



UMC Utrecht
Julius Center

Systematic review and meta-analysis of prognosis studies

Thomas Debray, PhD

Julius Center for Health Sciences and Primary Care

Cochrane Netherlands

Prognosis research

Prognosis research seeks to understand and improve future outcomes in people with a given disease or health condition.



Prognosis research

Investigation of the relations between future outcomes (endpoints) among people with a given baseline health state (startpoint) in order to improve health

- To provide evidence for translating findings from the laboratory to humans
- To provide evidence for translating findings from clinical research to clinical practice



Types of prognosis studies

- 1. Fundamental prognosis research:** What is most likely course (outcome) of individuals in certain health condition (often certain disease)?
- 2. Prognostic factor research:** Which factors are associated with specific outcome in individuals with certain health condition?
- 3. Prognostic model research:** What combination of prognostic factors predict, and how well, a certain outcome in individuals with a certain health condition?
- 4. Stratified medicine research:** Which factors lead to/predict different treatment effect/response in individuals to be treated?



Types of prognosis studies

1. **Fundamental prognosis research:** What is most likely course (outcome) of individuals in certain health condition (often certain disease)?
2. **Prognostic factor research:** Which factors are associated with specific outcome in individuals with certain health condition?
3. **Prognostic model research:** What combination of prognostic factors predict, and how well, a certain outcome in individuals with a certain health condition?
4. **Stratified medicine research:** Which factors lead to/predict different treatment effect/response in individuals to be treated?



Systematic review of prognosis studies

Need for evidence synthesis

- Number of studied prognostic factors increases per day
 - Biomarkers
 - Prognostic models
- Most studies conflicting results
 - much more than in therapeutic trials and in diagnostic test accuracy studies
- Relatively small studies
 - Kyzas Eur J Canc 2007; > 1500 studies cancer prognostic markers in 2005 → largest just over 1000 pts



Systematic review of prognosis studies



Maarten van Smeden @MaartenvSmeden · Mar 17
When are we going to stop using the word *validated* for prediction models to mean *valid*? Very few validated prediction models are actually valid

15 replies 9 retweets 65 likes

Ewout Steyerberg @ESteyerberg [Follow](#)

Replying to @MaartenvSmeden

Yes! We should assess performance of [#clinicalpredictionmodels](#) across a wide range of settings, and even then it is usually a leap of faith that a model is "valid" for a specific, new, setting.

7:05 AM - 17 Mar 2019

1 Retweet 14 Likes

1 reply 1 retweet 14 likes



Systematic review of prognosis studies

Synthesis of **published** prognosis studies may help


- To identify promising markers
 - By summarizing their (incremental) prognostic value
 - By exploring sources of between-study heterogeneity
- To identify promising prediction models
 - By summarizing their predictive performance
 - By exploring generalizability across different settings and populations
 - By evaluating the need for further improvements
- To improve estimation of prediction models
 - By avoiding overfitting in small samples



Systematic review of prognosis studies

Commentary | Open Access | Open Peer Review

The increasing need for systematic reviews of prognosis studies: strategies to facilitate review production and improve quality of primary research

Johanna A. A. G. Damen   and Lotty Hoofst

Diagnostic and Prognostic Research 2019 3:2

<https://doi.org/10.1186/s41512-019-0049-6> | © The Author(s) 2019

Received: 5 September 2018 | Accepted: 11 January 2019 | Published: 23 January 2019



Systematic review of prognosis studies

Research Methods & Reporting

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

BMJ 2019 ; 364 doi: <https://doi.org/10.1136/bmj.k4597> (Published 30 January 2019)

Cite this as: *BMJ* 2019;364:k4597

Research Methods & Reporting

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

BMJ 2017 ; 356 doi: <https://doi.org/10.1136/bmj.i6460> (Published 05 January 2017)

Cite this as: *BMJ* 2017;356:i6460

Systematic review of prognosis studies

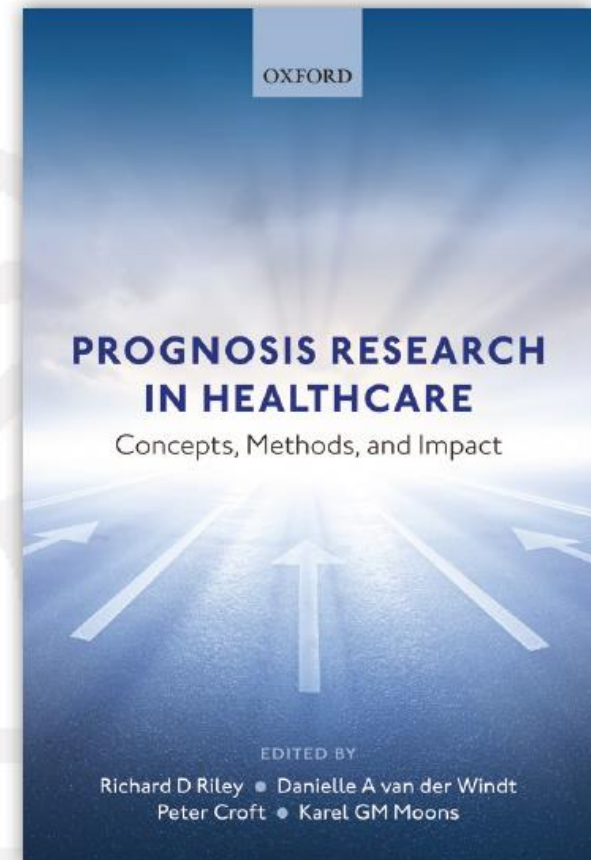
PROGNOSIS RESEARCH IN HEALTHCARE

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



Systematic review of prognosis studies

1. Well-formulated review question (PICO)
2. Extensive search for studies
3. Objective selection of studies
4. Objective extraction of data
5. Critical appraisal of methodological quality
6. Synthesis of data (meta-analysis)
7. Interpretation, conclusions, recommendations



Systematic review of prognosis studies

Formal review steps and tools

- Defining the review question ([PICOTS](#))
- Defining the search strategy
- Quality appraisal
 - Checklist for prognostic factor studies ([QUIPS](#))
 - Checklist for prognostic model studies ([PROBAST](#))
- Data extraction & meta-analysis
 - Focus on unadjusted and adjusted prognostic effects
 - Focus on model discrimination and calibration
- Interpretation ([GRADE](#))
- Reporting (guidelines: [REMARK](#), [PRISMA](#), [TRIPOD](#))

Prognostic factor studies

Systematic review & meta-analysis

What are prognostic factors?

Any information that, among people with a given health condition, is associated with a subsequent health outcome

- Routinely collected patient characteristics
 - Sex, Age, Body mass index, smoking status, blood pressure
 - Co-morbidities
 - Symptoms
- Biomarkers
 - Blood
 - Urine
 - Imaging
 - Electrophysiological
 - Physiological variables



Need for evidence synthesis

- There are many studies investigating prognostic factors
- There is often conflicting evidence about the prognostic value of certain variables
- The quality of many prognostic studies is poor

Education And Debate

Systematic reviews in health care

Systematic reviews of evaluations of prognostic variables

BMJ 2001 ; 323 doi: <https://doi.org/10.1136/bmj.323.7306.224> (Published 28 July 2001)

Cite this as: *BMJ* 2001;323:224

[Article](#)

[Related content](#)

[Metrics](#)

[Responses](#)


Douglas G Altman, director

Example review

A systematic review was performed

- to evaluate the quality of prognostic research evidence
- for the association of C-reactive protein (CRP)
- with fatal and nonfatal events
- among patients with stable coronary disease

Evaluating the Quality of Research into a Single Prognostic Biomarker: A Systematic Review and Meta-analysis of 83 Studies of C-Reactive Protein in Stable Coronary Artery Disease

Harry Hemingway , Peter Philipson, Ruoling Chen, Natalie K. Fitzpatrick, Jacqueline Damant, Martin Shipley, Keith R. Abrams, Santiago Moreno, Kate S. L. McAllister, Stephen Palmer, Juan Carlos Kaski, Adam D. Timmis, Aroon D. Hingorani



Step 1. Well-formulated review question

Guidance frame review question: CHARMS checklist

Item	Comments
1. <u>P</u>opulation	Target population in which the prognostic factor(s) under review will be used.
2. <u>I</u>ndex prognostic factor(s)	Index prognostic factor(s) for which the prognostic ability is under review.
3. <u>C</u>omparator prognostic factor(s)	Comparator prognostic factors can be considered in various ways. E.g. aim could be to compare prognostic ability of certain index factor with two or more other (that is, comparator) prognostic factors; or to review adjusted prognostic value of certain index factor over and above (adjusted for, independent of) other existing (comparator) prognostic factors. If aim is summarise unadjusted prognostic effect of certain index factor, then no comparator factor is being considered.
4. <u>O</u>utcome(s)	Outcome(s) of interest for the index factor(s) under review.
5. <u>T</u>iming (two elements)	(i) at what time-point(s) prognostic factors (index and comparators) are to be used (time point of prognostication); (ii) over what time period outcome(s) are predicted.
6. <u>S</u>etting	Define intended setting (role) of prognostic factor(s) under review.

Step 1. Well-formulated review question

CRP review

- **Population:** patients with stable coronary disease
- **Index prognostic factor:** C-reactive protein (CRP)
- **Comparative prognostic factors:** Adjustment for age, sex, smoking status, obesity, diabetes, and one or more lipid and inflammatory markers
- **Outcome:** coronary, cardiovascular, and all cause mortality.
- **Timing:** CRP measurement \geq 2 weeks after diagnosis. All follow-up information on the outcomes was extracted
- **Setting:** to provide prognostic information about patients diagnosed with coronary heart disease in primary & secondary care



Step 2. Extensive search for studies

CRP review

- Databases
 - MEDLINE (between 1966 and 25 November 2009)
 - EMBASE (between 1980 and 17 December 2009)
- Search string
 - terms for coronary disease, prognostic studies, and CRP.
- Search results
 - 1566 hits
 - 83 fulfilled the inclusion criteria.



Step 4. Objective extraction of data

CHARMS-PF 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)	

Step 4. Objective extraction of data

Key elements to extract for each factor of interest

- Estimates of the (adjusted) prognostic effect
 - risk ratio or odds ratio (for binary outcomes)
 - hazard ratio (for time-to-event outcome)
 - mean difference (for continuous outcomes such as pain or depression score)
- Corresponding standard errors or confidence intervals
- Other estimates of “incremental” value (e.g. changes in c-statistic, NRI, ...)



Step 4. Objective extraction of data

When not reported, unadjusted hazard ratios (and their variances) can be estimated:

- using the number of outcomes (events) and an available p -value (e.g. from a log-rank test or Cox regression)
- from survival proportions



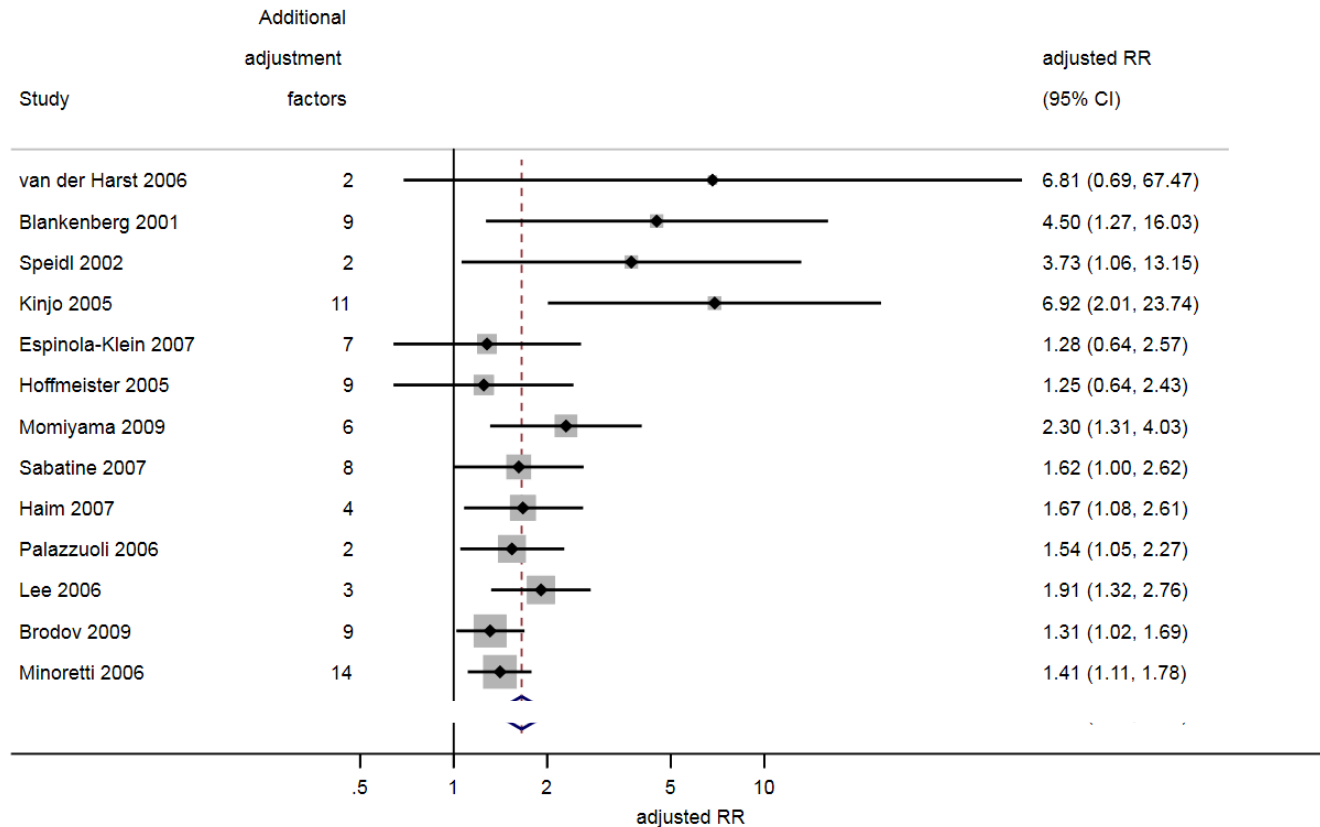
Step 4. Objective extraction of data

CRP review

- “We extracted the reported relative risk, odds ratio or hazard ratio, and 95% CIs from each study.”
- “We extracted the relative risks with the largest number of adjustment variables”
- “We converted the reported relative risk estimates onto a standard scale of effect, comparing the highest third with the lowest third of the CRP distribution”



Step 4. Objective extraction of data



Forest plot showing the study-specific estimates and meta-analysis summary result of the adjusted prognostic effect (risk ratio) of CRP. All studies were minimally adjusted for age, gender, smoking, diabetes, obesity, and lipids.



Step 5. Critical appraisal

QUIPS tool

Table 2 | QUIPS tool (quality in prognostic factor studies), which can be used to classify risk of bias of prognostic factor studies

Domains	Signalling items	Risk of bias ratings
1. Study participation	<ul style="list-style-type: none"> (a) Adequate participation in the study by eligible persons (b) Description of the target population or population of interest (c) Description of the baseline study sample (d) Adequate description of the sampling frame and recruitment (e) Adequate description of the period and place of recruitment (f) Adequate description of inclusion and exclusion criteria 	<p>High: the relationship between the PF and outcome is very likely to be different for participants and eligible non-participants</p> <p>Moderate: the relationship between the PF and outcome may be different for participants and eligible non-participants</p> <p>Low: the relationship between the PF and outcome is unlikely to be different for participants and eligible non-participants</p>
2. Study attrition	<ul style="list-style-type: none"> (a) Adequate response rate for study participants (b) Description of attempts to collect information on participants who dropped out (c) Reasons for loss to follow-up are provided (d) Adequate description of participants lost to follow-up (e) There are no important differences between participants who completed the study and those who did not 	<p>High: the relationship between the PF and outcome is very likely to be different for completing and non-completing participants</p> <p>Moderate: the relationship between the PF and outcome may be different for completing and non-completing participants</p> <p>Low: the relationship between the PF and outcome is unlikely to be different for completing and non-completing participants</p>
3. Prognostic factor measurement	<ul style="list-style-type: none"> (a) A clear definition or description of the PF is provided (b) Method of PF measurement is adequately valid and reliable (c) Continuous variables are reported or appropriate cutpoints are used (d) The method and setting of measurement of PF is the same for all study participants (e) Adequate proportion of the study sample has complete data for the PF (f) Appropriate methods of imputation are used for missing PF data 	<p>High: the measurement of the PF is very likely to be different for different levels of the outcome of interest</p> <p>Moderate: the measurement of the PF may be different for different levels of the outcome of interest</p> <p>Low: the measurement of the PF is unlikely to be different for different levels of the outcome of interest</p>
4. Outcome measurement	<ul style="list-style-type: none"> (a) A clear definition of the outcome is provided (b) Method of outcome measurement used is adequately valid and reliable (c) The method and setting of outcome measurement is the same for all study participants 	<p>High: the measurement of the outcome is very likely to be different related to the baseline level of the PF</p> <p>Moderate: the measurement of the outcome may be different related to the baseline level of the PF</p> <p>Low: the measurement of the outcome is unlikely to be different related to the baseline level of the PF</p>
5. Adjustment for other prognostic factors	<ul style="list-style-type: none"> (a) All other important PFs are measured (b) Clear definitions of the important PFs measured are provided 	<p>High: the observed effect of the PF on the outcome is very likely to be distorted by another factor related to PF and outcome</p>

Step 5. Critical appraisal

CRP review

- The median number of study quality items reported was 7 of a possible 17, and standards did not change between 1997 and 2009
- Only 2 studies referred to a study protocol, with none referring to a statistical analysis plan
- Only 2 studies reported the time elapsed between first lifetime presentation with coronary disease and assessment of CRP and this raised applicability concerns

Step 6. Synthesis of data (meta-analysis)

- Unexplained heterogeneity is likely to exist due to:
 - Publication bias
 - Variation in study design & quality
 - Variation in inclusion criteria
 - Differences in treatments received during follow-up
 - Different types of prognostic effect measures (e.g. odds ratio and hazard ratios)
 - Different sets of adjustment factors
 - ...
- A random effects meta-analysis approach is essential to allow for unexplained heterogeneity across studies

Step 6. Synthesis of data (meta-analysis)

CRP example

Random effects meta-analysis of 13 studies that adjusted for at least all six conventional prognostic factors

- The summary meta-analysis result was
 - Without Hartung-Knapp correction
1.65 (95% CI: 1.39 to 1.96)
 - With Hartung-Knapp correction
1.65 (95% CI: 1.34 to 2.04)
- which gives the average prognostic effect of CRP (for those in the top versus bottom third of CRP distribution),
- which suggests larger CRP values are associated with higher risk.

Step 6. Synthesis of data (meta-analysis)

If substantial heterogeneity is present, consider to ...

- Display the variability in estimates on a forest plot without showing an overall pooled estimate
- Quantify and report the magnitude of heterogeneity
 - Tau squared
 - (Approximate) prediction interval
- Perform subgroup analyses or meta-regression

Step 6. Synthesis of data (meta-analysis)

CRP review

- Studies originally reporting unequal CRP groups had stronger effects than those reporting CRP on a continuous scale
- For each additional adjustment factor, the summary risk ratio decreased by 3%.
- The summary risk ratio was smaller among studies with more than the median number of outcome events,
- The summary risk ratio was smaller among studies confined to stable coronary disease

Examining small-study effects

Systematic difference in prognostic effect estimates for small studies and large studies, e.g. due to

- Selective reporting
- Publication bias
- Between-study heterogeneity

The evidence for small-study can be evaluated in a funnel plot, which shows the study estimates (x-axis) against their precision (y-axis)



Examining small-study effects

- Well-known tests for detecting funnel plot asymmetry suffer from low power or excessive type-I error rates
- Evaluate funnel plot asymmetry in meta-analysis of survival data using the total number of observed events
- The use of funnel plot asymmetry tests should be avoided when there are few studies available

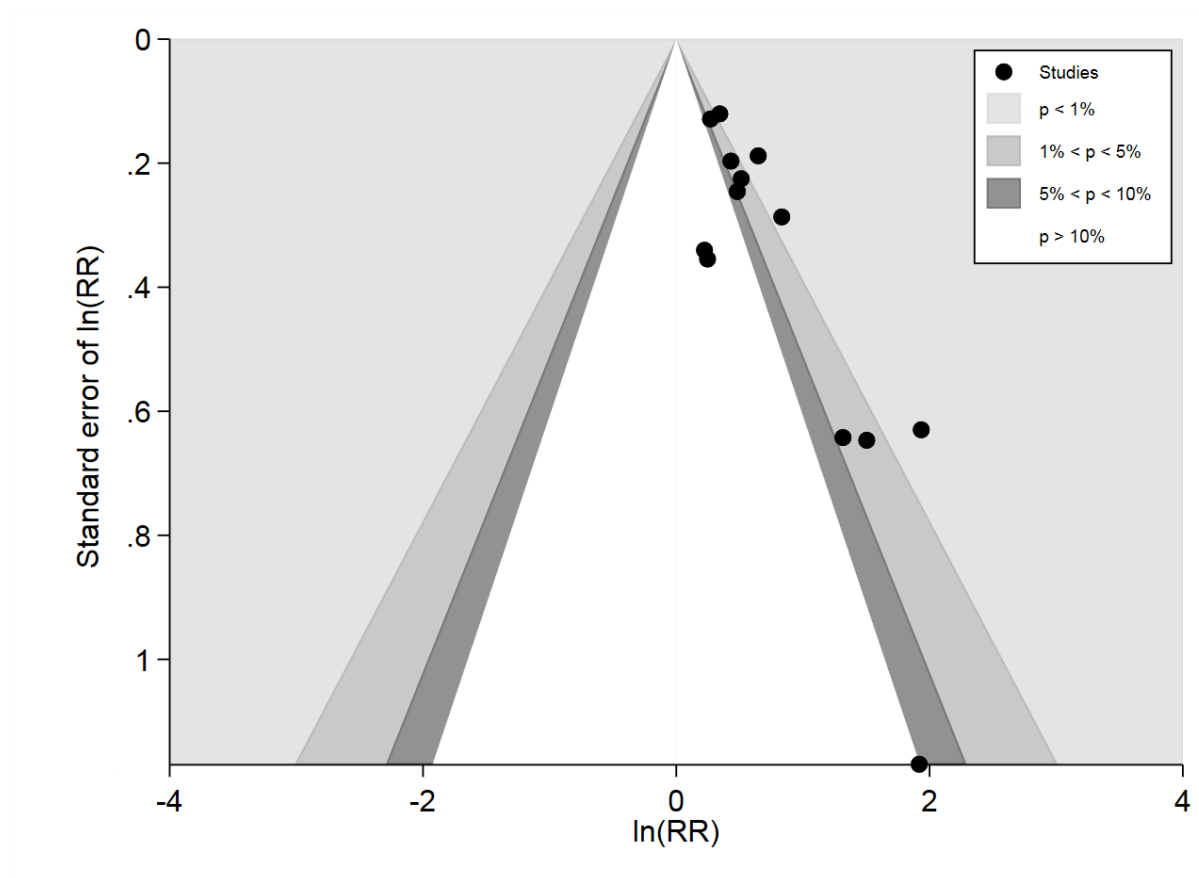
Detecting small-study effects and funnel plot asymmetry in meta-analysis of survival data: A comparison of new and existing tests

Thomas P. A. Debray^{1,2}  | Karel G. M. Moons^{1,2} | Richard D. Riley³



Examining small-study effects

CRP review



Step 7. Interpretation, conclusions, recommendations

GRADE (grades of recommendation, assessment, development, and evaluation)

- Risk of bias
- Inconsistency
- Imprecision
- Indirectness
- Publication bias



Step 7. Interpretation, conclusions, recommendations

CRP review



- Strong concern about the quality and reliability of the underlying evidence
- No firm conclusions about whether CRP has prognostic value after adjustment for established prognostic factors
- The concerns “explicitly challenge the statement for healthcare professionals made by the Centers for Disease Control that measuring CRP is both ‘useful’ and ‘independent’ as a marker of prognosis



Software

metamisc: Diagnostic and Prognostic Meta-Analysis

Meta-analysis of diagnostic and prognostic modeling studies. Summarize estimates of prognostic factors, diagnostic test accuracy and prediction model performance. Validate, update and combine published prediction models. Develop new prediction models with data from multiple studies.

Version: 0.2.0
Depends: R (≥ 3.2.0), stats, graphics
Imports: [metafor](#) (≥ 2.0.0), [mvtnorm](#), [lme4](#), [plyr](#), methods, [pROC](#), [ggplot2](#)
Suggests: [runjags](#), [rjags](#), [logistf](#) (≥ 1.23), [testthat](#) (≥ 1.0.2)
Published: 2019-02-07
Author: Thomas Debray  [aut, cre], Valentijn de Jong  [aut]
Maintainer: Thomas Debray <thomas.debray at gmail.com>
License: [GPL-3](#)
URL: <http://r-forge.r-project.org/projects/metamisc/>
NeedsCompilation: no
In views: [MetaAnalysis](#)
CRAN checks: [metamisc results](#)

Downloads:

Reference manual: [metamisc.pdf](#)
Package source: [metamisc_0.2.0.tar.gz](#)
Windows binaries: r-devel: [metamisc_0.2.0.zip](#), r-release: [metamisc_0.2.0.zip](#), r-oldrel: [metamisc_0.2.0.zip](#)
OS X binaries: r-release: [metamisc_0.2.0.tgz](#), r-oldrel: [metamisc_0.2.0.tgz](#)
Old sources: [metamisc archive](#)

Linking:

Please use the canonical form <https://CRAN.R-project.org/package=metamisc> to link to this page.



Acknowledgements

