



Download the workshop materials

# Introduction to systematic reviews of prognosis studies

**Carl Moons & Anneke Damen**  
on behalf of the Cochrane Prognosis Methods Group

Trusted evidence.  
Informed decisions.  
**Better health.**



# Agenda for today

10:00-11:30 Introduction to prognosis reviews

11:30-12:00 Risk of bias (lecture)

*12:00-12:45 Lunch*

12:45-14:45 Risk of bias (practical)

*14:45-15:15 Break*

15:15-16:30 Meta-analysis



# Introduction, design and protocol for systematic reviews of prognostic studies



# Outline

## Presentations:

- Introduction to types of prognosis research
- Introduction to types of SR of prognosis studies
- Defining the review question
- Data extraction and critical appraisal

Lecture + practicals



## Systematic reviews (SRs)

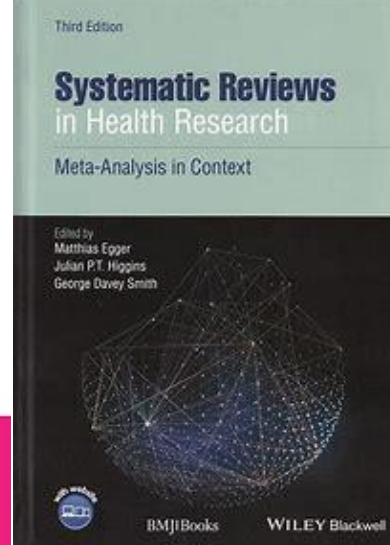
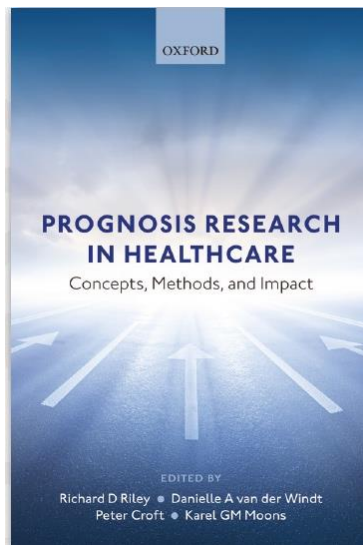
- Applicable to all fields of medical research
  - Therapeutic studies (RCTs): Cochrane Intervention Reviews
  - Diagnostic accuracy studies: Cochrane Diagnostic Test Accuracy Reviews
- Next: prognosis studies



# Why?

- Essence of precision, personalized or risk-based medicine
- Booming number of primary prognosis studies
  - Biomarkers, prognostic factors, models, algorithms
  - Further increase with introduction of AI/ML
- Reviews prognosis studies sharply increased past decade
  - Aggregate and IPD reviews
  - Cochrane library: Ongoing or published prognosis reviews = 42
- Reviews of prognosis studies ‘more’ challenging:
  - More variation in types of questions, designs, effect measures, analyses
  - But many recent method developments -> Cochrane PMG community





#### Welcome!

The **Cochrane Prognosis Methods Group (PMG)** focusses on the development of methods and guidance for performing Cochrane reviews of prognosis studies.

On this website you can find information about who we are, what guidance and tools are available, the training we offer and which reviews are ongoing. If you have any questions, please do not hesitate to [contact us](#).

For membership, please sign up to 'Join Cochrane' here or on the top navigation bar. On the webform, indicate whether you wish to become a full, active member or sign up to news and information only.

#### Planning to conduct a Cochrane prognosis review?

Check our PMG review process which contains important information for authors and CRGs



## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prognostic factor studies

Richard D Riley,<sup>1\*</sup> Karel G M Moons,<sup>2,4\*</sup> Kym I E Snell,<sup>1</sup> Joie Ensor,<sup>1</sup> Lotty Hooft,<sup>2,4</sup> Douglas G Altman,<sup>3</sup> Jill Hayden,<sup>5</sup> Gary S Collins,<sup>3</sup> Thomas P A Debray<sup>2,4</sup>

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prediction model performance

Thomas P A Debray,<sup>1,2</sup> Johanna A A G Damen,<sup>1,2</sup> Kym I E Snell,<sup>3</sup> Joie Ensor,<sup>3</sup> Lotty Hooft,<sup>1,2</sup> Johannes B Reitsma,<sup>1,2</sup> Richard D Riley,<sup>3</sup> Karel G M Moons<sup>1,2</sup>

## Group exercise – 10 minutes

1. What is prognosis?
2. Why do we prognosticate?
3. Types of prognosis studies?





## What is prognosis?

**Forecast** of the **course** and **outcome** for an **individual** in a **certain health state**

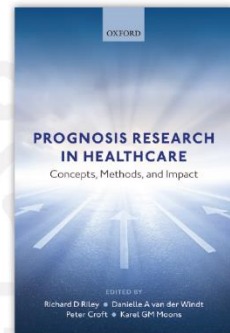
- Not necessarily sick people
- More technical: probable course/prediction of specific future outcomes in subjects with certain health condition or within a certain health state
- Disease does not have a prognosis → an individual does

**PROGNOSIS RESEARCH  
IN HEALTHCARE**  
Concepts, Methods, and Impact

**Edited by**

Richard D. Riley  
Danielle van der Windt  
Peter Croft  
Karel G.M. Moons

Paperback | 9780198796619  
January 2019 | 372 pages  
£44.99 **£31.49** | \$65.00 **\$45.50**



## Why prognosticate:

- To provide information to patients/individuals
- Identify groups for treatment or other management (e.g. lifestyle) – including abstine of management
- To target specific prognostic factors that modify treatment effects
- Select high/low risk patients for inclusion in RCTs
- Adjust for case-mix differences in comparison of healthcare institutes (benchmarking)



## Types of prognosis studies?

1. Average/overall prognosis: 'What is most likely course (outcome) of individuals with certain health condition?'
2. Prognostic factor studies: 'Which factors are associated with specific outcome in individuals with certain health condition?'
3. Prognostic modeling studies: 'What combination of prognostic factors predict, and how well, a certain outcome in individuals with a certain health condition?'

## Types of prognosis studies

1. Average/overall prognosis: 'What is most likely course (outcome) of individuals with certain health condition?'
2. Prognostic factor studies: 'Which factors are associated with specific outcome in individuals with certain health condition?'
3. Prognostic modeling studies: 'What combination of prognostic factors predict, and how well, a certain outcome in individuals with a certain health condition?'



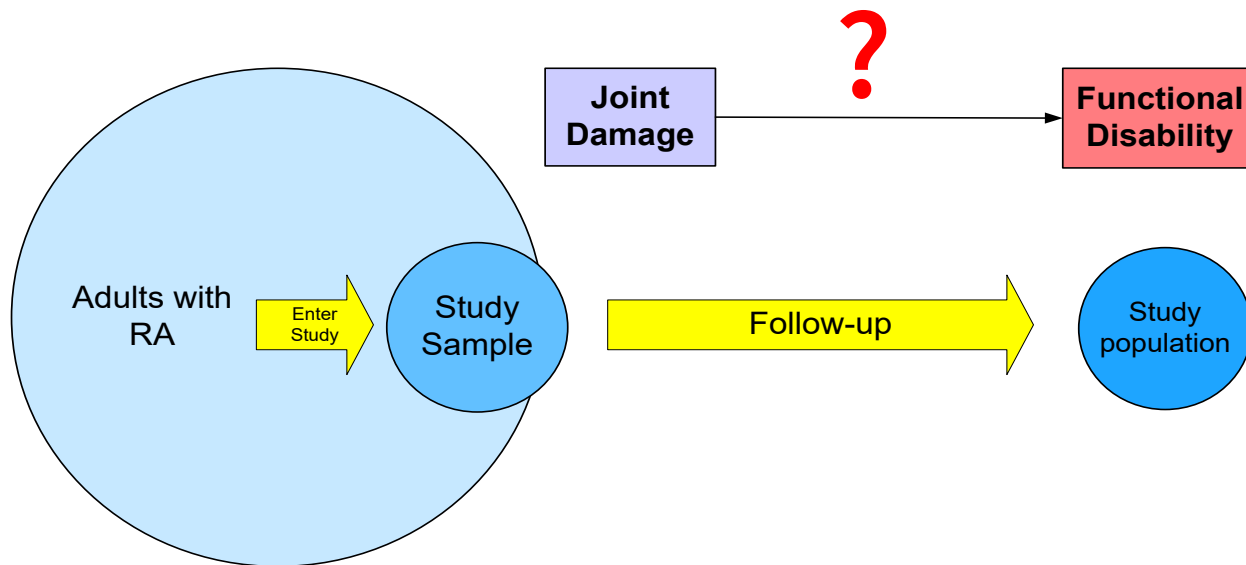
# Prognostic factor studies

Aim:

- To identify factors associated with subsequent outcomes in subjects with certain health condition
- Not necessarily sick people (patients)
- Independent predictors



# Prognostic Factor Study Example



## Types of prognosis studies

1. Average/overall prognosis: 'What is most likely course (outcome) of individuals with certain health condition?'
2. Prognostic factor studies: 'Which factors are associated with specific outcome in individuals with certain health condition?'
3. Prognostic modeling studies: 'What combination of prognostic factors predict, and how well, a certain outcome in individuals with a certain health condition?'



## **Exercise 10 min**

# Prognostic Model Studies

1. What is a prognostic model study, and what is the difference with multivariable analysis of multiple prognostic factors?
2. There are three phases of prediction modelling – which three?





## Prediction Model

Combination of 2 or more predictors in some kind of algorithm/formula that convert predictor values into an absolute probability of ...

...(presence of disease/result of reference test – diagnostic prediction model)

...future occurrence of certain outcome – prognostic prediction models

A prediction model is developed for use in new individuals to estimate their individual (diagnostic or prognostic) probability. Focus is on accuracy of entire model (discrimination + calibration). Predictors in the model not main interest.

Multivariable analysis of prognostic factors not focus on model, but rather on which are the independent predictors – Focus on HRs of the factors (adjusted HRs)

## What is the difference between 3 versus 1 and 2?

## 3 Phases of Prediction Modelling studies

Big difference = 3 are comparative studies → ideally randomised

1 and 2 are by definition single cohort studies- no inherent comparison

3 are thus ideally RCTs – for SRs of prediction model impact studies use the Cochrane tools available for RCTs of intervention studies

**Everything we say from here on also applies to  
SRs of diagnostic prediction model studies**

**You need no separate course for that!  
We use generic term: prediction model**

**Interesting and booming field – stay in it!**

# Conducting a systematic review of prognosis studies

1. Formulate review question (PICOTS)
2. Searching for studies
3. Screening and Selection of articles
4. Extraction of data
5. Risk of Bias assessments
6. Synthesis of data (meta-analysis)
7. Interpretation and conclusions

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prognostic factor studies

Richard D Riley,<sup>1\*</sup> Karel G M Moons,<sup>2,4\*</sup> Kym I E Snell,<sup>1</sup> Joie Ensor,<sup>1</sup> Lotty Hooft,<sup>2,4</sup>  
Douglas G Altman,<sup>3</sup> Jill Hayden,<sup>5</sup> Gary S Collins,<sup>3</sup> Thomas P A Debray<sup>2,4</sup>

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prediction model performance

Thomas P A Debray,<sup>1,2</sup> Johanna A A G Damen,<sup>1,2</sup> Kym I E Snell,<sup>3</sup> Joie Ensor,<sup>3</sup> Lotty Hooft,<sup>1,2</sup>  
Johannes B Reitsma,<sup>1,2</sup> Richard D Riley,<sup>3</sup> Karel G M Moons<sup>1,2</sup>

# Step 1. Well-formulated review question: PICOTS

## Guidance for framing review question: CHARMS checklist

Critical Appraisal and Data Extraction for Systematic  
Reviews of Prediction Modelling Studies: The CHARMS  
Checklist

Plos Med 2014

Karel G. M. Moons<sup>1†\*</sup>, Joris A. H. de Groot<sup>1†</sup>, Walter Bouwmeester<sup>1</sup>, Yvonne Vergouwe<sup>1</sup>, Susan Mallett<sup>2</sup>,  
Douglas G. Altman<sup>3</sup>, Johannes B. Reitsma<sup>1</sup>, Gary S. Collins<sup>3</sup>

RESEARCH METHODS AND REPORTING

A guide to systematic review and meta-analysis of prognostic  
factor studies **BMJ 2019**

Richard D Riley,<sup>1\*</sup> Karel G M Moons,<sup>2,4\*</sup> Kym I E Snell,<sup>1</sup> Joie Ensor,<sup>1</sup> Lotty Hooft,<sup>2,4</sup>  
Douglas G Altman,<sup>3</sup> Jill Hayden,<sup>5</sup> Gary S Collins,<sup>3</sup> Thomas P A Debray<sup>2,4</sup>

RESEARCH METHODS AND REPORTING

A guide to systematic review and meta-analysis of prediction  
model performance **BMJ 2017**

Thomas P A Debray,<sup>1,2</sup> Johanna A A G Damen,<sup>1,2</sup> Kym I E Snell,<sup>3</sup> Joie Ensor,<sup>3</sup> Lotty Hooft,<sup>1,2</sup>  
Johannes B Reitsma,<sup>1,2</sup> Richard D Riley,<sup>3</sup> Karel G M Moons<sup>1,2</sup>

# PICOTS for SRs of Prognostic factor(s)

Item	Comment
<b>1. <u>P</u>opulation</b>	Target population in which the prognostic factor(s) under review will be used.
<b>2. <u>I</u>ndex prognostic factor(s)</b>	Index prognostic factor(s) whose prognostic ability is under review.
<b>3. <u>C</u>omparator prognostic factor(s)</b>	One or more comparator prognostic factors can be reviewed, if applicable. E.g. comparing prognostic ability of certain index factor to other (i.e. comparator) prognostic factors. Or review of the adjusted prognostic ability of a certain index factor, adjusted for other (i.e. comparator) prognostic factors. If aim is summarise unadjusted prognostic effect of index factor, then no comparator factor is addressed.
<b>4. <u>O</u>utcome(s)</b>	Outcome(s) of interest for the factor(s) under review.
<b>5. <u>T</u>iming (two elements)</b>	<ul style="list-style-type: none"> <li>(i) at what time-point(s) prognostic factors (index and comparators) are to be used (time point of prognostication);</li> <li>(ii) over what time period outcome(s) are predicted.</li> </ul>
<b>6. <u>S</u>etting</b>	Define the intended setting (role) of the prognostic factor(s) under review.

# PICOTS of SRs of Prognostic (prediction) model(s)

Item	Comment
<b>1. <u>P</u>opulation</b>	Target population in which prediction model(s) under review will be used.
<b>2. <u>I</u>ndex prediction model(s)</b>	Index prediction model(s) under review
<b>3. <u>C</u>omparator prediction model(s)</b>	One can compare the predictive ability of the index model to one or more other prediction models, if applicable.
<b>4. <u>O</u>utcome(s)</b>	Outcome(s) of interest for the model(s) under review.
<b>5. <u>T</u>iming (two elements)</b>	<ol style="list-style-type: none"> <li>1. At what time-point(s) prediction models (index and comparators) are to be used (time point of prognostication);</li> <li>2. Over what time period (notably for prognostic prediction models) outcome(s) are predicted.</li> </ol>
<b>6. <u>S</u>etting</b>	Intended setting (role) of the prediction model(s) under review.



# Practical

## Exercise:

- **Define a review question + PICOTS**

## A Comprehensive Appraisal of Laboratory Biochemistry Tests as Major Predictors of COVID-19 Severity

*Elena Aloisio, MD; Mariia Chibireva, MD; Ludovica Serafini, MD; Sara Pasqualetti, MSc; Felicia S. Falvella, MSc; Alberto Dolci, MD; Mauro Panteghini, MD*

## Possible answer (but more answers are possible)

<b>Population</b>	Patients with COVID-19, proven with PCR
<b>Index factor(s)</b>	Any laboratory test (blood, urine, etc.)
<b>Comparator</b>	Not applicable (or , e.g., added to self tests or added to patient's symptoms and signs)
<b>Outcomes</b>	Overall Mortality (or, e.g., ICU admission, or combination)
<b>Timing</b>	1. Moment of prognostication: at COVID-19 diagnosis with PCR test 2. Time horizon: within 2 weeks
<b>Setting</b>	Secondary care

# BMC Medical Informatics and Decision Making



Research article

Open Access

## Systematic review of prognostic models in traumatic brain injury

Pablo Perel\*, Phil Edwards, Reinhard Wentz and Ian Roberts

Address: Nutrition and Public Health Intervention Research Unit, Epidemiology and Population Health Department, London School of Hygiene

Different clinical questions possible → depending on aim of the SR?

**Group exercise:**

- **Define a review question + PICOTS**

## Two possible answers (out of many possibilities)

<b>Population</b>	Patients with TBI surviving the first 24 hours	Patients with TBI right after accident
<b>Index model(s)</b>	All models	IMPACT model (i.e., focus on 1 specific model)
<b>Comparator</b>	Not applicable	CRASH model (i.e., one another specific model)
<b>Outcomes</b>	Daily functioning	Mortality
<b>Timing</b>	1. Prediction T0: 24 hours after accident/injury 2. Three months prediction of outcome	1. Prediction T0: right after accident/injury 2. Within 30 days
<b>Setting</b>	Patients in hospital surviving a TBI after 24 hours	Prediction in ambulance (or at battle field)

# Types of SR prognostic/prediction model questions

- Review all models for specific outcome in specific target population
  - Models predicting fatal/non-fatal coronary heart disease in middle-aged general population; models predicting stroke in 60+ of general population;
  - Models predicting survival after cardiac surgery ; predicting Length of stay after cardiac surgery ; predicting quality of life after surgery
- Review all existing models in a particular clinical field
  - e.g. all models for any cardiovascular disease outcome in general population; all developed models in obstetrics.



# Types of SR prognostic/prediction model questions

- How good is predictive performance of a specific model for a specific target population (validation studies only)
  - Predictive performance of Framingham risk model / GAIL model
- Review on added predictive value of a specific predictor/biomarker/test to a specific model
  - Adding CRP to Framingham risk score; D-dimer to Wells Rule
  - Adding imaging results to ‘basic risk scores’ (cancer models)



# Conducting a systematic review of prognosis studies

1. Formulate review question (PICOTS)
2. Searching for studies
3. Screening and Selection of articles
4. Extraction of data
5. Risk of Bias assessments
6. Synthesis of data (meta-analysis)
7. Interpretation and conclusions

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prognostic factor studies

Richard D Riley,<sup>1\*</sup> Karel G M Moons,<sup>2,4\*</sup> Kym I E Snell,<sup>1</sup> Joie Ensor,<sup>1</sup> Lotty Hooft,<sup>2,4</sup>  
Douglas G Altman,<sup>3</sup> Jill Hayden,<sup>5</sup> Gary S Collins,<sup>3</sup> Thomas P A Debray<sup>2,4</sup>

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prediction model performance

Thomas P A Debray,<sup>1,2</sup> Johanna A A G Damen,<sup>1,2</sup> Kym I E Snell,<sup>3</sup> Joie Ensor,<sup>3</sup> Lotty Hooft,<sup>1,2</sup>  
Johannes B Reitsma,<sup>1,2</sup> Richard D Riley,<sup>3</sup> Karel G M Moons<sup>1,2</sup>

# Search strategies

- No optimal, reliable methods for searching the literature for prognostic information
  - As for RCTs and Diagnostic Test Accuracy Studies
- A few published
  - Altman DG (2001): single prognostic factors
  - Wong SS (2003): very generic
  - Ingui BJ (2001): prediction models
  - Geersing (2012): validation Ingui (2001) and updated (new) search strategy
  - Kavanagh (2021): Optimizing a literature surveillance strategy to retrieve sound overall prognosis and risk assessment model papers.
  - Stallings (2022): Development and evaluation of a search filter to identify prognostic factor studies in Ovid MEDLINE.





# Search Filters for Finding Prognostic and Diagnostic Prediction Studies in Medline to Enhance Systematic Reviews

Geert-Jan Geersing<sup>1\*</sup>, Walter Bouwmeester<sup>1,2</sup>, Peter Zuithoff<sup>1</sup>, Rene Spijker<sup>2,4</sup>, Mariska Leeflang<sup>3,4</sup>, Karel Moons<sup>1</sup>

**1** Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands, **2** Medical Library Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, **3** Department of Clinical Epidemiology and Bio-Informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands, **4** Dutch Cochrane Center, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

**Table 1.** Search strategies for finding prediction research in Medline.

Filter	Search terms included in the filter*	Sensitivity# (95% CI)	Specificity# (95% CI)
Ingui filter	(Validat\$ OR Predict\$.ti. OR Rule\$) OR (Predict\$ AND (Outcome\$ OR Risk\$ OR Model\$)) OR ((History OR Variable\$ OR Criteria OR Scor\$ OR Characteristic\$ OR Finding\$ OR Factor\$) AND (Predict\$ OR Model\$ OR Decision\$ OR Identif\$ OR Prognos\$)) OR (Decision\$ AND (Model\$ OR Clinical\$ OR Logistic Models/)) OR (Prognostic AND (History OR Variable\$ OR Criteria OR Scor\$ OR Characteristic\$ OR Finding\$ OR Factor\$ OR Model\$))	0.98 (0.92–1.0)	0.86 (0.85–0.87)
Haynes broad filter	(Predict*[tiab] OR Predictive value of tests[mh] OR Scor*[tiab] OR Observ*[tiab] OR Observer variation[mh])	0.96	0.79

\*Using the Pubmed interface for MEDLINE.

#Sensitivity and specificity as reported by Ingui and Haynes in their original publication; CI = confidence interval, for the Haynes broad filter no confidence intervals were given in the original publication.

doi:10.1371/journal.pone.0032844.t001

# Geersing et al 2012

## Conclusions

Updated search strategy for prognosis research good in retrieving "Prediction model studies" (Se 0.78 to 0.89)

Less value in retrieving "Predictor Finding/prognostic factor" and "Prediction Model Impact Studies"

**Table 4.** Updated search string for finding prediction research.

"Stratification" OR "ROC Curve"[Mesh] OR "Discrimination" OR "Discriminate" OR "c-statistic" OR "c statistic" OR "Area under the curve" OR "AUC" OR "Calibration" OR "Indices" OR "Algorithm" OR "Multivariable"
--

doi:10.1371/journal.pone.0032844.t004

Strategy for "Predictor Finding / prognostic factor" studies still sub-optimal but good starting point!

# Conducting a systematic review of prognosis studies

1. Formulate review question (PICOTS)
2. Searching for studies
3. Screening and Selection of articles
4. Extraction of data
5. Risk of Bias assessments
6. Synthesis of data (meta-analysis)
7. Interpretation and conclusions

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prognostic factor studies

Richard D Riley,<sup>1\*</sup> Karel G M Moons,<sup>2,4\*</sup> Kym I E Snell,<sup>1</sup> Joie Ensor,<sup>1</sup> Lotty Hooft,<sup>2,4</sup>  
Douglas G Altman,<sup>3</sup> Jill Hayden,<sup>5</sup> Gary S Collins,<sup>3</sup> Thomas P A Debray<sup>2,4</sup>

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prediction model performance

Thomas P A Debray,<sup>1,2</sup> Johanna A A G Damen,<sup>1,2</sup> Kym I E Snell,<sup>3</sup> Joie Ensor,<sup>3</sup> Lotty Hooft,<sup>1,2</sup>  
Johannes B Reitsma,<sup>1,2</sup> Richard D Riley,<sup>3</sup> Karel G M Moons<sup>1,2</sup>

# Study selection

- Not different from other types of reviews



# Intermezzo Challenge

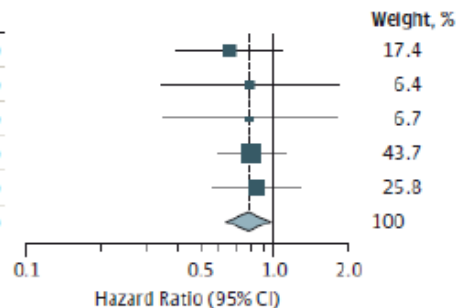
## Meta-analysis/Pooling of prognostic factor studies

**Exercise 10 minutes:**

1. Assume this forest plot is of RCTs on intervention X to prevent outcome Y in patients with disease Z.

- Is this pooling ok?
- Why or why not?

Study	Hazard Ratio (95% CI)
Kuyken et al, <sup>13</sup> 2008	0.66 (0.40-1.08)
Segal et al, <sup>18</sup> 2010	0.80 (0.35-1.82)
Huijbers et al, <sup>19</sup> 2015	0.80 (0.36-1.78)
Kuyken et al, <sup>21</sup> 2015	0.81 (0.59-1.11)
Williams et al, <sup>23</sup> 2014	0.85 (0.56-1.28)
Overall ( $I^2 = 0.0\%$ , $P = .96$ )	0.79 (0.64-0.97)



2. Assume this forest plot is of studies on prognostic factor X, to predict outcome Y in patients with disease Z.

- Is this pooling ok?
- Why or why not?

## Meta-analysis/Pooling in prognostic factor studies

- If RCTs
  - Pooling is ok – provided correctly randomised
  - Then the 3 HRs are unbiased (provided no other risks of biases) so can easily pool them
  - Clear effect of intervention X to prevent outcome Y
  - In frequentist world, at alpha 0.05 – even statistically significant result.
- If prognostic factor studies?
  - Non randomised -> even if a study was based on a RCT – the prognostic factor analysis is per arm and thus non randomised
  - Can not assume that the 3 HRs are unbiased
  - Only pool them if studies have adjusted for the same co-variables – or largely for the same co-variables – e.g. the same big 6 or 7 (the eighth co variate probably did not change the HR further)
  - So pooling of prognostic factor studies only if same adjustment -- otherwise do stratified pooling (e.g. over studies with similar adjustment)

# Conducting a systematic review of prognosis studies

1. Formulate review question (PICOTS)
2. Searching for studies
3. Screening and Selection of articles
4. Extraction of data
5. Risk of Bias assessments
6. Synthesis of data (meta-analysis)
7. Interpretation and conclusions

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prognostic factor studies

Richard D Riley,<sup>1\*</sup> Karel G M Moons,<sup>2,4\*</sup> Kym I E Snell,<sup>1</sup> Joie Ensor,<sup>1</sup> Lotty Hooft,<sup>2,4</sup>  
Douglas G Altman,<sup>3</sup> Jill Hayden,<sup>5</sup> Gary S Collins,<sup>3</sup> Thomas P A Debray<sup>2,4</sup>

## RESEARCH METHODS AND REPORTING

### A guide to systematic review and meta-analysis of prediction model performance

Thomas P A Debray,<sup>1,2</sup> Johanna A A G Damen,<sup>1,2</sup> Kym I E Snell,<sup>3</sup> Joie Ensor,<sup>3</sup> Lotty Hooft,<sup>1,2</sup>  
Johannes B Reitsma,<sup>1,2</sup> Richard D Riley,<sup>3</sup> Karel G M Moons<sup>1,2</sup>

- Extraction of characteristics/data of included studies + Critical appraisal
  - **CHARMS** – Table 2
  - 11 domains + signaling items

## Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies: The CHARMS Checklist

Karel G. M. Moons<sup>1†\*</sup>, Joris A. H. de Groot<sup>1†</sup>, Walter Bouwmeester<sup>1</sup>, Yvonne Vergouwe<sup>1</sup>, Susan Mallett<sup>2</sup>, Douglas G. Altman<sup>3</sup>, Johannes B. Reitsma<sup>1</sup>, Gary S. Collins<sup>3</sup>

### RESEARCH METHODS AND REPORTING

- Has been adapted for prognostic factors as well:

A guide to systematic review and meta-analysis of prognostic factor studies

Richard D Riley,<sup>1\*</sup> Karel G M Moons,<sup>2,4\*</sup> Kym I E Snell,<sup>1</sup> Joie Ensor,<sup>1</sup> Lotty Hooft,<sup>2,4</sup> Douglas G Altman,<sup>3</sup> Jill Hayden,<sup>5</sup> Gary S Collins,<sup>3</sup> Thomas P A Debray<sup>2,4</sup>



# Data Extraction

## Key issues CHARMS checklist

Domain	Key items	Reported on page #
SOURCE OF DATA	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	
PARTICIPANTS	Participant eligibility and recruitment method (e.g., consecutive participants, location, number of centers, setting, inclusion and exclusion criteria)	
	Participant description	
	Details of treatments received, if relevant	
	Study dates	
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	
	Was the same outcome definition (and method for measurement) used in all patients?	
	Type of outcome (e.g., single or combined endpoints)	
	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	
	Time of outcome occurrence or summary of duration of follow-up	
CANDIDATE PREDICTORS (OR INDEX TESTS)	Number and type of predictors (e.g., demographics, patient history, physical examination, additional testing, disease characteristics)	
	Definition and method for measurement of candidate predictors	
	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation)	
	Were predictors assessed blinded for outcome, and for each other (if relevant)?	
	Handling of predictors in the modelling (e.g., continuous, linear, non-linear transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	
	Number of outcomes/events in relation to the number of candidate predictors (Events Per Variable)	
MISSING DATA	Number of participants with any missing value (include predictors and outcomes)	
	Number of participants with missing data for each predictor	
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	

# Data Extraction

## Key issues CHARMS checklist

Domain	Key items	Reported on page #
SOURCE OF DATA	Source of data (e.g., cohort, case-control, randomized trial participants, or registry data)	
PARTICIPANTS	Participant eligibility and recruitment method (e.g., consecutive participants, location, number of centers, setting, inclusion and exclusion criteria)	
	Participant description	
	Details of treatments received, if relevant	
	Study dates	
OUTCOME(S) TO BE PREDICTED	Definition and method for measurement of outcome	
	Was the same outcome definition (and method for measurement) used in all patients?	
	Type of outcome (e.g., single or combined endpoints)	
	Was the outcome assessed without knowledge of the candidate predictors (i.e., blinded)?	
	Were candidate predictors part of the outcome (e.g., in panel or consensus diagnosis)?	
	Time of outcome occurrence or summary of duration of follow-up	
CANDIDATE PREDICTORS (OR INDEX TESTS)	Number and type of predictors (e.g., demographics, patient history, physical examination, additional testing, disease characteristics)	
	Definition and method for measurement of candidate predictors	
	Timing of predictor measurement (e.g., at patient presentation, at diagnosis, at treatment initiation)	
	Were predictors assessed blinded for outcome, and for each other (if relevant)?	
	Handling of predictors in the modelling (e.g., continuous, linear, non-linear transformations or categorised)	
SAMPLE SIZE	Number of participants and number of outcomes/events	
	Number of outcomes/events in relation to the number of candidate predictors (Events Per Variable)	
MISSING DATA	Number of participants with any missing value (include predictors and outcomes)	
	Number of participants with missing data for each predictor	
	Handling of missing data (e.g., complete-case analysis, imputation, or other methods)	

MODEL DEVELOPMENT	Modelling method (e.g., logistic, survival, neural network, or machine learning techniques)	
	Modelling assumptions satisfied	
	Method for selection of predictors <b>for inclusion</b> in multivariable modelling (e.g., all candidate predictors, pre-selection based on unadjusted association with the outcome)	
	Method for selection of predictors <b>during multivariable modelling</b> (e.g., full model approach, backward or forward selection) and criteria used (e.g., p-value, Akaike Information Criterion)	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform shrinkage, penalized estimation)	
MODEL PERFORMANCE	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination (C-statistic, D-statistic, log-rank) measures with confidence intervals	
	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification improvement) and whether a-priori cut points were used	
MODEL EVALUATION	Method used for testing model performance: development dataset only (random split of data, resampling methods e.g. bootstrap or cross-validation, none) or separate external validation (e.g. temporal, geographical, different setting, different investigators)	
	In case of poor validation, whether model was adjusted or updated (e.g., intercept recalibrated, predictor effects adjusted, or new predictors added)	
RESULTS	Final and other multivariable models (e.g., basic, extended, simplified) presented, including predictor weights or regression coefficients, intercept, baseline survival, model performance measures (with standard errors or confidence intervals)	
	Any alternative presentation of the final prediction models, e.g., sum score, nomogram, score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and validation datasets	
INTERPRETATION AND DISCUSSION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	

MODEL DEVELOPMENT	Modelling method (e.g., logistic, survival, neural network, or machine learning techniques)	
	Modelling assumptions satisfied	
	Method for selection of predictors <b>for inclusion</b> in multivariable modelling (e.g., all candidate predictors, pre-selection based on unadjusted association with the outcome)	
	Method for selection of predictors <b>during multivariable modelling</b> (e.g., full model approach, backward or forward selection) and criteria used (e.g., p-value, Akaike Information Criterion)	
	Shrinkage of predictor weights or regression coefficients (e.g., no shrinkage, uniform shrinkage, penalized estimation)	
MODEL PERFORMANCE	Calibration (calibration plot, calibration slope, Hosmer-Lemeshow test) and Discrimination (C-statistic, D-statistic, log-rank) measures with confidence intervals	
	Classification measures (e.g., sensitivity, specificity, predictive values, net reclassification improvement) and whether a-priori cut points were used	
MODEL EVALUATION	Method used for testing model performance: development dataset only (random split of data, resampling methods e.g. bootstrap or cross-validation, none) or separate external validation (e.g. temporal, geographical, different setting, different investigators)	
	In case of poor validation, whether model was adjusted or updated (e.g., intercept recalibrated, predictor effects adjusted, or new predictors added)	
RESULTS	Final and other multivariable models (e.g., basic, extended, simplified) presented, including predictor weights or regression coefficients, intercept, baseline survival, model performance measures (with standard errors or confidence intervals)	
	Any alternative presentation of the final prediction models, e.g., sum score, nomogram, score chart, predictions for specific risk subgroups with performance	
	Comparison of the distribution of predictors (including missing data) for development and validation datasets	
INTERPRETATION AND DISCUSSION	Interpretation of presented models (confirmatory, i.e., model useful for practice versus exploratory, i.e., more research needed)	
	Comparison with other studies, discussion of generalizability, strengths and limitations.	

## Take home messages

- 4 main types of prognosis studies
- 3 main types of prognostic model studies
- Systematic reviews of prognostic factor and model studies largely same as for intervention SRs
- Different and indeed more challenges in SRs of prognosis studies
- Tools available for all familiar steps of SR → prognosis studies



# Conducting systematic reviews of prediction model studies



Reporting of primary study

Transparent reporting of prediction models for prognosis and diagnosis (TRIPOD+AI) – *Collins et al. 2024 BMJ; Moons et al. 2015 Ann Intern Med*

Defining review question and developing criteria for including studies

Guidance for defining review question, design of the review and checklist for critical appraisal and data extraction (CHARMS) – *Moons et al 2014 PLOS Med*

Searching for studies

Search filters for prediction studies – *Geersing et al. 2012 PLOS One; Ingui et al. 2002 J Am Med Inform Assoc; Wong et al. 2003 AMIA Annual Symp Proc*

Selecting studies and collecting data

Guidance for defining review question, design of the review and checklist for critical appraisal and data extraction (CHARMS) – *Moons et al 2014 PLOS Med*

Assessing risk of bias and applicability in included studies

Assessment of risk of bias and applicability (PROBAST+AI) – *Moons et al. 2025 BMJ; Moons et al. E&E 2019 Ann Intern Med*

Analysing data and undertaking meta-analyses

Meta-Analysis of clinical prediction models  
*Ahmed et al. BMC Res Meth 2014; Debray et al. Stat Med 2012; Debray et al. Stat Med 2014; Debray et al BMJ 2016*

Interpreting results and drawing conclusions

Guidance for interpretation of results  
*Iorio et al. BMJ 2015; Huguier 2013 Syst Rev; Foroutan 2020&2022&2024 JCE*

Reporting of systematic reviews

Transparent reporting of systematic reviews and meta-analysis (PRISMA 2020 & TRIPOD-SRMA)  
*Page et al. BMJ 2021; Snell et al. BMJ 2023*

Assessing risk of bias of systematic reviews

Risk of bias in systematic reviews (ROBIS)  
*Whiting et al. J Clin Epid 2015*