From the Department of Public Health Sciences
Karolinska Institutet, Stockholm, Sweden

NOVEL METHODS FOR DOSE-RESPONSE META-ANALYSIS

Alessio Crippa

Karolinska Institutet

Stockholm 2018

All published papers reproduced with permission
Published by Karolinska Institutet
Printed by E-Print AB 2018

Edited in R using knitr

ISBN <include number>

NOVEL METHODS FOR DOSE-RESPONSE META-ANALYSIS

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Alessio Crippa

Principal supervisor:
Associate Professor Nicola Orsini
Karolinska Institutet
Department of Public Health Sciences

Co-supervisor:
Professor Alicja Wolk
Karolinska Institutet
Institute of Environmental Medicine

Professor Matteo Bottai
Karolinska Institutet
Institute of Environmental Medicine

Professor Donna Spiegelman
Harvard T.H. Chan School of Public Health
Department of Epidemiology

Opponent:
Professor Christopher H. Schmid
Brown University
Center for Evidence Based Medicine

Examination board:
Associate Professor Nele Brusselaers
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Associate Professor Antonio Gasparrini
London School of Hygiene and Tropical Medicine
Department of Social & Environmental Health Research

Professor Paul Lambert
University of Leicester
Department of Health Sciences

"The function of the expert reviewer is not to be more right than other people, but to be wrong for more sophisticated reasons."

—Iain Chalmers and Douglas G. Altman

Systematic Reviews, 1995

Abstract

Dose-response meta-analysis is a statistical technique increasingly used to combine and contrast the evidence on the association between a continuous exposure and the risk of a health outcome. Several papers refined selected aspects of the methodology, such as implementations of flexible strategies and considering extensions to multivariate meta-analysis. However, there were still several relevant questions that needed to be addressed. This thesis aims to address these issues by developing and implementing (Paper I) new strategies and ad-hoc measures, including tools for evaluating the goodness-of-fit (Paper II), a new measure for quantifying the impact of heterogeneity (Paper III), a strategy to deal with differences in the exposure range across studies (Paper IV), and a one-stage approach to estimate complex models without excluding relevant studies (Paper V).

In Paper I, we described the implementation of the main aspects of the methodology in the `dosresmeta` R package available on CRAN. Dedicated functions was written to facilitate specific tasks such as definition of the design matrix and prediction of the pooled results. We illustrated how to estimate both linear and non-linear curves, conduct test of hypotheses, and present the results in a grapichal way using summarized data on alcohol intake and colorectal cancer risk.

In Paper II, we discussed how to evaluate the goodness-of-fit for dose-response meta-analysis. The proposed solutions consist of descriptive measures to summarize the agreement between fitted and observed data (the deviance and the coefficient of determination), and graphical tools to visualize the fit of the model (decorrelated residuals-versus-exposure plot). Data from two published meta-analyses were used to show how these tools can improve the practice of quantitative synthesis of aggregated dose-response data.

In Paper III, we proposed and discussed a new measure, R_b, to quantify the proportion of the variance of the pooled estimate attributable to the between-study heterogeneity. R_b does not make any assumption about the distribution of the within-study error variances, nor does it require specification of a typical value for these quantities. The performance of the proposed measure was evaluated in an extensive simulation study. We demonstrated how to present and interpret the R_b re-analyzing three published meta-analyses.

In Paper IV, we extended a point-wise approach to dose–response meta-analysis of aggregated data. The strategy consists of combining predicted relative risks for a fine grid of exposure values based on potentially different dose-response models. Thus, a point-wise approach can improve the flexibility in modelling the study-specific curves and may limit the impact of extrapolation by predicting the study-specific relative risk based on the observe exposure range. We illustrated the methodology using both individual and aggregated participant data.

In Paper V, we formalized a one-stage approach for dose–response meta-analysis in terms of a linear mixed model. We explained the main aspects of the methodology and how to extend the measures typically presented in a two-stage analysis. Using both hypothetical and real data, we showed how the one-stage approach can facilitate investigation the impact of heterogeneity over the exposure range, model comparison, and prediction of individual dose-response associations. The main advantage is that complex curves can be estimated without excluding relevant studies.

In conclusion, the methods presented in this thesis enrich the set of tools available to apply dose–response meta-analyses and to address specific questions including goodness-of-fit evaluation (Paper II) and a new measure of heterogeneity (Paper III). In addition, we presented alternative models for pooling results in case of heterogeneous exposure range (Paper IV) and for estimating complex models without excluding relevant studies (Paper V). The proposed methods have been illustrated using real data and implemented in the `dosresmeta` and `hetmeta` R packages available on CRAN (Paper I).

List of publications

I. Alessio Crippa, and Nicola Orsini
Multivariate dose–response meta-analysis: the dosresmeta R Package
Journal of Statistical Software, Code Snippets 2016; 72(1), 1–15

II. Andrea Discacciati, Alessio Crippa, and Nicola Orsini
Goodness of fit tools for dose–response meta-analysis of binary outcomes
Research Synthesis Methods 2015

III. Alessio Crippa, Polyna Khudyakov, Molin Wang, Nicola Orsini, and Donna Spiegelman
A new measure of between-studies heterogeneity in meta-analysis
Statistics in medicine 2016; 35(21), 3661–75

IV. Alessio Crippa, Ilias Thomas, and Nicola Orsini
A pointwise approach to dose-response meta-analysis of aggregated data
Manuscript 2018

V. Alessio Crippa, Andrea Discacciati, Matteo Bottai, Alicja Wolk, and Nicola Orsini
One-stage dose–response meta-analysis for aggregated data
Manuscript 2018

The articles will be referred to in the text by their Roman numerals, and are reproduced in full at the end of the thesis.

Related publications

- Alessio Crippa, Susanna C. Larsson, Andrea Discacciati, Alicja Wolk, and Nicola Orsini
  Red and processed meat consumption and risk of bladder cancer: a dose–response meta-analysis of epidemiological studies
  European journal of nutrition 2016, 1–13

- Andrea D. Smith, Alessio Crippa, James Woodcock, and Søren Brage
  Physical activity and incident type 2 diabetes mellitus: a systematic review and dose–response meta-analysis of prospective cohort studies
  Diabetologia 2016, 1–19

- Marco Vinceti, Tommaso Filippini, Alessio Crippa, Agnès de Sesmaisons, Lauren A. Wise, and Nicola Orsini
  Meta-Analysis of Potassium Intake and the Risk of Stroke
  Journal of the American Heart Association 2016, 5(10), e004210

- Alessio Crippa, and Nicola Orsini
  Dose–response meta-analysis of differences in means
  BMC medical research methodology 2016, 16(1), 91

- Emir Veledar, Alessio Crippa, Chukwuemeka U. Osondu, Adnan Younus, and Khurram Nasir
  Letter to Editor: Ideal cardiovascular health metrics and risk of cardiovascular disease or mortality
  International journal of cardiology 2016, 222, 737

- Alessio Crippa, Andrea Discacciati, Nicola Orsini, and Viktor Oskarsson
  Letter: coffee consumption and gallstone disease—a cautionary note on the assignment of exposure values in dose–response meta-analyses
  Alimentary Pharmacology & Therapeutics 2016, 43(1), 166-167

- Susanna C. Larsson, Alessio Crippa, Nicola Orsini, Alicja Wolk, and Karl Michaëlsson
  Milk consumption and mortality from all causes, cardiovascular disease, and cancer: a systematic review and meta-analysis
  Nutrients 2016, 7(9), 7749-7763

- Daniela Di Giuseppe, Alessio Crippa, Nicola Orsini, and Alicja Wolk
  Fish consumption and risk of rheumatoid arthritis: a dose-response meta-analysis
  Arthritis research & therapy 2014, 16(5), 446

- Alessio Crippa, Andrea Discacciati, Susanna C. Larsson, Alicja Wolk, and Nicola Orsini
  Coffee consumption and mortality from all causes, cardiovascular disease, and cancer: a dose–response meta-analysis
  American journal of epidemiology 2014, 180(8), 763-775

Contents

List of abbreviations

AIC	Akaike Information Criterion
CI	Confidence Interval
df	Degrees of Freedom
GLS	Generalized Least Squares
GRSS	Generalized Residual Sum of Squares
GTSS	Generalized Total Sum of Squares
FP2	Second-degree Fractional Polynomials
HRR	Hazard Rate Ratio
IR	Incidence Rate
IRR	Incidence Rate Ratio
logRR	log–Relative Risk
MR	Mortality Rate
MRR	Mortality Rate Ratio
RCS	Restricted Cubic Splines
R^2	Coefficient of Determination
RR	Relative Risk
WLS	Weighted Least Squares

Chapter 1

Introduction

A single experiment or study can hardly provide a definitive answer to a specific scientific question. Science is oftentimes referred to as a cumulative process where results from many studies, aiming to address the common question of interest, contribute to create and update the scientific evidence. In the cumulative paradigm, meta-analysis is the statistical methodology to combine and compare the current evidence in the field. This process lies at the heart of evidence-based medicine, and plays a major role in informing policy and practice.

Many epidemiological studies assess whether the occurrance of a health outcome (e.g. mortality, incidence of a disease) varies according to a quantitative exposure (e.g. amount of physical activity, alcohol intake). The quantitative exposure is frequently divided in intervals and the results are expressed in a tabular format as relative risks for different exposure categories. A high vs. low meta-analysis contrasts the relative risks for the highest exposure versus the lower one. This approach, however, discards the results for intermediate categories and thus provides only a limited picture. The information of the quantitative exposure is also lost, and the estimates being combined may be associated to different exposure values.

A dose–response meta-analysis, instead, has the potential to be more informative and powerful since it uses the whole available information to estimate the dose–response association. Because the estimates depend on the same reference category, it is not possible to regress the relative risks on the assigned dose using oridinaly least square. Greenland and Longnecker described in their seminal paper in 1992 how to reconstruct the correlation within set of relative risks and incorporate it in the dose–response analysis using generalized least square regression. Since then, the number of published dose–response meta-analysis has rapidly increased in many fields of application including oncology, public health, environmental sciences, nutrition, endocrinology, and internal medicine. Additional papers refined selected aspects of the proposed methodology, mainly focusing on the implementation of flexible strategies to model the dose-response curve and incorporating the advances of multivariate meta-analysis. However, there were still several relevant questions that needed to be addressed, including how to assess the goodness-of-fit, how to quantify the impact of heterogeneity, how to deal with differences in the exposure range across studies, and how to estimate complex models without excluding relevant studies.

This thesis aims to address these issues by developing and implementing (Paper I) new strategies and ad-hoc measures, including tools to evaluate the goodness-of-fit (Paper II), a new measure to quantify the impact of heterogeneity (Paper III), a strategy to deal with differences in the exposure range across studies (Paper IV), and a one-stage approach to estimate complex models without excluding relevant studies (Paper V). This thesis aims to address these issues by developing new strategies and ad-hoc measures. The proposed methodologies are demonstrated reanalyzing published meta-analyses and are implemented in user friendly packages written in the free and open source R language, to bridge the gap between theory and application.

Chapter 2

Background

Write my background with subsections.

Here an example of a figure (Figure 2.1).

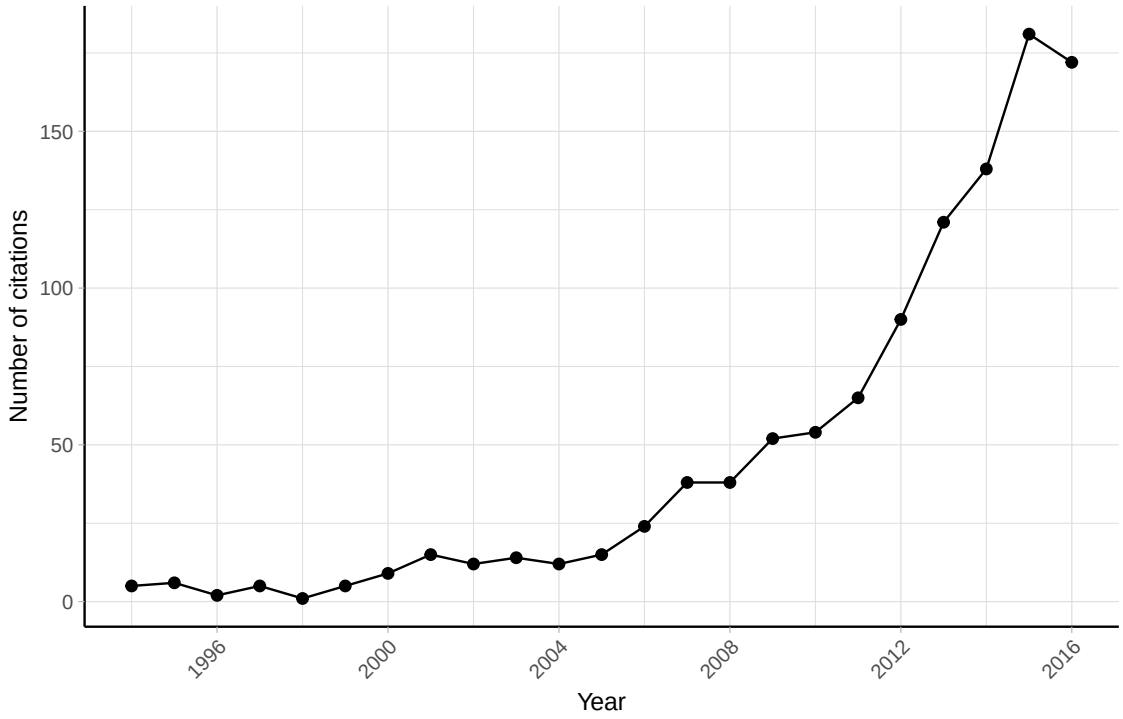

Figure 2.1

Chapter 3

Aims of the thesis

The overall aims of this thesis were to develop and implement new methods for dose–response in meta-analysis, in order to deal with the methodological aspects that have not yet been addressed.

More specifically, the aims were:

- To describe the implementation of the main aspects of a dose–response meta-analysis and the usage of the `dosresmeta` R package.
- To present and discuss relevant measures and graphical tools to assess the goodness-of-fit in dose–response meta-analysis of binary outcome.
- To develop a new measure of between-study heterogeneity in the broader context of meta-analysis, and assess its statistical properties as compared to other available measures.
- To explore possible advantages of a point-wise approach, especially, in case of dose–response meta-analysis where the exposure range varies substantially across the studies.
- To formalize and present an alternative one-stage random-effects model for dose-response meta-analysis of aggregated data, formulating the meta-analytic model in terms of a general linear mixed-effect model.

Chapter 4

Materials and methods

Write materials and methods with subsections as in the background section

Chapter 5

Results

Write the results with subsections as in the background section

Chapter 6

Discussion

Write the discussion with subsections as in the background section

Chapter 7

Conclusions

The methods presented in this thesis enrich the set of tools available to apply dose–response meta-analyses and to address specific questions including how to evaluate the goodness-of-fit and how to measure the impact of the between-studies heterogeneity. Furthermore, this thesis presents alternative models for pooling results in case of heterogeneous exposure range and for estimating complex models without excluding relevant studies. The proposed methods have been illustrated using real data from published meta-analyses and implemented in the `dosresmeta` and `hetmeta` R packages available on CRAN.

More specifically we conclude the following:

- The `dosresmeta` R package can help practitioners to conduct dose–response meta-analyses and to apply the methods presented in this thesis. Dedicated functions help to avoid common pitfalls frequently encountered in published meta-analyses, such as definition of the design matrix and prediction of the pooled results (Paper I).

- Goodness-of-fit should be regularly evaluated in applied dose-response meta-analysis. The proposed solutions consist of descriptive measures to summarize the agreement between fitted and observed data (the deviance and the coefficient of determination), and graphical tools to visualize the fit of the model (decorrelated residuals-versus-exposure plot). These tools can be employed to identify systematic dose–response patterns and possible source of heterogeneity, and to support the conclusions in applied meta-analyses (Paper II).

- The new measure of heterogeneity, R_b, quantifies the proportion of the variance of the pooled estimate attributable to the between-study heterogeneity. It does not make any assumption about the distribution of the within-study error variances, nor does it require specification of a typical value for these quantities. Therefore, we recommend the use of the R_b as preferred measure for quantifying the impact of heterogeneity (Paper III).

- A point-wise strategy for dose-response meta-analysis does not require the specification of a unique model as in the traditional approaches, and thus allows for more flexibility in modeling the individual curves. In addition, the extent of extrapolation is limited by

predicting the study-specific relative risk based on the observe exposure range. The use of the described strategy may improve the robustness of the results, especially in case of heterogeneous exposure range (Paper IV).

- A one-stage approach for dose–response meta-analysis consists of a linear mixed-effects model, which offer useful tools for describing the impact of heterogeneity over the exposure range, for comparing the fit of different models, and for predicting individual dose-response associations. The main advantage as compared to a two-stage analysis is that complex curves can be estimated without excluding relevant studies (Paper V).

Chapter 8

Future research

Based on the conclusions presented in this thesis, future research includes:

- <>
- <>
- <>

Appendix A

Supplementary figures

Figures.

Appendix B

Supplementary tables

Tables.

References

Crippa A, Discacciati A, Bottai M, Spiegelman D, Orsini N (2018a). "One-stage dose–response meta-analysis for aggregated data." *Manuscript*.

Crippa A, Khudyakov P, Wang M, Orsini N, Spiegelman D (2016). "A new measure of between-studies heterogeneity in meta-analysis." *Statistics in medicine*, **35**(21), 3661–3675.

Crippa A, Orsini N (2016). "Multivariate dose-response meta-analysis: The dosresmeta R Package." *Journal of statistical software, Code Snippets*, **72**(1), 1–15. doi:10.18637/jss.v072.c01.

Crippa A, Thomas I, Orsini N (2018b). "A pointwise approach to dose-response meta-analysis of aggregated data." *Submitted*.

Discacciati A, Crippa A, Orsini N (2015). "Goodness of fit tools for dose–response meta-analysis of binary outcomes." *Research synthesis methods*.

Acknowledgements

There are many people that I would like to thank for their contributions to this thesis, and for their support and encouragement during these years.

Nicola Orsini, my main supervisor for the second half of my doctoral education.

This work was supported by **Karolinska Institutet**'s funding for doctoral students (KID-funding).