Systematic reviews of prognosis studies

Carl Moons, Lotty Hooft, Anneke Damen, Katrina Williams, Nicole Skoetz, Marialena Trivella, Angela Aldin, Thomas Debray, Richard Riley, Jill Hayden

Cochrane Methods Prognosis

We have no actual or potential conflict of interest in relation to this presentation

Contents

1. Introduction to systematic reviews of prognosis
2. Formulate a research question
   - Searching for studies
3. Screening and selection of articles

BREAK

4. Data extraction
5. Risk of bias (short introduction)

Systematic reviews (SRs)

- Applicable to all fields of medical research
  - Therapeutic studies (RCTs): Cochrane Intervention Reviews
  - Diagnostic accuracy studies: Cochrane Diagnostic Test Accuracy Reviews
    - Both including meta-analytical approaches

- Next: prognosis studies

Why?

- Increasing interest in and demand for the evaluation of prognostic factors, biomarkers & models
- Growing number of primary studies
- Reviews more challenging: more variation in questions, designs, effect measures, analyses
- Several recent methodological developments & remaining challenges

Key course objectives

i. To know the main types of prognostic studies
ii. To understand the different aims of systematic reviews of (prognostic) prediction modeling studies
iii. To describe the similarities and differences between intervention and prediction modeling reviews
iv. To learn about data extraction and screening of articles

1. What is prognosis, as compared to treating and diagnosis?

2. Why do we prognosticate?

3. Types of prognosis studies?

Answer

(BMJ series 2009 (Altman, Moons, Royston, Vergouwe) + Progress series BMJ/Plos Med 2013

Forecast of the **course** and **outcome** for an **individual** in a **certain health state** (given a **specific treatment** management)

  - Not necessarily sick people

- More technical: probable course/prediction of specific future outcomes in subjects with certain health condition

- Disease does not have a prognosis → an individual does

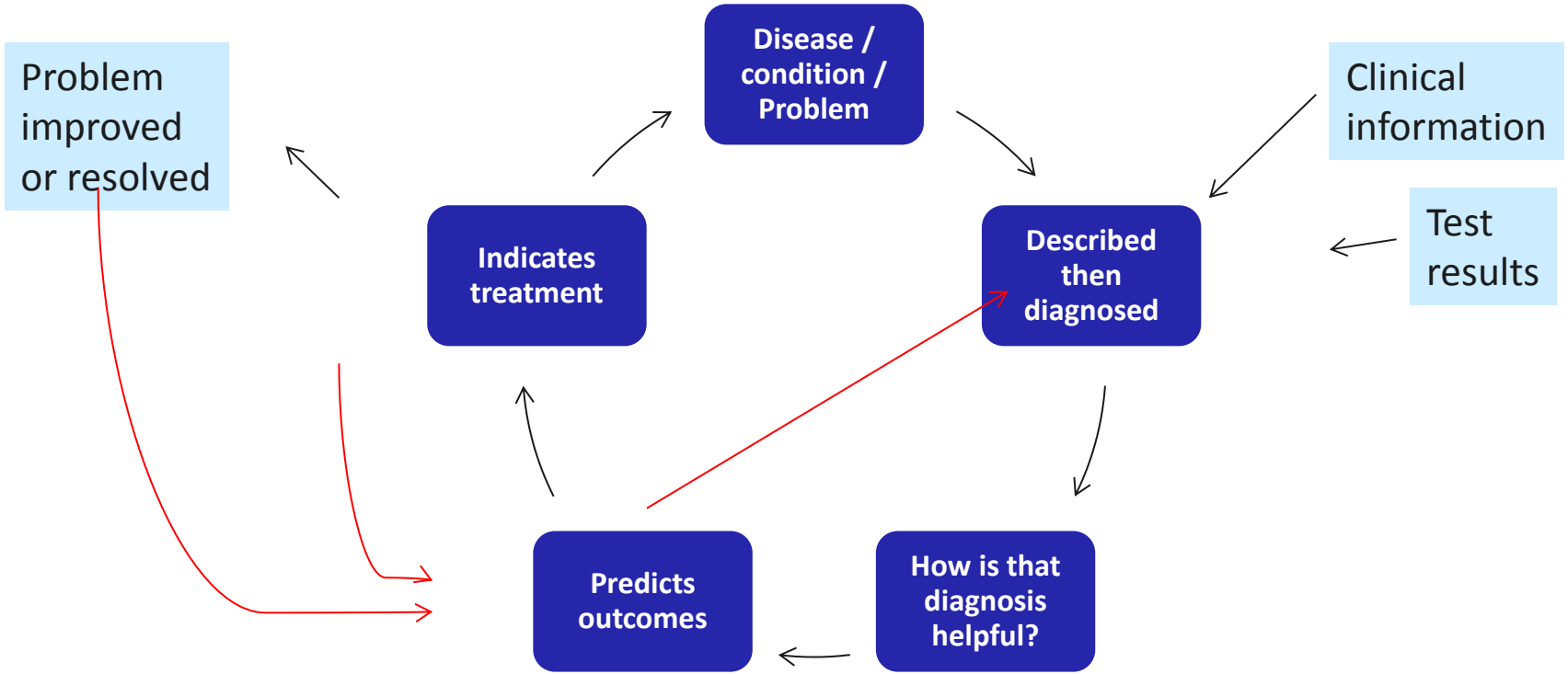
Answer

All is interconnected

Cause	*What factors/conditions result in disease?*
Diagnosis	*How accurate are diagnostic tests to find disease?*
Prognosis	*What are the consequences of having disease?*
Treatment	*Does Rx make a difference to the course of disease?*
Prevention	*Does early Dx and Rx improve outcome?*

Answer

Diagnostic-Prognostic-Therapeutic cycle



Answer

Health outcome impact and interconnectedness

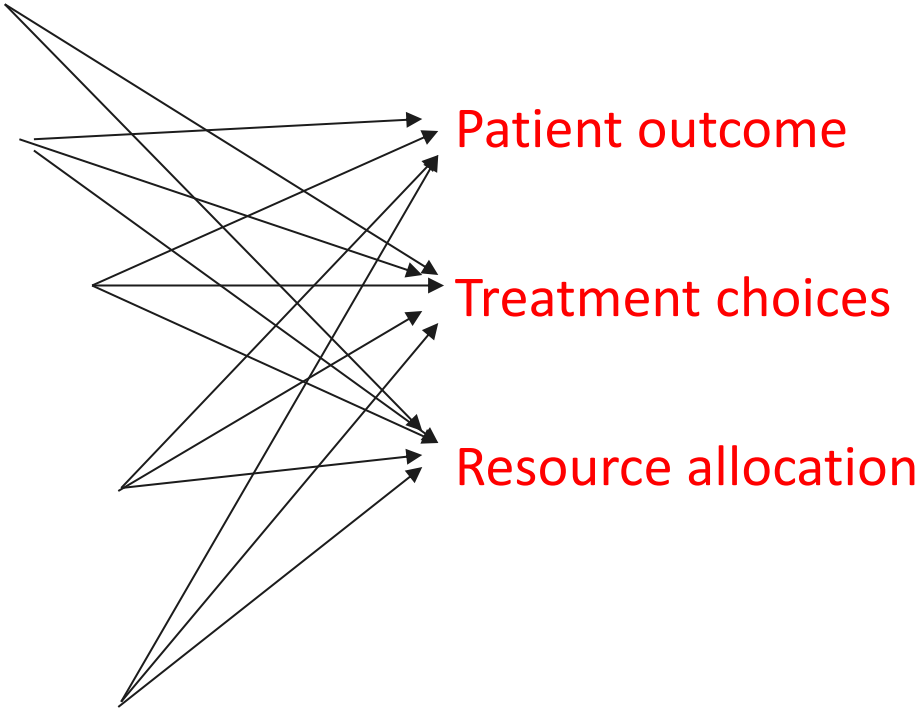
How common is a problem

What is the cause

How accurately can we diagnose a problem

How effective are treatments and what are their risks

What is the prognosis



Answer

- Why prognosticate:
  - To provide information to patients
  - Identify groups for treatment or other management – including abstine
  - 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 health care of institutes
  - Service developers make decisions about what services are needed
  - policy makers what to support/advocate

Answer

Types of prognosis studies?

PROGRESS series 2013: BMJ and Plos Med

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?'

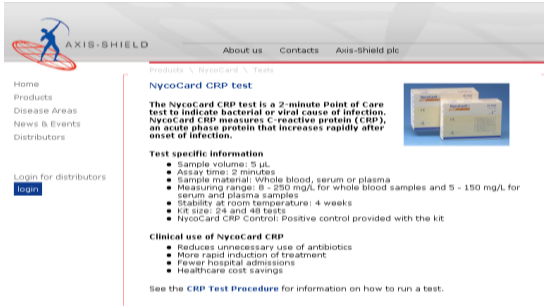
4. Treatment selection factors: 'Which factors lead to/predict different treatment effect in individuals to be treated?'

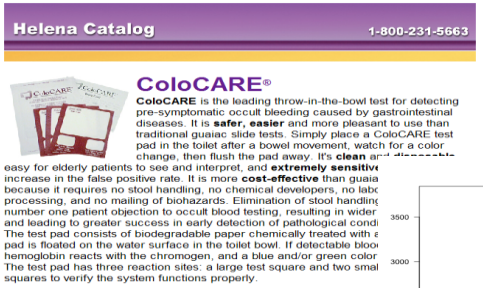
Focus on 2 +3

Why SRs Prognostic studies?

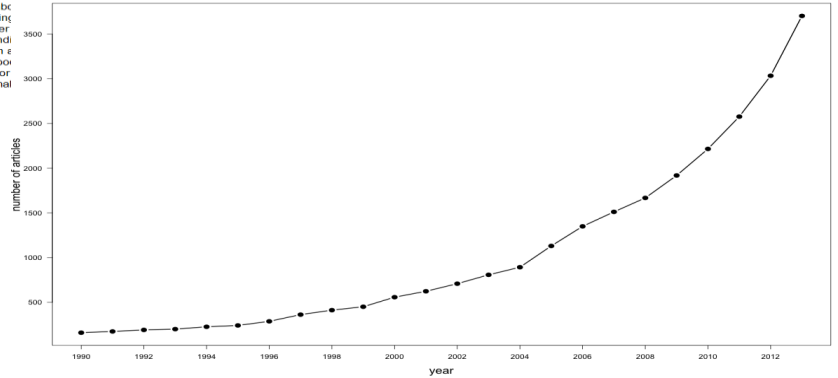
1. Number of studied prognostic factors increases per day due to precision/personalized medicine focus

- Biomarkers (all types)





- Also prognostic models



Why SRs Prognostic studies?

2. Most studies conflicting results
   - much more than in therapeutic trials and in diagnostic test accuracy studies
   - **Non-randomised (often not predesigned studies)**
   - **Often retrospective using existing data sets**

3. Relatively small studies (compared to therapeutic studies)
   - Kyzas Eur J Canc 2007; > 1500 studies cancer prognostic markers in 2005 → largest just over 1000 pts.

Hence …

… prognosis studies are hot

… SR's and notably MA of prognosis studies as well

- highly desired and well received by journals/policy makers →

- to systematically summarise the existing evidence in the field

Types of prognosis studies

PROGRESS series 2013: BMJ and Plos Med

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?'

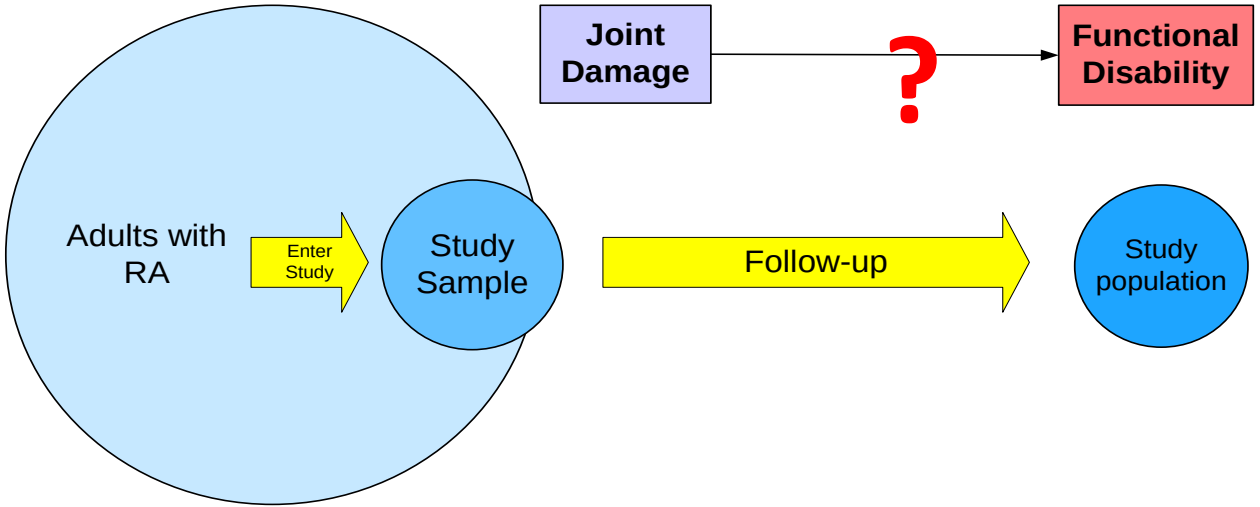
4. Treatment selection factors: 'Which factors lead to/predict different treatment effect in individuals to be treated?'

Prognostic Factor Studies

Aim:

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

Prognostic Factor Study Example



Adapted from: *Fletcher & Fletcher, Clinical Epidemiology – The Essentials. Chapter 6. Williams & Wilkins, Baltimore. 1996*

Types of prognosis studies

PROGRESS series 2013: BMJ and Plos Med

1. Average/overall prognosis: 'What is most likely course (outcome) of individuals in certain health condition (often certain disease)?'

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?'

4. Treatment selection factors: 'Which factors lead to/predict different treatment effect/response in individuals to be treated?'

Prognostic Prediction Model Studies

1. What is a prognostic prediction model study, and what is difference with multivariable analysis of prognostic factors?

2. There are three phases of prediction modelling – which three?

3. What is the biggest difference between phase 1+2 versus 3?

Answers

Risk prediction = foreseeing / foretelling

… (probability) of something that is yet unknown

Combination of 2 or more predictors 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 (diagnostic or prognostic) probability. Focus is on accuracy of entire model (discrimination + calibration). Factors of the model not at interest.

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

Answers

3 Phases of Prediction Modelling studies

BMJ series 2009/Bouwmeester 2012/PROGRESS series 2013 (BMJ/Plos Med)

1. Model development studies – to develop prediction model from data: identify important predictors; estimate predictor weights; construct model for individualised predictions; quantify predictive performance; internal validation

2. Model validation studies – test (validate) predictive performance of previously developed model in participant data other than development set

3. Model impact studies – quantify effect/impact of actually using model on participant/physician management and health outcomes – relative to not using the model

What is the difference between 3 versus 1 and 2?

Answers

3 Phases of Prediction Modelling studies

BMJ series 2009/Bouwmeester 2012/PROGRESS series 2013 (BMJ/Plos Med)

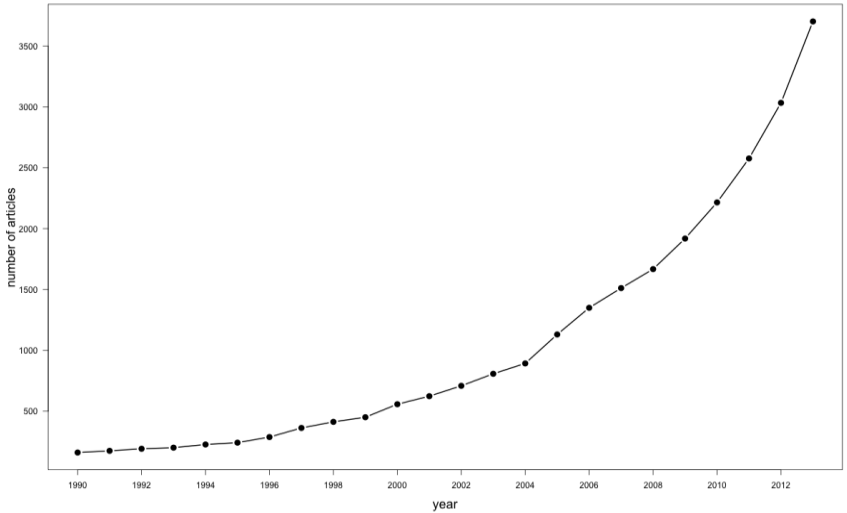
- Big difference = 3 are comparative studies → ideally randomised
- 1 and 2 are by definition single cohort studies- no inherent comparison
- 3 are thus ideally RCTs – for SRs of prediction model impact studies use the Cochrane tools available for RCTs of intervention studies
- This course provides tool for prediction model development and validation studies (type 1 and 2)

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

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

Interesting and booming field – stay in it!

Prediction models are hot



(("prognostic model") OR ("prediction model") OR ("risk score") OR ("clinical prediction rule") OR ("decision rule") OR ("prognostic index") OR ("prognostic indices") OR ("prediction index") OR ("risk algorithm") OR ("risk stratification") OR ("multivariable prediction"))

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. Meta-analysis of prognostic factor studies
7. Meta-analysis of prognostic model studies
8. Interpretation and conclusions

Formulating the review question

Step 1. Well-formulated review question

Actually: define the PICO → stands for?

Guidance frame 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[1¶*], Joris A. H. de Groot[1¶], Walter Bouwmeester[1], Yvonne Vergouwe[1], Susan Mallett[2], Douglas G. Altman[3], Johannes B. Reitsma[1], Gary S. Collins[3]

RESEARCH METHODS AND REPORTING

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

BMJ 2017

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

PICOTS SR Prognostic factor(s)

Item	Comments
1. Population	Define target population in which prognostic factor(s) under review will be used.
2. Index factor(s)	Define the prognostic factor(s) under review.
3. Comparator(s)	If applicable, one can review more than one factor for the target population and outcome under review.
4. Outcome(s)	Define the outcome(s) of interest for the factor(s) under review.
5. Timing	Define at what time-points the prognostic factor(s) are to be used and over what time period the outcome(s) are predicted
6. Setting	Define the intended role or setting of the prognostic factor(s) under review.

PICOTS SR Prognostic (prediction) model(s)

Item	Comments
1. Population	Define target population in which prediction model(s) under review will be used.
2. Index model(s)	Define the prediction model(s) under review.
3. Comparator(s)	If applicable, one can review more than one model for the target population and outcome under review.
4. Outcome(s)	Define the outcome(s) of interest for the model(s) under review.
5. Timing	Define when prediction model(s) under review is intended to be used and over what time period (notably for prognostic prediction models) the outcome(s) is predicted.
6. Setting	Define the intended role or setting of the prediction model(s) under review.

Practical

Exercise:

- **Define a review question + PICOTS**

Research

BMJ

Value of sentinel node status as a prognostic factor in melanoma: prospective observational study

Stephen Kettlewell, Colin Moyes, Caroline Bray, David Soutar, Alan MacKay, Dominique Byrne, Taimur Shoaib, Barun Majumder, Rona MacKie

Abstract

Objective To establish the prognostic value of knowledge of sentinel node status in melanoma.

Design Single centre prospective observational study, with sentinel nodes identified by lymphoscintigraphy, γ probe, and intraoperative blue dye and examined by both conventional histopathology and immunopathology

multicentre randomised trial (MSLT1) is in progress with the aim of determining if patients with melanoma who have a positive SNB and proceed immediately to full node dissection have a superior disease-free survival or overall survival compared with patients who have node dissection only when nodes draining the site of the primary melanoma are clinically palpable. Definitive results are awaited.[12]

Suggested answer

Population	• Patients with melanoma
Index factor	• Sentinel node status
Comparator	• Not applicable
Outcomes	• Recurrence • Mortality
Timing	• Prediction at preoperative visit • 3 months prediction of outcomes
Setting	• Patients scheduled for surgery to remove the melanoma

Save a Copy Search Select 165% Help

BMC Medical Informatics and Decision Making

BioMed Central

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 → different aims of SR of prediction models?

Group exercise:

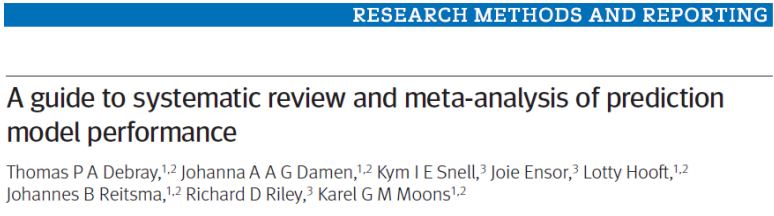
- **Define a review question + PICOTS**

Suggested answer

Population	• Patients with TBI (e.g. surviving the first 24 hours)
Index model(s)	• All developed+validated models for patients with TBI (surviving the first 24 hours)
Comparator	• All existing developed+validated models
Outcomes	• Mortality • Or: Daily functioning
Timing	• Prediction T0 24 hours after accident/injury • 3 months predciton of outcome (or: 12 months)
Setting	• Patients in hospital surviving a TBI after 24 hours • Or: battlefield TBI • Or: prediction in ambulance or at the site of the accident

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. Meta-analysis of prognostic factor studies
7. Meta-analysis of prognostic model studies
8. Interpretation and conclusions



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

OPEN ACCESS Freely available online

PLoS one

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

Geert-Jan Geersing[1*☯], Walter Bouwmeester[1☯], Peter Zuithoff[1], Rene Spijker[2,4], Mariska Leeflang[3,4], Karel Moons[1]

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!

Study selection

- Selecting studies involves judgement, and is highly influential on the outcomes of the review

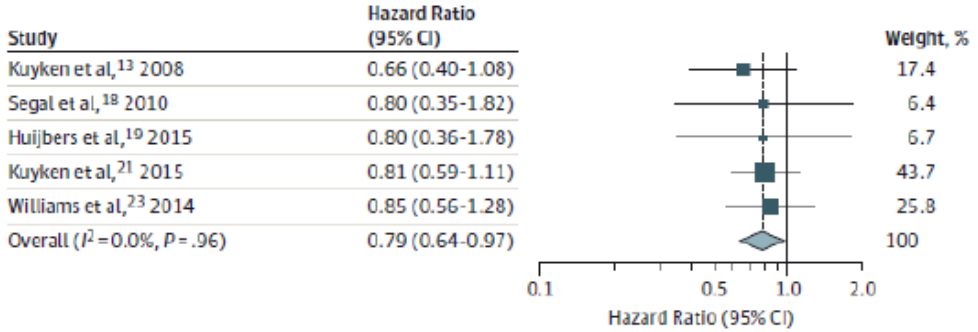
- Two (or more…) reviewers, independently
  - minimizing bias
  - pilot selection on a few papers first: substantial variation
  - Determine how will disagreements be managed

- Examine titles and abstracts
- Flow chart of included/excluded studies
- Retrieve and examine full text reports

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?



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

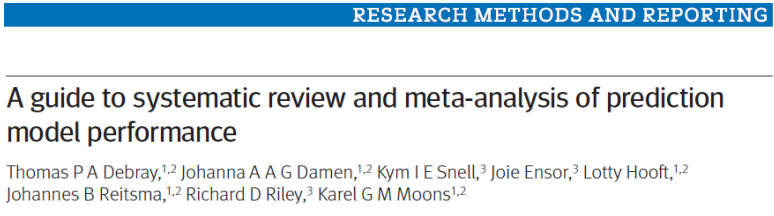
Answers:

- 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 frequentistic 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-variates – or largely for the same co-variates – 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. Meta-analysis of prognostic factor studies
7. Meta-analysis of prognostic model studies
8. Interpretation and conclusions



Identifying prognostic factor studies

- Identification of prognostic studies can be quite challenging!
- Lack of (PF-) specific search filters
- PF of interest not necessarily named as such
- Poor reporting of PF studies
- In addition to prognostic factor, look out for forms such as prediction, predictive, sometimes even etiologic factor

Hints and tips for screening

- A PF is usually an attribute of the condition or of the patient
- Within a PF-SR, patients should have similar baseline of illness or health (distinguish between risk and prognosis)
- PFs may be dichotomous (or many categories) or continuous; interestingly, a PF may possibly be treated as continuous and categorised/dichotomous
- The PF must never be the reason for treatment modification within a study
- A basic PF analysis would be univariable (unadjusted), but adjusted analysis is often included in a study
- Distinguish between a PF and a prognostic model study. Adjusted analysis (simple regression analysis does not constitute a prognostic model)

PICOTS

P = Patients with Hodgkin Lymphoma

I = Interim-PET

C = Any other PF (multivariable/adjusted analysis)

O = Overall Survival (OS), Progression-Free Survival (PFS)

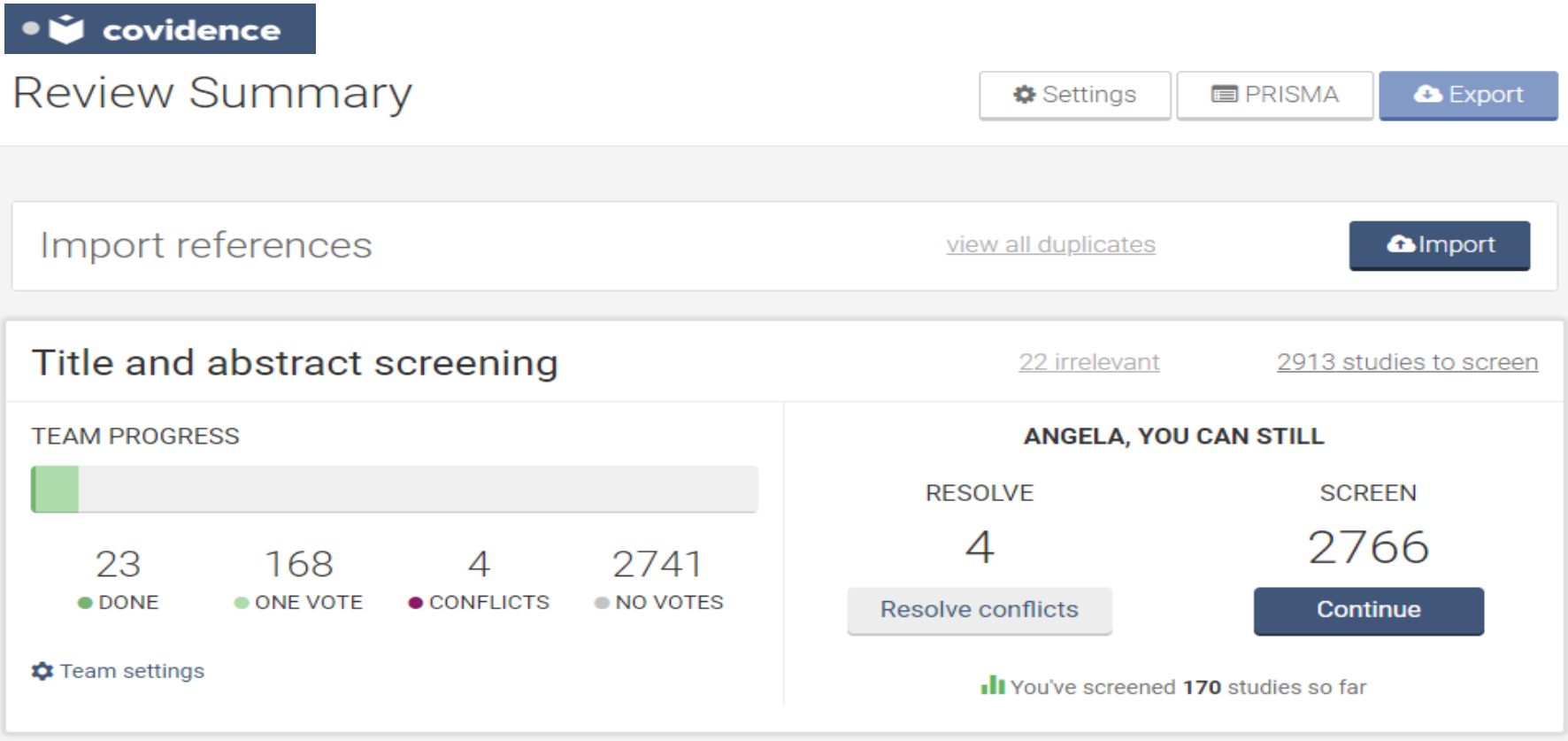
T = after 2 (or more) cycles of chemotherapy

S = Hospital/treating center

Interim-PET example

- Previously untreated patients with Hodgkin Lymphoma
- Interim-PET scan results after a few cycles of chemotherapy as a prognostic factor
- Interim-PET distinguishes between
  - PET-positive patients (= poor prognosis)
  - PET-negative patients (= good prognosis)

Screening with Covidence

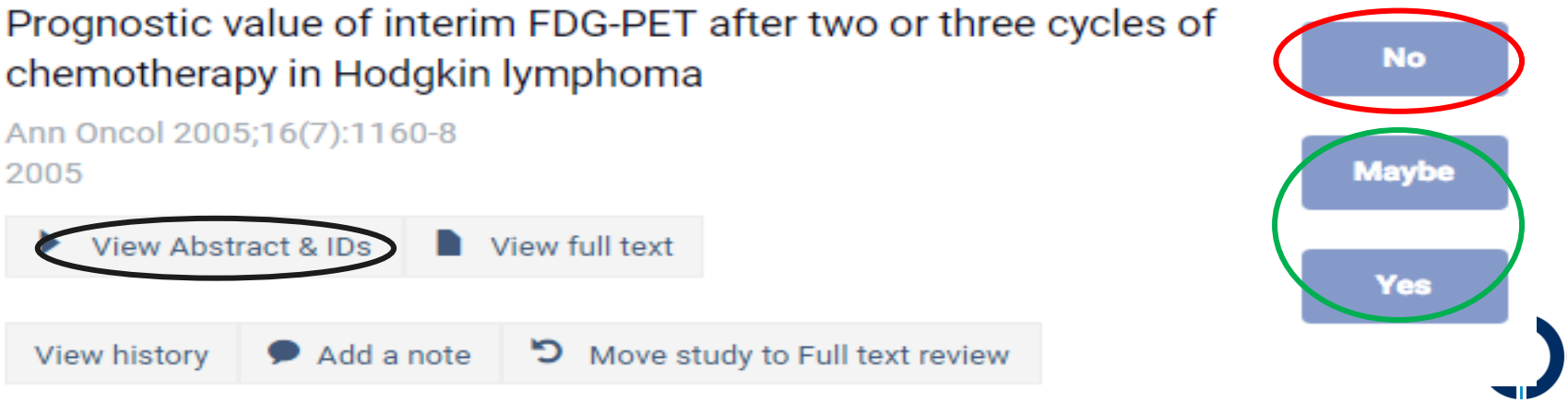


Picture source: Covidence

Step 1: Title and abstract screening

Two independent reviewers:

- If yes/maybe + yes/maybe → moves to full-text screening
- If yes/maybe + no → moves to conflicts
- If no + no → irrelevant (excluded)



Picture source: Covidence

Title and abstract screening

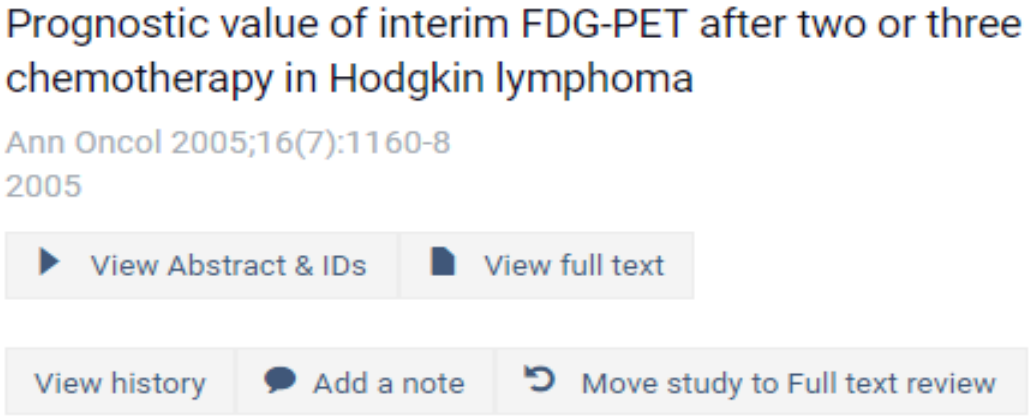
Inclusion criteria:

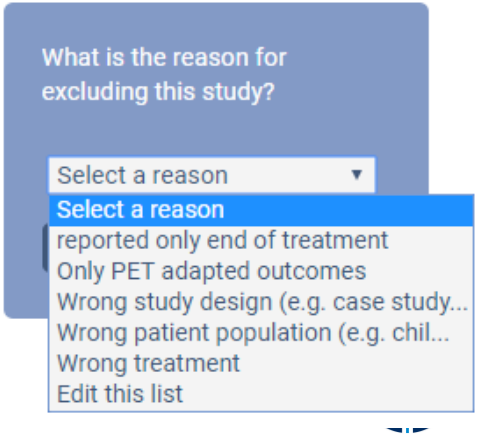
- Patients newly diagnosed with Hodgkin Lymphoma (HL)
- Receiving first-line chemotherapy (e.g. ABVD)
- Interim-PET scan after a few (2, 3 or 4) cycles of chemotherapy
- No treatment modification according to the result of the interim PET
- Outcome of interest: Overall Survival (OS) and Progression-Free Survival (PFS)
- Hint: look out for the red and green highlighted keywords

Step 2: full text screening

Two independent reviewers:

- If include + include → study included
- If include + exclude or exclude + exclude (different reasons) → conflict
- If exclude + exclude (same reason) → excluded reference





Picture source: Covidence

Full-text screening

Hints (despite abstract, the full text provides information that determines final eligibility)

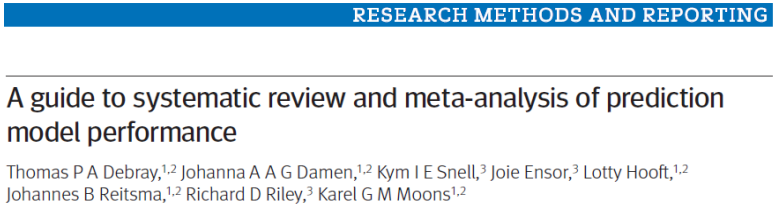
- Methods vs. analysis: what they planned vs. what they actually did
- Do not base inclusion on our outcomes of interest
- Was the treatment adapted after interim PET was conducted?
- Are really only HL patients evaluated?
- Are patients with recurrent disease included?
- When was the PET conducted (interim or maybe at the end of chemotherapy?)

Exercise **answers**

Included	Excluded
Barnes 2011 – PET2 or 3 (included, although some indication of additional RT for some patients stated in the discussion section. In cases like this, authors should be contacted to ask for clarification (and/or additional/separate data) in order to decide upon inclusion). Cerci 2010 – PET 2 Simon 2015 – PET2 Zinzani 2012 – PET 2 Zinzani 2006 – associated with Zinzani 2012 (includes subpopulation) Kobe 2018 – RCT	Dann 2010 – Interim-PET result adapted therapy EL-Galaly 2012 – wrong study design (inclusion criteria were patients who had already received first-line treatment) Dann 2017 – Interim-PET result adapted therapy Kobe 2014 – wrong study design (end-of-chemotherapy PET)

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. Meta-analysis of prognostic factor studies
7. Meta-analysis of prognostic model studies
8. Interpretation and conclusions



CHARMS

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

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

Karel G. M. Moons[1¶*], Joris A. H. de Groot[1¶], Walter Bouwmeester[1], Yvonne Vergouwe[1], Susan Mallett[2], Douglas G. Altman[3], Johannes B. Reitsma[1], Gary S. Collins[3]

  - Does not exist for prognostic factor studies – though can use CHARMS for that

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 **for inclusion** in multivariable modelling (e.g., all candidate predictors, pre-selection based on unadjusted association with the outcome)	
	Method for selection of predictors **during multivariable modelling** (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 **for inclusion** in multivariable modelling (e.g., all candidate predictors, pre-selection based on unadjusted association with the outcome)	
	Method for selection of predictors **during multivariable modelling** (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.	

Hints and tips for data extraction

- For any MA the first rule is to pool similar studies ie record
  - ✓ different ways of measuring the PF
  - ✓ different timings at which PF is measured
  - ✓ different timings at which outcomes are measured
  - ✓ the challenge (or is it the peril?) of different cutpoints
  - ✓ the covariates, if an adjusted regression model is included
- The usual effect estimate is a univariable time-to-event, e.g. Hazard Ratio (HR)
- Data may not be readily available, and would need to be estimated, hence some statistical knowledge may be necessary for which items to extract, e.g. data from graphs may be needed

PICOTS

P = Patients with Hodgkin Lymphoma

I = Interim-PET

C = Any other PF (multivariable/adjusted analysis)

O = Overall Survival (OS), Progression-Free Survival (PFS)

T = after 2 (or more) cycles of chemotherapy

S = Hospital/treating center

Data extraction (Exercise 1)

1. Split in groups and read one of the provided studies (abstract, methods, results, discussion):
   - Simon (2015) “Combined prognostic value of absolute lymphocyte/monocyte ratio in peripheral blood and interim PET/CT results in Hodgkin lymphoma”
   - Barnes (2011): “End-of-treatment but not interim PET scan predicts outcome in nonbulky limited-stage Hodgkin’s lymphoma“
2. Extract the results for interim PET as a prognostic factor

Data extraction made easy for you

- Data extraction form is colour-coded (orange and purple sections)
- Your facilitator will allocate you in groups and colours
- Work within your group, fill in as much as you can, but focus on your group's colour

Data extraction form (CHARMS checklist)

Analysis	
Univariable analysis	
Total number of patients included in univariable analysis	
Method used	
Comments	
How was the factor treated?	
Comments	
Criteria for choice of cut points	
Comments	
OS in univariable analysis/unadjusted	
Univariable analysis for OS?	
Comments	
Summary of follow-up	
Duration of follow-up **(in months)**	
Total no. of patients included in OS analysis	
Total no. of PET+ patients	
Total no. of PET- patients	

Exercise 2

Discuss in groups of two (5 min.): Should we pool results from different studies for progression-free survival in multivariable analysis? Why, why not?

	PFS ($n = 121$)		
	HR	95 % CI	*P*
Multivariate analysis			
LMR (≤2.11)	4.39	1.87–10.27	**0.001**
PET2 (positive)	17.74	6.61–47.57	**<0.001**

Cox regression analysis

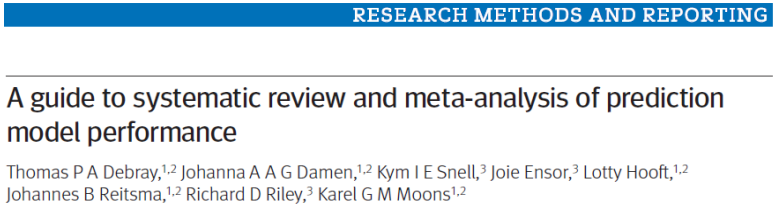
		PFS	
		Univariate	Multivariate
Parameter		HR (95% CI) *P*	HR (95% CI) *P*
Symptoms	A	1	1
	B	4.12 (1.26-13.49) .019	2.87 (0.86-9.59) .09
PET result	Negative	1	1
	Positive	9.11 (2.67-31.12) .0001	7.64 (2.19-26.69) .001

Table 4. Multivariate analyses of progression-free survival

			95% CI for HR	
	P	HR	Lower	Upper
Clinical stage, extranodal disease and early interim FDG-PET				
Step 1				
Clinical stage	.378	1.397	0.664	2.939
Extranodal disease	.346	2.018	0.469	8.678
PET2 (qualitative)	< .001	36.281	7.179	183.4

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. **Meta-analysis of prognostic factor studies**
7. **Meta-analysis of prognostic model studies**
8. **Interpretation and conclusions**



Risk of Bias tools

Prognostic factor/predictor finding studies

- **QUIPS** → J Haydn, Ann Int Med 2006 + 2013

Prediction model studies (development and validation)

- **PROBAST** – Ann Int Med (fall 2018)

QUIPS (J Hayden, Ann Int Med 2006 + 2013)

Table 2. Domains Included in the Framework of Potential Biases and the Proportion of Reviews Assessing the Biases*

Potential Bias	Studies Adequately Assessing Bias, %†	Domains Addressed	Studies Assessing Domain, %
1. The study sample represents the population of interest on key characteristics, sufficient to limit potential bias to the results (study participation).	55	1. Source population clearly defined 2. Study population described 3. Study population represents source population or population of interest	50 21 50
2. Loss to follow-up (from sample to study population) is not associated with key characteristics, sufficient to limit potential bias (i.e., the study data adequately represent the sample) (study attrition).	42	4. Completeness of follow-up described 5. Completeness of follow-up adequate	19 42
3. The prognostic factor of interest is adequately measured in study participants to sufficiently limit potential bias (prognostic factor measurement).	59	6. Prognostic factors defined 7. Prognostic factors measured appropriately	31 59
4. The outcomes of interest are adequately measured in study participants to sufficiently limit potential bias (outcome measurement).	51	8. Outcome defined 9. Outcome measured appropriately	42 51
5. Important potential confounders are appropriately accounted for, limiting potential bias with respect to the prognostic factor of interest (confounding measurement and account).	13	10. Confounders defined and measured 11. Confounding accounted for	21 53
6. The statistical analysis is appropriate for the design of the study, limiting potential for presentation of invalid results (analysis).	33	12. Analysis described 13. Analysis appropriate 14. Analysis provides sufficient presentation of data	8 33 32

* Data are from 163 prognostic systematic reviews with quality items that could be extracted.

†Adequate assessment included 1) study participation: "source population clearly defined" and "study population described" or "study population represents source

Six Opportunities for Bias

1. Study participation
2. Study attrition
3. Prognostic factor measurement
4. Outcome measurement
5. Covariate measurement & accounting
6. Analysis & presentation

Julius Center
UMC Utrecht

PROBAST

Prediction model Risk Of Bias Assessment Tool

Karel Moons, Robert Wolff, Penny Whiting, Richard Riley, Gary Collins, Johannes Reitsma, Marie Westwood, Jos Kleijnen, Sue Mallett

Annals of Internal Medicine 2018

Structure of PROBAST

- Also domain-based: each with risk of bias + applicability
- Follows QUADAS-2, ROBINS-I, ROB 2.0 tool

Bias Likelihood that a prediction model leads to distorted predictive performance for its intended use in the targeted individuals.

Applicability refers to extent to which prediction model from primary study matches your systematic review question, in terms of participants, predictors or outcomes of interest

PROBAST 4 phases

Step	Task	When to complete
1	Specify your systematic review question	Once per systematic review
2	Classify the type of prediction model evaluation	Once for each model of interest in each publication being assessed, for each relevant outcome
3	Assess risk of bias and applicability	Once for each evaluation (development and/or validation) of each distinct model
4	Overall judgment	Once for each evaluation (development and/or validation) of each distinct model

Cochrane PMG title registration form for SRs of prognostic studies

Cochrane Methods Prognosis

Prognosis Studies review proposal form

Review Proposal Form

Please complete this form to outline your proposal for a Cochrane systematic review. Email the completed form to [email address], or send to [name], Managing Editor, Cochrane XXX Group, [postal address]. Ph: +XX XXXXXXXXX Fax: +XX XXXXXXXXX.

Before completing this form:

- Read "Managing expectations: what does The Cochrane Collaboration expect of authors, and what can authors expect of The Cochrane Collaboration?" (see http://community.cochrane.org/editorial-and-publishing-policy-resource/cochrane-review-development/managing-expectations Note: this information is particularly for systematic reviews of intervention studies. A page for prognosis reviews is under construction.)
- Note that a Cochrane review of prognosis studies clearly differs from that of intervention studies and diagnostic test accuracy studies, in, e.g., searching, data extraction, critical appraisal and meta-analysis. Step-by-step guidance to help you understanding prognosis studies and the processes of conducting a review of prognosis studies is given in the papers in the reference list below.
- Cochrane reviews of prognosis require a multidisciplinary team. Below you find several question addressing the available expertise in the author team, and whether external expertise (e.g. from information specialists or methodologists) is needed to conduct this review. If additional expertise is needed, e.g. an information specialist, or methodological or statistical expertise, please provide this request to the Prognosis Methods Group (PMG) timely.

Proposed title

Choose one of the formats below. See also the generic guidance on defining a review question for prognosis studies in the CHARMS checklist.

Incidence of [outcome] within [time] in [population]
[Prognostic factors] for predicting incidence of [outcome] in [population]
Prediction of [outcome] in [population] using [prognostic factors]
Prognostic models for predicting [outcome] in [population]
Predictive performance of [prognostic model] for predicting [outcome] in [population]
Added value of [prognostic factor] on top of [existing prognostic factors/prognostic model] for predicting [outcome] in [population]
[Predictive factors] predicting the [outcome of treatment] in [population]
[Factors / Models] predicting differential treatment response in [population]
[Factors / Models] for predicting treatment response in [population]

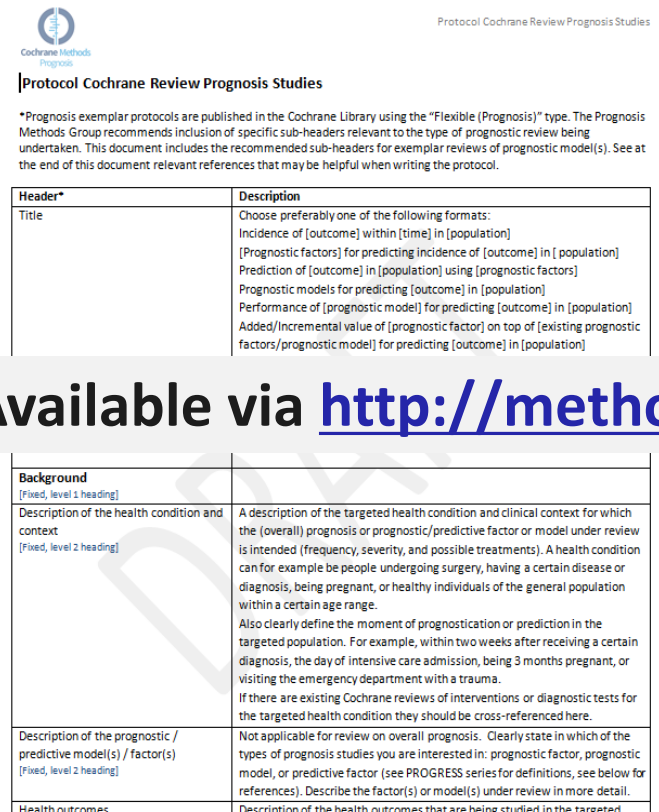
Short description of review proposal

Provide brief but enough information to make sure that the clinical context and the actual question that is being asked is clear for non-content experts as well.

For explicit guidance to help filling in this title registration form and for the conduct of the review, from framing the review question, search strategy, study in/exclusion criteria, critical appraisal, risk of bias assessment, meta-analysis and reporting, please see the papers mentioned in the reference list below.

Type of prognosis review *Indicate what type of review you are going to perform (double click to check a box). See PROGRESS series in the reference list.*	☐ Overall prognosis ☐ Prognostic factors ☐ Prognostic models ☐ Predictive/Treatment selection factors
Motivation for the review *For example, is this going to be part of a PhD thesis; is it part of a larger project; is it particularly topical at the present time?*	
Background *i) The clinical problem. A short description of the existing clinical pathway of the targeted individuals/patients; their starting condition and moment of prognostication (time point in the clinical pathway); what prognostic outcomes are relevant to the targeted individuals. For predictive factor reviews also refer to the role of treatment.* *ii) Why is this review relevant, including how might the results of the review be used: e.g., the prognostic or predictive factor(s) or model(s) under review may be used to determine treatment allocation or abstention, decide on closer follow-up or monitoring, etc. Reference to an existing systematic review on this topic outside Cochrane is helpful.*	
Review objective(s) *What is the review question, according to the PICOTS format? (see Box 1 in the paper of Debray et al, BMJ 2017, see reference list below.)*	Primary objective: Secondary objective(s):
Participants / setting *Short outline of the targeted population and clinical setting, to be included and excluded for the review.*	

Cochrane PMG Protocol Template for SRs of prognostic studies

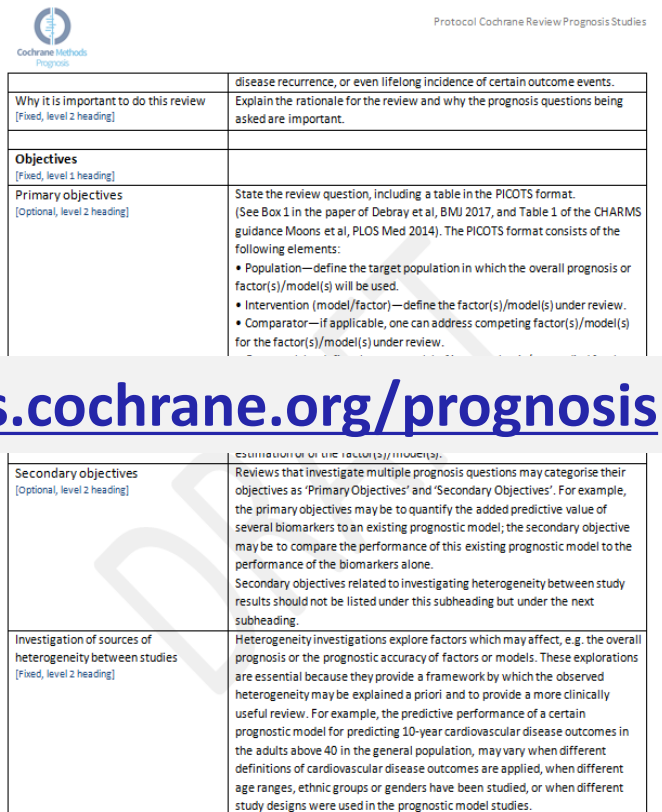


Protocol Cochrane Review Prognosis Studies

Protocol Cochrane Review Prognosis Studies

*Prognosis exemplar protocols are published in the Cochrane Library using the "Flexible (Prognosis)" type. The Prognosis Methods Group recommends inclusion of specific sub-headers relevant to the type of prognostic review being undertaken. This document includes the recommended sub-headers for exemplar reviews of prognostic model(s). See at the end of this document relevant references that may be helpful when writing the protocol.

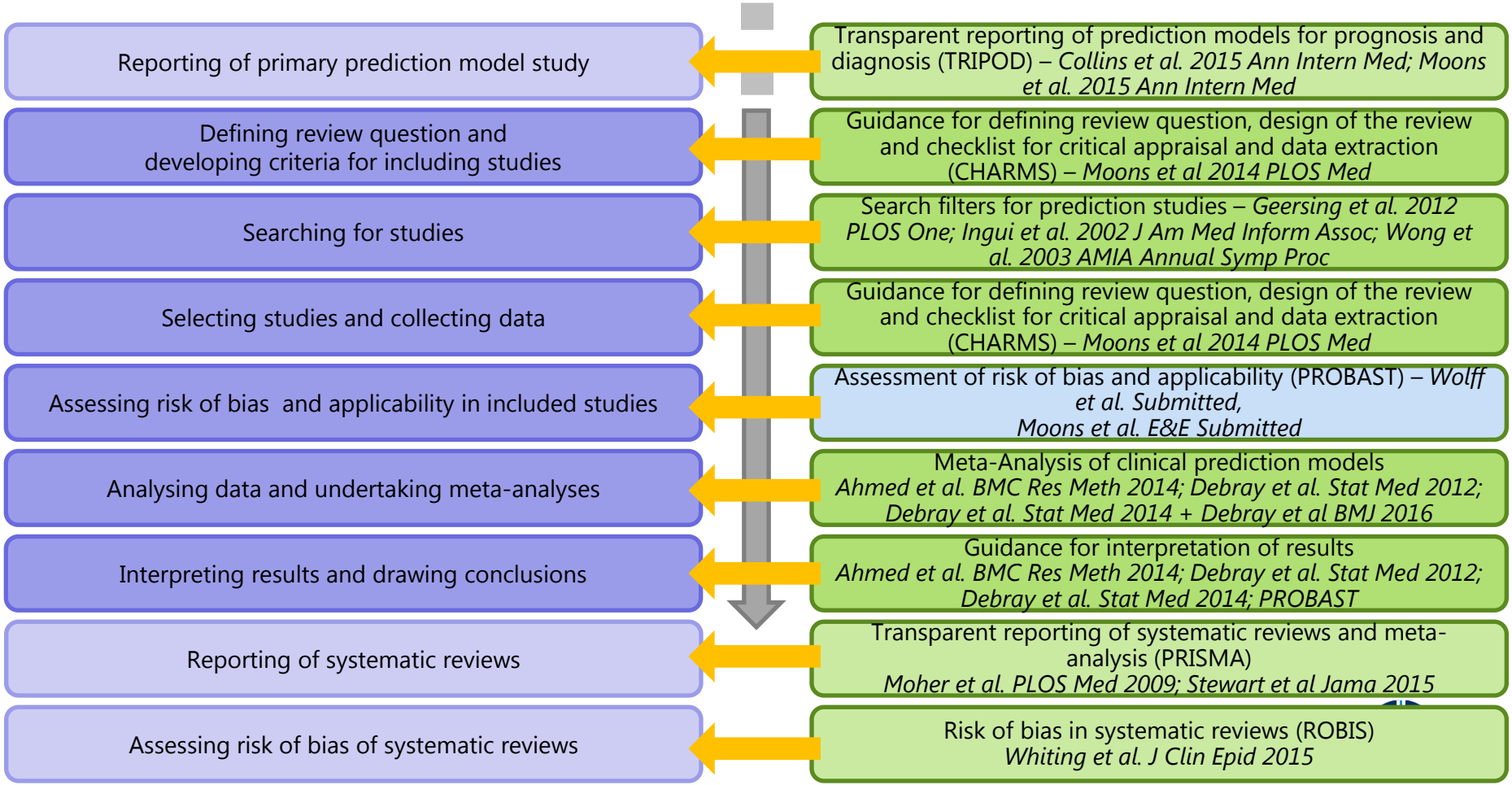
Header*	Description
Title	Choose preferably one of the following formats: Incidence of [outcome] within [time] in [population] [Prognostic factors] for predicting incidence of [outcome] in [population] Prediction of [outcome] in [population] using [prognostic factors] Prognostic models for predicting [outcome] in [population] Performance of [prognostic model] for predicting [outcome] in [population] Added/Incremental value of [prognostic factor] on top of [existing prognostic factors/prognostic model] for predicting [outcome] in [population]
Background [Fixed, level 1 heading]	
Description of the health condition and context [Fixed, level 2 heading]	A description of the targeted health condition and clinical context for which the (overall) prognosis or prognostic/predictive factor or model under review is intended (frequency, severity, and possible treatments). A health condition can for example be people undergoing surgery, having a certain disease or diagnosis, being pregnant, or healthy individuals of the general population within a certain age range. Also clearly define the moment of prognostication or prediction in the targeted population. For example, within two weeks after receiving a certain diagnosis, the day of intensive care admission, being 3 months pregnant, or visiting the emergency department with a trauma. If there are existing Cochrane reviews of interventions or diagnostic tests for the targeted health condition they should be cross-referenced here.
Description of the prognostic / predictive model(s) / factor(s) [Fixed, level 2 heading]	Not applicable for review on overall prognosis. Clearly state in which of the types of prognosis studies you are interested in: prognostic factor, prognostic model, or predictive factor (see PROGRESS series for definitions, see below for references). Describe the factor(s) or model(s) under review in more detail.
Health outcomes	Description of the health outcomes that are being studied in the targeted



Protocol Cochrane Review Prognosis Studies

	disease recurrence, or even lifelong incidence of certain outcome events.
Why it is important to do this review [Fixed, level 2 heading]	Explain the rationale for the review and why the prognosis questions being asked are important.
Objectives [Fixed, level 1 heading]	
Primary objectives [Optional, level 2 heading]	State the review question, including a table in the PICOTS format. (See Box 1 in the paper of Debray et al, BMJ 2017, and Table 1 of the CHARMS guidance Moons et al, PLOS Med 2014). The PICOTS format consists of the following elements: • Population—define the target population in which the overall prognosis or factor(s)/model(s) will be used. • Intervention (model/factor)—define the factor(s)/model(s) under review. • Comparator—if applicable, one can address competing factor(s)/model(s) for the factor(s)/model(s) under review.
Secondary objectives [Optional, level 2 heading]	Reviews that investigate multiple prognosis questions may categorise their objectives as 'Primary Objectives' and 'Secondary Objectives'. For example, the primary objectives may be to quantify the added predictive value of several biomarkers to an existing prognostic model; the secondary objective may be to compare the performance of this existing prognostic model to the performance of the biomarkers alone. Secondary objectives related to investigating heterogeneity between study results should not be listed under this subheading but under the next subheading.
Investigation of sources of heterogeneity between studies [Fixed, level 2 heading]	Heterogeneity investigations explore factors which may affect, e.g. the overall prognosis or the prognostic accuracy of factors or models. These explorations are essential because they provide a framework by which the observed heterogeneity may be explained a priori and to provide a more clinically useful review. For example, the predictive performance of a certain prognostic model for predicting 10-year cardiovascular disease outcomes in the adults above 40 in the general population, may vary when different definitions of cardiovascular disease outcomes are applied, when different age ranges, ethnic groups or genders have been studied, or when different study designs were used in the prognostic model studies.

Available via http://methods.cochrane.org/prognosis



Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 - http://handbook.cochrane.org/

Example SRs of prognostic factor studies

Is P53 a prognostic (molecular) marker for bladder cancer (Malats, Lancet Oncol, 2005)

SR on all molecular and biological prognostic markers of tumours in the Ewing's Sarcoma Family (Riley Eur J Canc 2003)

Prognostic markers for death or tumour recurrence in patients with neuroblastoma (Riley, Br J Canc 2003)

Prognostic factors of sequelae and death after bacterial meninigitis in childhood (de Jonge, BMC Infectious Diseases 2010)

Added value of carotid imaging markers in the prediction of fatal or non-fatal CVD events in general population (Peters, Heart 2011)

Examples SRs of prognostic modeling studies

Risk prediction models for the development of type 2 DM: SR (Collins 2011, Plos Med)

Prediction models for CVD in patients with type 2 DM (van Dieren, Heart 2011)

Risk prediction models for prolonged ICU stay after cardiac surgery (Ettema, Circulation 2010)

Risk prediction models for outcome after traumatic brain injury (Perel BMC Med Informatic and Decis Making 2006)

Prediction models for cardiovascular disease risk in the general population: systematic review (Damen, BMJ 2016)

Other workshops

- Systematic reviews of prognostic studies II: risk of bias assessment in systematic reviews of prognostic studies (Sunday, September 16th, 16:00)
- Systematic reviews of prognostic studies III: meta-analytical approaches in systematic reviews of prognostic studies (Monday, September 17th, 11:00)
- Systematic reviews of prognostic studies IV: meta-analysis of prognostic studies using individual participant data (Tuesday, September 18th, 11:00).
- Systematic reviews of prognostic studies I: introduction, design and protocol for systematic reviews of prognostic studies (Sunday, September 16th, 11:00)