Project summary

Created 19 October 2021

Last updated 5 November 2021

* clinicaltrials.gov is a publicly available, international registry of human clinical trials and intervention studies.
* Study investigators are responsible for registering their study, providing details about their proposed study design, outcome measures, patient recruitment and resulting publications
* Increasingly, studies on the development and validation of clinical prediction models are being registered on clinicaltrials.gov, covering a range of diseases and clinical outcomes
  + Some examples:

<https://clinicaltrials.gov/ct2/show/NCT02861547>

<https://clinicaltrials.gov/ct2/show/NCT03973437>

<https://clinicaltrials.gov/ct2/show/NCT04292158>

<https://clinicaltrials.gov/ct2/show/NCT04525287>

* The development of clinical prediction models requires careful consideration of several factors, including minimum sample sizes considerations, appraisal of model performance and the assessment of potential biases that could influence a model’s generalisability to new patient cohorts.
* Failure to adequately consider these factors during the planning phase is likely to severely limit model applicability in practice and inevitably leads to research waste.

The aim of this study is to evaluate the completion and publication of clinical prediction modelling studies registered on clinicaltrials.gov. Analyses will be based on the results of systematic database searches using predefined search strings. Initial search strings have been adapted from an ongoing review of diagnostic and prognostic models for COVID-19 (Wynants et al, BMJ, 2020). [To be finalized - requires further testing/manual review]:

titles=diagnostic OR prognostic OR prediction model OR machine learning OR artificial intelligence OR algorithm OR score OR deep learning OR regression

The above search query returned 12,770 studies.

Identified studies will be evaluated against the following criteria:

1. The number of studies marked as completed
2. The proportion of registered studies with a resulting publication on the proposed prediction model(s). Information on publications will be sourced from publications data entered by study investigators on clinicaltrials.gov and from search of PubMed for publications linked by NCT number (via metadata or [A Web-based Tool for Automatically linking Clinical Trials to their Publications | medRxiv](https://www.medrxiv.org/content/10.1101/2021.06.24.21259481v1))
3. Times from study registration to study completion; study completion to publication (time-to-event analysis)
4. Original estimated enrolment versus actual enrolment

Revised search terms

5/11/21:

Other terms: (Prediction OR Predictive OR Warning OR Scoring) AND (score OR model OR algorithm OR rule OR tool) AND (development OR establishment OR develop OR establish OR build OR create) AND (prognosis OR diagnosis OR prognostic OR diagnostic); returned 6,524 studies

25/11/21: Updated searches stratified by prognosis, diagnostic

|  |  |  |  |
| --- | --- | --- | --- |
| **Common search string** | **Prognostic**  AND (prognosis OR prognostic) | **Diagnostic**  AND (diagnosis OR diagnostic) | **Excluding prognostic/diagnostic terms** |
| (Prediction OR Predictive OR Warning OR Scoring) AND (score OR model OR algorithm OR rule) | 3,184 | 19,173 | 104,939 |
| ("Machine learning" OR "Artificial Intelligence" OR "Deep Learning" OR Prediction OR Predictive OR Warning OR Scoring) AND (score OR model OR algorithm OR rule) | 3,197 | 19,386 | 105,281 |
| **Combined terms for models** | | |  |
| (“Machine learning” OR “Artificial Intelligence” OR “Deep Learning” OR “Prediction model” OR “Predictive model” OR “Prediction score” OR “Predictive score” OR “warning score” OR “Risk prediction”) | 326 | 1,143 | 2,578 |
| ("Machine learning" OR "Artificial Intelligence" OR "Deep Learning" OR "Prediction model" OR "Predictive model" OR "Prediction score" OR "Predictive score" OR "warning score" OR "Risk prediction" OR "risk score") | 399 | 1,435 | 3,491 |
| **Add model development or establishment** | | |  |
| (Prediction OR Predictive OR Warning OR Scoring) AND (score OR model OR algorithm OR rule) AND (development OR establishment OR develop OR establish) | 940 | 5,332 | 20,810 |
| ("Machine learning" OR "Artificial Intelligence" OR "Deep Learning" OR Prediction OR Predictive OR Warning OR Scoring) AND (score OR model OR algorithm OR rule) AND (development OR establishment OR develop OR establish) | 947 | 5,439 | 20,974 |

First attempt to get code setup and working:

downloaded studies found using: ("Machine learning" OR "Artificial Intelligence" OR "Deep Learning" OR "Prediction model" OR "Predictive model" OR "Prediction score" OR "Predictive score" OR "warning score" OR "Risk prediction" OR "risk score") AND (prognosis OR prognostic); 399 studies