

Trends in the development and validation of patient-level prediction models using electronic health record data: a systematic review

PRESENTER: Cynthia Yang

INTRO:

- The aim of this systematic review is to provide further insights in the development of the field over time, with a focus on the transparent reporting of model development and validation using electronic health record (EHR) data to enable external validation by other investigators.

METHODS:

- We searched Embase, Medline, Web-of-Science, Cochrane Library and Google Scholar. The search was limited to papers written in English and published between January 1, 2009, and November 15, 2019.
- We included all papers that described the development of one or more multivariable prognostic prediction models using EHR data.
- To investigate trends, we assessed differences in items between the periods 2009-2014 and 2015-2019.

RESULTS:

- Our literature search resulted in a total of 9,942 papers. After deduplication, 6,235 titles and abstracts were screened. From this, 1,075 potentially eligible papers were identified. Upon full text inspection, 422 papers were eventually included for synthesis. In total, we extracted items for 579 models from 422 papers (1 to 6 models per paper). We observed an increase from 135 models in 101 papers in the period 2009-2014 to 444 models in 321 papers in the period 2015-2019.

We found limited improvement in the methodological conduct and reporting of prognostic model development and validation using EHR data in the period 2009-2019.

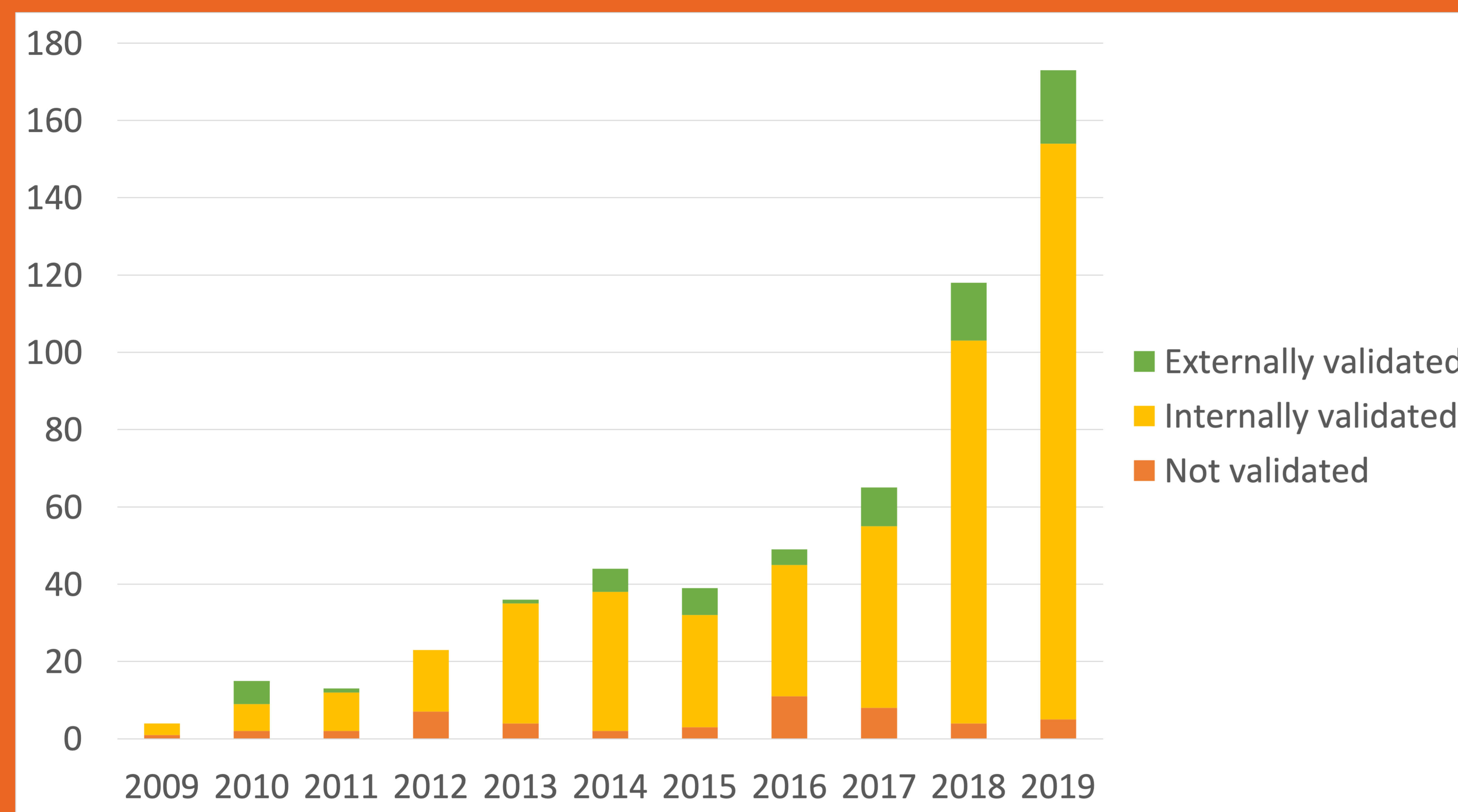
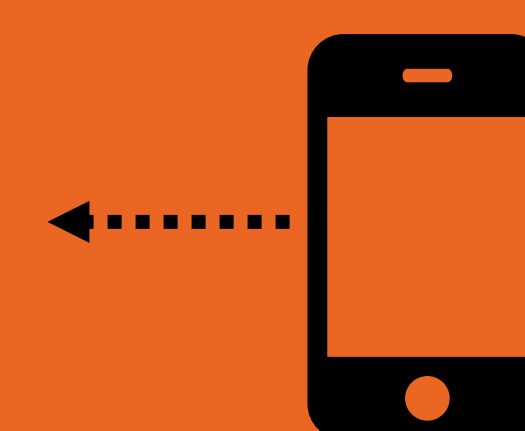


Figure 1. Reporting of validation results



Scan the QR code to access the brief report.

Main findings:

- The percentage of models for which code lists were provided to define the target population, outcome, and candidate predictors was very low and remained below 20% over time. In both periods, the prediction horizon was reported for 84% of all models. The percentage of models for which the time window for candidate predictor measurement was reported increased from 46% to 50%, while the percentage of models for which the final model was completely presented decreased from 49% to 39%.
- External validation increased from 10% to 12%, internal validation only increased from 76% to 81%, and no validation decreased from 13% to 7% (see Figure 1). The percentage of externally validated models that were validated using data from a different country increased from 7% to 9%.

Cynthia Yang, MSc,
Jan A. Kors, PhD,
Solomon Ioannou, MSc,
Luis H. John, MSc,
Aniek F. Markus, MSc,
Alexandros Rekkas, MSc,
Maria de Ridder, PhD,
Tom Seinen, MSc,
Ross D. Williams, MSc,
Peter R. Rijnbeek, PhD



This project has received funding from the Innovative Medicines Initiative 2 Joint Undertaking (JU) under grant agreement No 806968. The JU receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA.

