration will be performed by the same two independent reviewers with consensus to settle disagreements.

2.4.3 Data Collection Process

A pre-specified, standardized data extraction form will be used to collect data from full-text articles. The draft data extraction form is displayed in Appendix 2 but the final data extraction form could be modified after pilot-testing this draft form on 20 randomly selected articles. Two reviewers will independently extract data and any disagreement will be resolved by consensus.

We will extract data about: general study characteristics (date of publication, journal, country where the study was conducted); objective of the study (development and internal/external validation of AI algorithm, external validation only, evaluation of an

intervention based on an AI algorithm); design (observational study or clinical trial, cross-sectional, prospective or retrospective studies); aim of the algorithm (screening, diagnosis, prognostic, responder identifier, side-effects predictions); the type of algorithm that was used, its performance measures (accuracy, confusion matrices, ROC curves, calibration measure) and the tuning of hyperparameters. We will extract information about all limitations reported; details about limitations i.e. if separate section and heading for limitations, if embedded within Discussion section, size (number of words); countermeasures proposed or taken by authors to overcome potential limitations.

We will extract information separately for the training and the test datasets when possible: sample selection, sample size, data preparation (e.g., image pre-processing) and the type of data used (radiological images, pathological images, clinical data, data from EHR or medical claims, other). We will assess transparency (whether the data and algorithm are available), as well as measures taken to be sure that the algorithm did not have access to any part of the test dataset during its training. We will evaluate whether patients are representative of a clinical problem and determine whether the algorithm should be implemented in clinics (yes, too early, no).

We will extract the limits listed in the article and classify each one to do a mapping. Hence, we will obtain a list of limits and their frequency in articles. We will assess the saturation of the data by the method of Tran et al. (7,8)

2.5 Data Synthesis

Descriptive statistics will be applied to present and summarize the extracted data. If the amount of articles is appropriate, we intend to perform subgroup analyses on studies of imaging vs. non-imaging, as well as oncology vs. non-oncology studies.

3. Expected results

Our study will be the first to evaluate current conducting and reporting of studies on AI in medicine and to provide guidance on how to develop and assess AI algorithms. This project aims to increase transparency of the research, the quality of evidence and hence the success of AI in medicine. Sound methods underpin sound applied research. Biased evaluation methods

could compromise findings from AI studies, therefore, it is crucial that we critically appraise and improve the evaluation methods supporting the use of AI in medicine.

4. REFERENCES

1. Becker A, Marcon M, Ghafoor S, Wurnig M, Frauenfelder T, Boss A. Deep Learning in Mammography. Invest Radiol. 2017 Jul 1;52(7):434–40.

2. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017 02;542(7639):115–8.

3. Yu K-H, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun. 2016 16;7:12474.

4. Wolff RF, Moons KGM, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A Tool to Assess the Risk of Bias and Applicability of Prediction Model Studies. Ann Intern Med. 2019 Jan 1;170(1):51.

5. Luxton DD. Should Watson Be Consulted for a Second Opinion? AMA J Ethics. 2019 Feb 1;21(2):131–7.

6. Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, van Calster B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. J Clin Epidemiol. 2019 Feb 11;

7. Tran V-T, Porcher R, Tran V-C, Ravaud P. Predicting data saturation in qualitative surveys with mathematical models from ecological research. J Clin Epidemiol. 2017 Feb 1;82:71-78.e2.

8. Tran V-T, Porcher R, Falissard B, Ravaud P. Point of data saturation was assessed using resampling methods in a survey with open-ended questions. J Clin Epidemiol. 2016;80:88–96.

Appendix 1. Search Equation

#1	Mesh Terms	("Artificial Intelligence"[Mesh] NOT "Robotics"[Mesh]) OR "Neural Networks (Computer)"[Mesh]
#2	Free Text	Deep learning OR neural network OR gradient boosting OR support vector OR random forest OR ensemble learning OR Machine Learning OR Artificial Intelligence OR Reinforcement learning OR Unsupervised learning OR Supervised learning OR Semi-Supervised Learning OR classifier OR algorithm
#3	Full article	Has abstract [text]
	Search	(#1 OR #2) AND #3

Date of Publication	("2018/01/01"[Date - Publication] : "3000/01/01"[Date - Publication])
Top 10 journals in GIM	"The New England journal of medicine"[Journal] OR "Lancet (London, England)"[Journal] OR "JAMA"[Journal] OR "BMJ (Clinical research ed.)"[Journal] OR "JAMA internal medicine"[Journal] OR "Annals of internal medicine"[Journal] OR "PLoS medicine"[Journal] OR "BMC medicine"[Journal] OR "Mayo Clin Proc"[jour] OR "CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne"[Journal]
Top 5 journals in Allergy	OR “The Journal of allergy and clinical immunology"[Journal] OR "The journal of allergy and clinical immunology. In practice"[Journal] OR "Allergy"[Journal]

	OR "The World Allergy Organization journal"[Journal] OR "Clinical and experimental allergy : journal of the British Society for Allergy and Clinical Immunology"[Journal]
ANATOMY & MORPHOLOGY	OR "Brain structure & function"[Journal] OR "Frontiers in neuroanatomy"[Journal] OR "Developmental dynamics: an official publication of the American Association of Anatomists"[Journal] OR "Journal of anatomy"[Journal] OR "Advances in anatomy, embryology, and cell biology"[Journal]
ANDROLOGY	OR "Asian journal of andrology"[Journal] OR "Andrology"[Journal] OR "The world journal of men's health"[Journal] OR "Andrologia"[Journal] OR "Systems biology in reproductive medicine"[Journal]
ANAESTHESIOLOGY	OR "Anesthesiology"[Journal] OR "British journal of anaesthesia"[Journal] OR "pain"[Journal] OR "Anaesthesia"[Journal] OR "Regional anesthesia and pain medicine"[Journal]
CARDIAC AND CARDIOVASCULAR SYSTEMS	OR "European heart journal"[Journal] OR "circulation"[Journal] OR "Journal of the American College of Cardiology"[Journal] OR "Circulation research"[Journal] OR "European journal of heart failure"[Journal]
CLINICAL NEUROLOGY	OR "The Lancet. Neurology"[Journal] OR "Acta neuropathologica"[Journal] OR "Alzheimer's & dementia: the journal of the Alzheimer's Association"[Journal] OR "JAMA neurology"[Journal]

	OR "Brain : a journal of neurology"[Journal]
CRITICAL CARE MEDICINE	OR "The Lancet. Respiratory medicine"[Journal] OR "American journal of respiratory and critical care medicine"[Journal] OR "Intensive care medicine"[Journal] OR "Chest"[Journal] OR "Critical care medicine"[Journal]
DENTISTRY, ORAL SURGERY & MEDICINE	OR "Periodontology 2000"[Journal] OR "Journal of dental research"[Journal] OR "Oral oncology"[Journal] OR "Clinical oral implants research"[Journal] OR "International journal of oral science"[Journal]
DERMATOLOGY	OR "JAMA dermatology"[Journal] OR "Journal of the American Academy of Dermatology"[Journal] OR "The Journal of investigative dermatology"[Journal] OR "The British journal of dermatology"[Journal] OR "Advances in wound care"[Journal]
EMERGENCY MEDICINE	OR "Resuscitation"[Journal] OR "Annals of emergency medicine"[Journal] OR "Emergencias : revista de la Sociedad Española de Medicina de Emergencias"[Journal] OR "World journal of emergency surgery: WJES"[Journal] OR "Academic emergency medicine : official journal of the Society for Academic Emergency Medicine"[Journal]
ENDOCRINOLOGY & METABOLISM	OR "The lancet. Diabetes & endocrinology"[Journal] OR "Diabetes care"[Journal] OR "Trends in endocrinology and metabolism: TEM"[Journal]

	OR "Thyroid : official journal of the American Thyroid Association"[Journal] OR "Diabetes"[Journal]
GASTROENTEROLOGY & HEPATOLOGY	OR "Gastroenterology"[Journal] OR "Gut"[Journal] OR "Journal of hepatology"[Journal] OR "Hepatology (Baltimore, Md.)"[Journal] OR "The American journal of gastroenterology"[Journal]
GERIATRICS & GERONTOLOGY	OR "Journal of cachexia, sarcopenia and muscle"[Journal] OR "Aging cell"[Journal] OR "Journal of the American Medical Directors Association"[Journal] OR "Aging"[Journal] OR "Aging and disease"[Journal]
GERONTOLOGY	OR "The journals of gerontology. Series A, Biological sciences and medical sciences"[Journal] OR "Journal of the American Geriatrics Society"[Journal] OR "The Gerontologist"[Journal] OR "The American journal of geriatric psychiatry: official journal of the American Association for Geriatric Psychiatry"[Journal] OR "The journals of gerontology. Series B, Psychological sciences and social sciences"[Journal]
HEMATOLOGY	OR "Circulation research"[Journal] OR "Blood"[Journal] OR "The Lancet. Haematology"[Journal] OR "Leukemia"[Journal] OR "Haematologica"[Journal]
IMMUNOLOGY	OR "Annual review of immunology"[Journal] OR "Nature immunology"[Journal]

	OR "Immunity"[Journal] OR "Trends in immunology"[Journal] OR "The Journal of allergy and clinical immunology"[Journal]
INFECTIOUS DISEASES	OR "The Lancet. Infectious diseases"[Journal] OR "The lancet. HIV"[Journal] OR "Clinical infectious diseases: an official publication of the Infectious Diseases Society of America"[Journal] OR "Emerging infectious diseases"[Journal] OR "Euro Surveill"[jour]
MEDICINE, LEGAL	OR "Forensic science international. Genetics"[Journal] OR "International journal of legal medicine"[Journal] OR "Forensic science, medicine, and pathology"[Journal] OR "Forensic science international"[Journal] OR "Legal medicine"[Journal]
MICROBIOLOGY	OR "Cell host & microbe"[Journal] OR "Nature microbiology"[Journal] OR "Trends in microbiology"[Journal] OR "Annual review of microbiology"[Journal] OR "The ISME journal"[Journal]
NEURO IMAGING	OR "NeuroImage"[Journal] OR "Neuroimage Clin"[jour] OR "Human brain mapping"[Journal] OR "Brain imaging and behavior"[Journal] OR "AJNR. American journal of neuroradiology"[Journal]
NUTRITION & DIETETICS	OR "Annual review of nutrition"[Journal] OR "Advances in nutrition (Bethesda, Md.)"[Journal] OR "The American journal of clinical nutrition"[Journal]

	OR "Critical reviews in food science and nutrition"[Journal] OR "The international journal of behavioral nutrition and physical activity"[Journal]
OBSTETRICS & GYNECOLOGY	OR "Human reproduction update"[Journal] OR "American journal of obstetrics and gynecology"[Journal] OR "Ultrasound in obstetrics & gynecology: the official journal of the International Society of Ultrasound in Obstetrics and Gynecology"[Journal] OR "Human reproduction (Oxford, England)"[Journal] OR "Obstetrics and gynecology"[Journal]
ONCOLOGY	OR "The Lancet. Oncology"[Journal] OR "Journal of clinical oncology: official journal of the American Society of Clinical Oncology"[Journal] OR "JAMA oncology"[Journal] OR "Annals of oncology: official journal of the European Society for Medical Oncology"[Journal] OR "Journal of the National Cancer Institute"[Journal]
OPHTHALMOLOGY	OR "Ophthalmology"[Journal] OR "JAMA ophthalmology"[Journal] OR "The ocular surface"[Journal] OR "American journal of ophthalmology"[Journal] OR "Retina (Philadelphia, Pa.)"[Journal]
ORTHOPAEDICS	OR "The American journal of sports medicine"[Journal] OR "Osteoarthritis and cartilage"[Journal] OR "The Journal of bone and joint surgery. American volume"[Journal] OR "Journal of physiotherapy"[Journal] OR "Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy

	Association of North America and the International Arthroscopy Association"[Journal]
OTORHINOLARYNGOLOGY	OR "JAMA otolaryngology-- head & neck surgery"[Journal] OR "Ear and hearing"[Journal] OR "Rhinology"[Journal] OR "Journal of vestibular research: equilibrium & orientation"[Journal] OR "Hearing research"[Journal]
PARASITOLOGY	OR "Cell host & microbe"[Journal] OR "Trends in parasitology"[Journal] OR "PLoS pathogens"[Journal] OR "PLoS neglected tropical diseases"[Journal] OR "Advances in parasitology"[Journal]
PATHOLOGY	OR "Annual review of pathology"[Journal] OR "Acta neuropathologica"[Journal] OR "Modern pathology: an official journal of the United States and Canadian Academy of Pathology, Inc"[Journal] OR "Seminars in immunopathology"[Journal] OR "The Journal of pathology"[Journal]
PEDIATRICS	OR "JAMA pediatrics"[Journal] OR "Journal of the American Academy of Child and Adolescent Psychiatry"[Journal] OR "Pediatrics"[Journal] OR "Pediatric allergy and immunology: official publication of the European Society of Pediatric Allergy and Immunology"[Journal] OR "The Journal of adolescent health : official publication of the Society for Adolescent Medicine"[Journal]
PERIPHERAL VASCULAR DISEASE	OR "Circulation"[Journal] OR "Circulation research"[Journal]

	OR "Hypertension (Dallas, Tex.: 1979)"[Journal] OR "Stroke"[Journal] OR "Arteriosclerosis, thrombosis, and vascular biology"[Journal]
PRIMARY HEALTH CARE	OR "Annals of family medicine"[Journal] OR "The British journal of general practice: the journal of the Royal College of General Practitioners"[Journal] OR "Journal of the American Board of Family Medicine: JABFM"[Journal] OR "Primary care respiratory journal: journal of the General Practice Airways Group"[Journal] OR "BMC family practice"[Journal]
PSYCHIATRY	OR "World psychiatry: official journal of the World Psychiatric Association (WPA)"[Journal] OR "JAMA psychiatry"[Journal] OR "The lancet. Psychiatry"[Journal] OR "The American journal of psychiatry"[Journal] OR "Psychotherapy and psychosomatics"[Journal]
RADIOLOGY, NUCLEAR MEDICINE & MEDICAL IMAGING	OR "JACC. Cardiovascular imaging"[Journal] OR "European heart journal cardiovascular Imaging"[Journal] OR "European journal of nuclear medicine and molecular imaging"[Journal] OR "Radiology"[Journal] OR "Journal of nuclear medicine : official publication, Society of Nuclear Medicine"[Journal]
REHABILITATION	OR "Neurorehabilitation and neural repair"[Journal] OR "Journal of physiotherapy"[Journal] OR "IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society"[Journal]

	OR "Journal of neuroengineering and rehabilitation"[Journal] OR "Journal of neurologic physical therapy : JNPT"[Journal]
REPRODUCTIVE BIOLOGY	OR "Human reproduction update"[Journal] OR "Human reproduction (Oxford, England)"[Journal] OR "Fertility and sterility"[Journal] OR "Molecular human reproduction"[Journal] OR "Biology of reproduction"[Journal]
RESPIRATORY SYSTEM	OR "The Lancet. Respiratory medicine"[Journal] OR "American journal of respiratory and critical care medicine"[Journal] OR "The European respiratory journal"[Journal] OR "Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer"[Journal] OR "Thorax"[Journal]
RHEUMATOLOGY	OR "Annals of the rheumatic diseases"[Journal] OR "Arthritis & rheumatology (Hoboken, N.J.)"[Journal] OR "Osteoarthritis and cartilage"[Journal] OR "Rheumatology (Oxford, England)"[Journal] OR "Seminars in arthritis and rheumatism"[Journal]
SPORT SCIENCES	OR "British journal of sports medicine"[Journal] OR "Sports medicine (Auckland, N.Z.)"[Journal] OR "The American journal of sports medicine"[Journal] OR "Exercise and sport sciences reviews"[Journal] OR "Arthroscopy : the journal of arthroscopic & related surgery : official publication of the Arthroscopy Association of North America and the International Arthroscopy Association"[Journal]
SUBSTANCE ABUSE	OR "Alcohol research: current reviews"[Journal]

	OR "Addiction (Abingdon, England)"[Journal] OR "Nicotine & tobacco research: official journal of the Society for Research on Nicotine and Tobacco"[Journal] OR "The International journal on drug policy"[Journal] OR "Drug and alcohol dependence"[Journal]
SURGERY	OR "Annals of surgery"[Journal] OR "JAMA surgery"[Journal] OR "The Journal of heart and lung transplantation: the official publication of the International Society for Heart Transplantation"[Journal] OR "Journal of neurology, neurosurgery, and psychiatry"[Journal] OR "Endoscopy"[Journal]
TRANSPLANTATION	OR "The Journal of heart and lung transplantation: the official publication of the International Society for Heart Transplantation"[Journal] OR "American journal of transplantation: official journal of the American Society of Transplantation and the American Society of Transplant Surgeons"[Journal] OR "Nephrology, dialysis, transplantation: official publication of the European Dialysis and Transplant Association - European Renal Association"[Journal] OR "Bone marrow transplantation"[Journal] OR "Transplantation"[Journal]
TROPICAL MEDICINE	OR "PLoS neglected tropical diseases"[Journal] OR "Malaria journal"[Journal] OR "Memórias do Instituto Oswaldo Cruz"[Journal] OR "Transactions of the Royal Society of Tropical Medicine and Hygiene"[Journal] OR "The American journal of tropical medicine and hygiene"[Journal]
UROLOGY & NEPHROLOGY	OR "European urology"[Journal]

	OR "Journal of the American Society of Nephrology: JASN"[Journal] OR "Kidney international"[Journal] OR "American journal of kidney diseases: the official journal of the National Kidney Foundation"[Journal] OR "Clinical journal of the American Society of Nephrology : CJASN"[Journal]
VIROLOGY	OR "Cell host & microbe"[Journal] OR "PLoS pathogens"[Journal] OR "Current opinion in virology"[Journal] OR "Advances in virus research"[Journal] OR "Annual review of virology"[Journal]

Appendix 2: Data Extraction Form Draft

We will pilot test this draft based on 20 randomly selected studies among the one that will be included at the end of the full-text stage and hence it is likely to be modified.

General study characteristics

ID:

Journal:

Regions where the study was conducted:

- USA
- Europe
- Middle-East
- Africa
- South-America
- Asia
- Australia/New-Zealand

Medical Area:

Objective of the study

Development and validation (internal +/- external) of an algorithm / External Validation of an already developed algorithm / Clinical implementation of an algorithm (including pilot-testing) / Evaluation of an intervention based on an AI algorithm

Aim of the Algorithm

aim of the algorithm:

- Screening
- Triage
- diagnosis
- prognosis
- prediction of response
- prediction of side-effects
- other

Prediction Model? Y/N

Classification algorithm? Y/N

What does the model predict or class? Outcome

Dimension/nature of the output: Binary / Multiple classes / Regression (number) / Survival

Target population? Outpatient / Inpatient / Emergency / Primary Care / At Home

Gold standard: Study Experts / Follow-up / Routinely collected data

Design Characteristics

observational study / clinical trial

cross-sectional / prospective / retrospective study

Performance

Primary objective (accuracy / confusion matrices / ROC curves / calibration measure)

Accuracy:

Sensitivity:

Specificity:

AUROC, c-index:

95% CI lower:

95% CI upper:

Other:

Calibration assessed: Y/N

Calibration method:

Algorithm Characteristics

Type of algo:

- Deep Learning
- Gradient Boosting
- Regression
- Support Vector machines
- Random Forests
- Other:

Tuning of hyperparameters: not reported / default / manually tuned /automated tuning

Loss minimized: Mean squared Error (L2) / Binary cross-entropy / Categorical cross-entropy / Penalized Likelihood /Other
If transfer from existive algorithm Y/N or validation, name of the algorithm:

Training and the test datasets:
No of centers:
No of countries:
sample selection: consecutive patients Y/N
Sample size analyzed:
No. of unique patients:
type of data used:

- radiological images
- pathological images
- clinical data
- Data from an EHR
- Medical claim databases
- Other

Setting: Outpatient / Inpatient / Emergency / Primary Care / At Home
Included patients are representative of the patients in which the prediction model will be applied?: Y/N

Name of image bank for the input data:
Reporting the data preparation (e.g., image pre- processing) Y/N
Report on how information leakage was prevented?
Is there a flowchart reported?
Handling of missing data: not reported / complete cases / imputation / other

Method for internal validation: Splitting / cross-validation / bootstrap / other
External validation: Y/N

Transparency

whether the data are available.

whether the algorithm is available.

Limitations Reported

Details about limitations

Separate Heading: Y/N

Embedded within Discussion: Y/N

Size: no of words

Countermeasures taken / proposed to overcome limits

Conclusion of the paper:

- Algo > humans
- Algo + humans > humans
- Algo < humans
- Algo = humans

Implementation in clinic: yes / too early / no