EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

ERJ methods

Developing, validating, updating and judging the impact of prognostic models for respiratory diseases

Florien S. van Royen, Karel G.M. Moons, Geert-Jan Geersing, Maarten van Smeden

Please cite this article as: van Royen FS, Moons KGM, Geersing G-J, *et al*. Developing, validating, updating and judging the impact of prognostic models for respiratory diseases. *Eur Respir J* 2022; in press (https://doi.org/10.1183/13993003.00250-2022).

Developing, validating, updating and judging the impact of prognostic models for respiratory diseases

Florien S. van Royen[1], Karel G.M. Moons[2], Geert-Jan Geersing[1], Maarten van Smeden[2]*

[1]Dept. General Practice, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

[2]Dept. Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

*Corresponding author

Introduction

Prognostic models predict an individual's risk (i.e. probability) of future health outcomes, such as the development of respiratory disease, complications due to disease progression, intended or adverse outcomes of treatments, or any other health outcome.(1–3) In clinical practice, prognostic models are used to inform patients and healthcare providers about likely health outcomes and to guide shared medical decision-making in initiating treatment, preventive strategies or watchful waiting.(4)

Before implementing a newly developed prognostic model in clinical practice, it is generally accepted that the prediction model needs, at the very least, to be evaluated on its performance (i.e. to be validated) on other individuals than from which it was developed. Additionally, prognostic models developed in another setting may also need to be updated for the application setting at hand to better tailor its predictions to that setting. Ideally, a prognostic model is also assessed on its impact to improve decision making and patient outcomes.(1,5–8)

While many prognostic models for relevant health outcomes and conditions can be found in the literature, e.g. over 400 models for COPD progression (9), 37 on treatment response in pulmonary tuberculosis (10) and 17 on the prognosis of acute pulmonary embolism (11), only few models reach the validation or implementation phases.(3,6,12,13) Reasons for failed implementation of prediction models are illustrated in Figure 1. With such a large pool of already existing prognostic models, we argue that – in general – the first step of prognostic model research should be a critical appraisal and validation of existing prognostic models, before embarking on the development of a new prognostic model.(14) Following this line of reasoning, in this paper we will first discuss critical appraisal and validation of existing prediction models in the respiratory disease domain, before discussing aspects of model development.

Reviewing and critical appraisal of existing prognostic models

A first natural step in prognostic model research is to search, review, critically appraise and externally validate already existing prediction models. For instance, a systematic review and critical appraisal on existing COVID-19 models revealed over 100 prognostic models.(15) All

models were critically appraised using the Prediction model Risk Of Bias ASsessment Tool (PROBAST), to assess the applicability and risk of bias of prediction model studies in the intended population and context.(16,17). Most of these prognostic COVID-19 models, usually with a mortality or clinical deterioration outcome, were judged to be at high risk of bias, for which the main reasons were the use of inappropriate data sources, inadequate low sample sizes, inappropriate statistical model evaluations and overall poor reporting. Moreover, for 22 prognostic models that provided sufficient information to be externally validated, almost all showed poor predictive performance in a dataset from London hospitals.(18)

Tools for designing, searching, critical appraisal, data extraction and interpretation of existing prognostic models in systematic reviews, are easily available with hand on guidance at the Cochrane Prognosis Review Methods Group website (www.methods.cochrane.org/prognosis) and at the PROBAST website (www.probast.org).(16,19)

External validation and model updating

External validation is the evaluation of the predictive performance of an existing prognostic model in new patients.(6,8,13,20,21) Predictive performance of a prognostic model is usually assessed by the *calibration*, typically investigated by visual inspection of a calibration curve and quantified by calibration statistics such as the calibration slope and calibration-in-the-large, expressing how well the predicted risks correspond to the actual observed outcome frequencies in the validation dataset, and *discrimination*, typically quantified by a concordance statistic expressing how well the predicted risks distinguish between those who experience the outcome and those who do not in the validation dataset.(22,23) A detailed description of these and various other performance metrics for prediction models, including decision curve analysis, are described by Steyerberg et al.(24) Methods for calculating the minimum sample size required for external validation studies have recently been suggested.(25–27)

External validation of existing prognostic models in new populations or settings also motivates updating of the prognostic model to the new situation at hand, especially when model miscalibration (i.e. the predicted risks do not correspond to the actual observed risks in a new population) is detected or model discrimination is less than anticipated. In these

situations, updating of underperforming existing models improves their performance in the validation setting and can increase the likelihood of transportability to other settings.(8) Updating of a prediction model that is based on a regression model, such as a logistic regression or a Cox regression model, often starts with recalibration of the intercept or baseline hazard (i.e. re-estimating the intercept or baseline hazard using the validation data set at hand, which are important for accurate estimates of risk), but could require several more steps including the complete re-estimation of all the coefficients in the model.(8,28,29)

When the validation of an existing prognostic model in a new setting did not provide satisfying results, it is generally recommended to first undertake model updating to determine to what extent the existing model can be easily improved, before developing an entirely new prognostic model using the validation dataset only.(2,14) Repeated external validation and updating of a prognostic model may often be necessary to ensure the model remains relevant over time and place, meaning that validation and updating should not be considered a one-time activity.(30)

Impact and effectiveness assessment

When validation studies with updating (if indicated) show indeed sufficient predictive performance of a prognostic model (where what is considered "sufficient" is context specific and to some degree subjective), a natural next step is to empirically test its applicability, effectiveness and impact. Such empirical effectiveness studies are often referred to as comparative prediction model impact studies.(5,6,8,31) Good predictive performance of a prognostic model in validation studies is unfortunately no guarantee for usefulness and effectiveness of the model in clinical practice. An example of such model impact study is the VISTA trial that compares the use of the Vienna risk prediction model combined with risk-tailored management with usual care in patients with venous thromboembolism. The discrimination was judged to be "good" (c-statistic 0.76) and calibration was judged as "moderate" to "good", with best calibration performance in the lower predicted risk groups. However, despite the promising predictive performance, no evidence was found for risk reduction in venous thromboembolism recurrence (the primary outcome) in the risk-tailored management applying the Vienna risk prediction model in the VISTA trial.(32,33)

In contrast to validation studies in which the model's predictive performance is assessed in a single cohort, model impact studies typically have a comparative nature, where the outcome is not the predictive accuracy of the model, but the difference in effect of actual use of a model as compared to not using it. Usually, a group of individuals, e.g. clinicians, using the newly developed prognostic model versus a group of similar individuals not using that prognostic model, i.e. following common clinical practice, are compared on decision-making, behavioural changes, patient outcomes and/or cost-effectiveness. The choice of outcome may partly be based on whether the prognostic model is meant to be assistive (e.g. a prognostic model aim to inform about risk without explicit treatment guidance) or directive (e.g. a prognostic model with explicit treatment guidance).

The ideal design for such a comparative study is a cluster randomised trial where the prognostic model is only available in randomised clusters (e.g. healthcare professionals, practices or hospitals). Randomisation of clusters is often preferred over randomising individual patients because in the latter, a learning effect may influence the results of the trial.(8,31) Non-randomised designs for prediction model impact studies are also possible, for instance a prospective before-after study, although this type of design is prone to time and thus confounding effects.(8,31) Assessing effectiveness of prediction models may even call for mixed-methods designs, including also qualitative approaches such as interviewing of physicians and patients that actually used the model during the study.(34)

Decision analytic modelling studies can assist in estimating the possible (cost-)effectiveness by combining information from various sources, including simulation of hypothetical clinical scenarios based on all the available evidence.(35,36) If the results are negative, i.e. the implementation or use of a prognostic model does not seem to improve the (cost-)effectiveness, the prognostic model may require revision or updating before further steps towards prospective implementation or impact studies are taken.(8,34)

Developing a new prognostic model

Reasons for developing a new prognostic model, ideally as discussed above after validation and or updating of existing prediction models, are considered in Box 1. In general, the first consideration should be the availability and selection of a relevant and large enough patient

population from which data of predictors and outcomes are registered. In general, the selected patients should reflect the target domain or setting in which the model is intended to be used.(6,7,37) A too small development sample size may create a prediction model that generates imprecise predictions and risks *overfitting*, which means that the prediction model is likely to perform poorly on individuals who were not part of the dataset used to develop the model. Recent studies have shown that overfitting is difficult to overcome even when using modern methods.(38,39) To avoid overfitting and ensure precise predictions, considerations on sample size and model complexity must be made during the design phase of the study. Guidance for these considerations in the form of minimal sample size criteria have been proposed, with simple to apply sample size calculators currently available in the statistical programs *R* and *Stata*.(40–42)

Selecting the relevant predictors to include in the prediction model is another crucial step in prognostic model development. Ideally, the pre-selection of potential predictors should focus on those variables that, prior to data collection, are known to be related to the prognostic outcome, based on a combination of clinical expertise and evidence from literature, for instance through prognostic factor studies.(43) Beyond prior evidence and statistical criteria (i.e. statistical evidence that the predictor has incremental predictive value over other predictors), the selection of predictors should also consider practical implications of predictors. For instance, prognostic models developed using invasive or expensive to measure predictors may create a model that will become largely inapplicable to patients for whom the model was intended.

The success of developing a relevant, applicable and well-performing prognostic model depends on many other factors beyond those already mentioned. The outcome to be predicted should be clinically useful and measured as precisely as possible for all individuals. Patients for whom the outcome is not observed, due to drop-out, being at risk of developing the outcome at the end of the study period or so-called competing risks, should be handled using the appropriate statistical techniques, for instance through survival analyses.(29) Handling of missing data and modelling of continuous predictors while avoiding wasting important prognostic information using unnecessary categorization, are other important

aspects of prognostic modelling.(44–46) For a more thorough discussion on prognostic model development steps we refer to the already existing literature.(5,21,29,37)

Concluding remarks

A useful clinical prediction model is a model that predicts accurately (good performance as shown in calibration and discrimination), has been shown to work in settings other than its development setting (is externally validated) and is of value to clinicians by improving relevant patient outcomes (impact).

Developing new prediction models is often not difficult, but deriving a clinical useful model that stands the test of time exemplified by satisfactory validation studies, proves to be more challenging. It should also be noted that recent regulatory developments, in particular medical device regulation in the EU and the US on required regulatory approval for prediction models (notably for those developed using artificial intelligence methodology), may impact on the requirements, such as results from impact studies or decision analytic modelling, for deriving a clinically useful prognostic model.

A key aspect of developing, validating, updating and assessing the impact of a model is its transparent reporting. The TRIPOD guideline for developing, validating and updating prognostic models has been widely used.(2) Updates of this guideline, including a version specifically designed for application of prognostic modelling based on machine learning and artificial intelligence, are soon to be expected.(47)

References

1. Moons KGM, Royston P, Vergouwe Y, Grobbee DE, Altman DG. Prognosis and prognostic research: What, why, and how? BMJ. 2009;338(7706):1317–20.
2. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): Explanation and elaboration. Ann Intern Med. 2015;162(1):W1–73.
3. van Smeden M, Reitsma JB, Riley RD, Collins GS, Moons KG. Clinical prediction models: diagnosis versus prognosis. J Clin Epidemiol. 2021;132:142–5.
4. Hingorani AD, Van Der Windt DA, Riley RD, Abrams K, Moons KGM, Steyerberg EW, et al. Prognosis research strategy (PROGRESS) 4: Stratified medicine research. BMJ. 2013;346(February):1–9.
5. Riley R, van der Windt D, Croft P, Moons K. Prognosis research in healthcare: concepts, methods, and impact. Oxford University Press; 2019.
6. Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis Research Strategy (PROGRESS) 3: Prognostic Model Research. PLOS. 2013;10(2):e1001381.
7. Moons KGM, Kengne AP, Woodward M, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker. Heart. 2012;98(9):683–90.
8. Moons KGM, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, et al. Risk prediction models: II. External validation, model updating, and impact assessment. Heart. 2012;98(9):691–8.
9. Bellou V, Belbasis L, Konstantinidis AK, Tzoulaki I, Evangelou E. Prognostic models for outcome prediction in patients with chronic obstructive pulmonary disease: Systematic review and critical appraisal. BMJ. 2019;367.
10. Peetluk LS, Ridolfi FM, Rebeiro PF, Liu D, Rolla VC, Sterling TR. Systematic review of prediction models for pulmonary tuberculosis treatment outcomes in adults. BMJ Open. 2021;11(3).
11. Elias A, Mallett S, Daoud-Elias M, Poggi JN, Clarke M. Prognostic models in acute pulmonary embolism: A systematic review and meta-analysis. BMJ Open. 2016;6(4).
12. Bouwmeester W, Zuithoff NPA, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and methods in clinical prediction research: A systematic review. PLoS

Med. 2012;9(5).

13. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med. 2000;19(4):453–73.
14. Collins GS, Moons KGM. Comparing risk prediction models: Should be routine when deriving a new model for the same purpose. BMJ. 2012;344(7859):1–2.
15. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. BMJ. 2020;369.
16. 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;170(1):51–8.
17. Moons KGM, Wolff RF, Riley RD, Whiting PF, Westwood M, Collins GS, et al. PROBAST: A tool to assess risk of bias and applicability of prediction model studies: Explanation and elaboration. Ann Intern Med. 2019;170(1):W1–33.
18. Gupta RK, Marks M, Samuels THA, Luintel A, Rampling T, Chowdhury H, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: An observational cohort study. Eur Respir J [Internet]. 2020;56(6). Available from: http://dx.doi.org/10.1183/13993003.03498-2020
19. Moons KGM, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG, et al. Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist. PLoS Med. 2014;11(10).
20. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: Seven steps for development and an ABCD for validation. Eur Heart J. 2014;35(29):1925–31.
21. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation and Updating. Vol. 66, Biometrics. Springer; 2009.
22. Van Calster B, McLernon DJ, Van Smeden M, Wynants L, Steyerberg EW, Bossuyt P, et al. Calibration: The Achilles heel of predictive analytics. BMC Med. 2019;17(1):1–7.
23. Altman DG, Vergouwe Y, Royston P, Moons KGM. Prognosis and prognostic research: Validating a prognostic model. BMJ. 2009;338(7708):1432–5.
24. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: A framework for traditional and novel measures. Epidemiology. 2010;21(1):128–38.

25. Archer L, Snell KIE, Ensor J, Hudda MT, Collins GS, Riley RD. Minimum sample size for external validation of a clinical prediction model with a continuous outcome. Stat Med. 2021;40(1):133–46.
26. Riley RD, Debray TPA, Collins GS, Archer L, Ensor J, van Smeden M, et al. Minimum sample size for external validation of a clinical prediction model with a binary outcome. Stat Med. 2021;40(19):4230–51.
27. Riley RD, Collins GS, Ensor J, Archer L, Booth S, Mozumder SI, et al. Minimum sample size calculations for external validation of a clinical prediction model with a time-to-event outcome. Stat Med. 2021;(May):1–16.
28. Vergouwe Y, Nieboer D, Oostenbrink R, Debray TPA, Murray GD, Kattan MW, et al. A closed testing procedure to select an appropriate method for updating prediction models. Stat Med. 2017;36(28):4529–39.
29. Harrell Jr FE. Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis. Springer. 2015.
30. Jenkins DA, Martin GP, Sperrin M, Riley RD, Debray TPA, Collins GS, et al. Continual updating and monitoring of clinical prediction models: time for dynamic prediction systems? Diagnostic Progn Res. 2021;5(1):1–7.
31. Moons KGM, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: Application and impact of prognostic models in clinical practice. BMJ. 2009;338(7709):1487–90.
32. Eichinger S, Heinze G, Jandeck LM, Kyrle PA. Risk assessment of recurrence in patients with unprovoked deep vein thrombosis or pulmonary embolism: The vienna prediction model. Circulation. 2010;121(14):1630–6.
33. Geersing GJ, Hendriksen JMT, Zuithoff NPA, Roes KC, Oudega R, Takada T, et al. Effect of tailoring anticoagulant treatment duration by applying a recurrence risk prediction model in patients with venous thromboembolism compared to usual care: A randomized controlled trial. PLoS Med [Internet]. 2020;17(6):1–17. Available from: http://dx.doi.org/10.1371/journal.pmed.1003142
34. Kappen TH, van Klei WA, van Wolfswinkel L, Kalkman CJ, Vergouwe Y, Moons KGM. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. Diagnostic Progn Res. 2018;2(1):1–11.
35. Schaafsma JD, van der Graaf Y, Rinkel GJE, Buskens E. Decision analysis to complete

diagnostic research by closing the gap between test characteristics and cost-effectiveness. J Clin Epidemiol [Internet]. 2009;62(12):1248–52. Available from: http://dx.doi.org/10.1016/j.jclinepi.2009.01.008

36. Jenniskens K, Lagerweij GR, Naaktgeboren CA, Hooft L, Moons KGM, Poldervaart JM, et al. Decision analytic modeling was useful to assess the impact of a prediction model on health outcomes before a randomized trial. J Clin Epidemiol [Internet]. 2019;115:106–15. Available from: https://doi.org/10.1016/j.jclinepi.2019.07.010
37. Royston P, Moons KGM, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ. 2009;338(7707):1373–7.
38. Van Calster B, van Smeden M, De Cock B, Steyerberg EW. Regression shrinkage methods for clinical prediction models do not guarantee improved performance: Simulation study. Stat Methods Med Res. 2020;29(11):3166–78.
39. Riley RD, Snell KIE, Martin GP, Whittle R, Archer L, Sperrin M, et al. Penalization and shrinkage methods produced unreliable clinical prediction models especially when sample size was small. J Clin Epidemiol [Internet]. 2021;132:88–96. Available from: https://doi.org/10.1016/j.jclinepi.2020.12.005
40. Riley RD, Ensor J, Snell KIE, Harrell FE, Martin GP, Reitsma JB, et al. Calculating the sample size required for developing a clinical prediction model. BMJ [Internet]. 2020;368(March):1–12. Available from: http://dx.doi.org/doi:10.1136/bmj.m441
41. van Smeden M, Moons KGM, de Groot JAH, Collins GS, Altman DG, Eijkemans MJC, et al. Sample size for binary logistic prediction models: Beyond events per variable criteria. Stat Methods Med Res. 2019;28(8):2455–74.
42. Riley RD, Snell KIE, Ensor J, Burke DL, Harrell FE, Moons KGM, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. Stat Med. 2019;38(7):1276–96.
43. Riley RD, Hayden JA, Steyerberg EW, Moons KGM, Abrams K, Kyzas PA, et al. Prognosis research strategy (PROGRESS) 2: Prognostic factor researh. PLoS Med. 2013;10(2):e1001380.
44. Sauerbrei W, Perperoglou A, Schmid M, Abrahamowicz M, Becher H, Binder H, et al. State of the art in selection of variables and functional forms in multivariable analysis—outstanding issues. Diagnostic Progn Res. 2020;4(1).
45. Sperrin M, Martin GP, Sisk R, Peek N. Missing data should be handled differently for

prediction than for description or causal explanation. J Clin Epidemiol [Internet]. 2020;125:183–7. Available from: https://doi.org/10.1016/j.jclinepi.2020.03.028

46. van Smeden M, Groenwold RHH, Moons KG. A cautionary note on the use of the missing indicator method for handling missing data in prediction research. J Clin Epidemiol [Internet]. 2020;125:188–90. Available from: https://doi.org/10.1016/j.jclinepi.2020.06.007
47. Collins GS, Dhiman P, Andaur Navarro CL, Ma J, Hooft L, Reitsma JB, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open. 2021;11(7):1–7.
48. Celli BR, Cote CG, Marin JM, Casanova C, Montes de Oca M, Mendez RA, et al. The Body-Mass Index, Airflow Obstruction, Dyspnea, and Exercise Capacity Index in predicting hospitaliZation for Chronic Obstructive Pulmonary Disease. N Engl J Med. 2004;350(10):1005–12.
49. Puhan MA, Garcia-Aymerich J, Frey M, ter Riet G, Antó JM, Agustí AG, et al. Expansion of the prognostic assessment of patients with chronic obstructive pulmonary disease: the updated BODE index and the ADO index. Lancet [Internet]. 2009;374(9691):704–11. Available from: http://dx.doi.org/10.1016/S0140-6736(09)61301-5
50. Lim WS, Van Der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, et al. Defining community acquired pneumonia severity on presentation to hospital: An international derivation and validation study. Thorax. 2003;58(5):377–82.
51. Fine MJ, Auble TE, Yealy DM, Hanusa BH, Weissfeld LA, Singer DE, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med. 1997;336(4):243–50.
52. Jiménez D, Aujesky D, Moores L, Gómez V, Lobo JL, Uresandi F, et al. Simplification of the pulmonary embolism severity index for prognostication in patients with acute symptomatic pulmonary embolism. Arch Intern Med. 2010;170(15):1383–9.
53. Zondag W, Mos ICM, Creemers-Schild D, Hoogerbrugge ADM, Dekkers OM, Dolsma J, et al. Outpatient treatment in patients with acute pulmonary embolism: The Hestia Study. J Thromb Haemost. 2011;9(8):1500–7.
54. Clift AK, Coupland CAC, Keogh RH, Diaz-Ordaz K, Williamson E, Harrison EM, et al. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality

from coronavirus 19 in adults: national derivation and validation cohort study. BMJ. 2020;371:1–20.

55. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: Development and validation of the 4C Mortality Score. BMJ. 2020;370(September):1–13.
56. Castro-Rodríguez JA, Holberg CJ, Wright AL, Martinez FD. A clinical index to define risk of asthma in young children with recurrent wheezing. Am J Respir Crit Care Med. 2000;162(4 I):1403–6.
57. Ley B, Ryerson CJ, Vittinghoff E, Ryu JH, Tomassetti S, Lee JS, et al. A multidimensional index and staging system for idiopathic pulmonary fibrosis. Ann Intern Med. 2012;156(10):684–95.

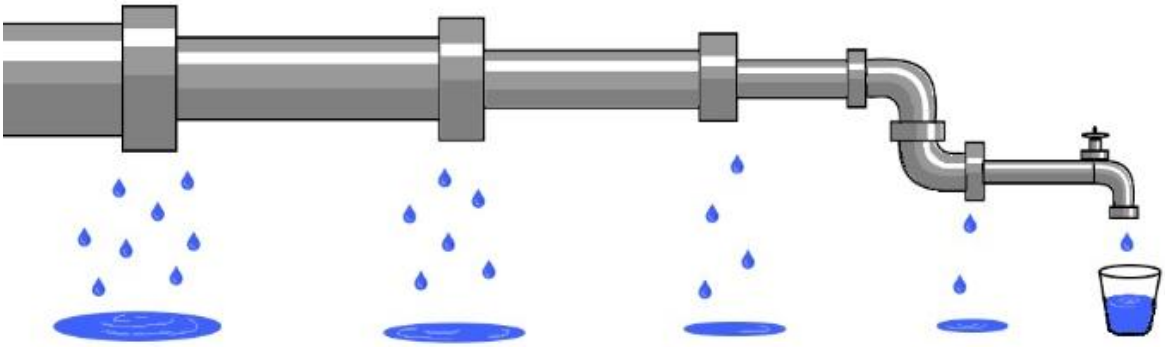
Box 1 - Examples of situations in which to move from validating existing models to developing a new prognostic model

When to proceed from validating existing models to developing a new model?

Reasons to start developing a new prognostic model are for instance when:

- There is no prognostic model currently available for the same outcome and patient population
- There are prognostic models available for the same outcome and patient population, but these existing models:
  - are insufficiently reported to be applied or validated, e.g. no intercept, baseline hazard or regression coefficients presented.
  - have been developed in a patient population that is substantially different from the intended population, e.g. a very selective set of the intended population.
  - use predictors that are expensive, impractical or use predictor definitions that are different than usual in the setting in which the model is intended to be used, e.g. based on advanced imaging techniques for a prognostic model intended to be used in primary care.
  - use predictors that are incompatible with time of prediction, e.g. post-operative predictors for a prognostic model developed prior to operations.
  - predict an outcome that was measured using unreliable methods.
  - showed poor validation performance overall or in important groups of individuals and updating of the model did not improve that performance.
  - have been shown to have no impact or a negative impact on patient outcomes in impact studies.

Figure 1 – Leaky prognostic model adoption pipeline. Examples of reasons for failed prediction model adoption in clinical practice.



Not fit for purpose	No validation	No implementation	Not adopted
Developed on wrong patient population	Lack of data or incentive to pursue validation studies	No impact on decision making or patient (health) outcomes	Prediction (perceived as) not useful
Expensive or non-available predictors	Incompletely reported prediction model	No software developed to implement and use the model	Predictions not trusted
Time intensive to use model	Poorly developed or overfitted model	Requirements for adherence to (medical device) regulations	Model not transparent enough, or no tools available to enhance its use in practice
Outcome measured unreliably	Proprietary model code	Cost(-effectiveness) of use proprietary model	Model (perceived as) outdated