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 procedure for combining and contrasting 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 implementation of flexible strategies and 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 new strategies and ad-hoc measures (Paper I), 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 tabular and grapichal format using summarized data on alcohol intake and colorectal cancer risk.

In Paper II, we discussed how to evaluate the goodness-of-fit. 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). A reanalysis of two published meta-analyses exemplified how these tools can improve the practice of quantitative synthesis of aggregated dose–response data.

In Paper III, we proposed and characterized a new measure, R_b, to quantify the proportion of the variance of the pooled estimate attributable to the between-study heterogeneity. Contrary to the available measures of 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. A point-wise approach can improve the flexibility in modeling 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 flexible curves can be estimated regardless of the number of data-points in the individual analyses.

In conclusion, the methods presented in this thesis enrich the set of tools available for applying dose–response meta-analyses and for addressing specific questions including goodness-of-fit evaluation (Paper II) and quantification 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 can hardly provide a definitive answer to a 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 the concept of evidence-based medicine, and plays a major role in informing policy and practice.

Epidemiological studies often 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 groups. A high verusu low meta-analysis contrasts the relative risks for the highest exposure category compared to 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 compared 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 group, 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 in modeling non-linear associations and incorporating the advances of multivariate meta-analysis. However, there were still several relevant questions that needed to be addressed such as how to assess the goodness-of-fit (Paper II), how to quantify the impact of heterogeneity (Paper III, how to deal with differences in the exposure range across studies (Paper IV), and how to estimate complex models without excluding relevant studies (Paper V).

This thesis aims to address these issues by developing and implementing new strategies and ad-hoc measures (Paper I). 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

2.1 Meta-analysis

Relevant research questions are typically addressed by independent investigators in multiple studies. The sampling error and possibly differences in the investigations will inevitably produce diverse results, sometimes even conflicting. Evidence-based medicine requires a synthesis of the available evidence to optimize the decision-making process (Haidich, 2010).

Meta-analysis, or more generally quantitative review synthesis, is the statistical methodology for integrating and synthetizing the information arising from multiple studies (Borenstein *et al.*, 2009). Using appropriate statistical models, quantitative reviews contrast and pool results in the hope of identifying similarities or explain differences across study findings. Meta-analysis represents the state of the art for systematically reviewing the evidence, as indicated by the increasing number of published meta-analyses over the last 40 years (Figure 2.1).

The classical approach for meta-analysis consists of a weighted average of the study-specific results or estimates. A fixed-effect model for meta-analysis assumes that all the studies estimate a single common parameter (Rice *et al.*, 2017). The hypothesis of homogeneity of the estimates is rarely applicable in biomedical and social sciences where studies typically differ in terms of design, disease classification, exposure measurement, and implemented statistical analyses (Colditz *et al.*, 1995). In such cases, heterogeneity across estimates is expected and should be considered in the analysis (Higgins *et al.*, 2009). If the parameters estimated in the studies are not identical but similar, a random-effects models can be used to identify those similarities or to explain the observed heterogeneity (Higgins *et al.*, 2009).

2.1.1 Random-effects meta-analysis

In a meta-analysis of K studies indexed by $i = 1, \ldots, K$, we denote $\hat{\beta}_i$ the estimate of an effect of interest (effect size) in the i-th study.

A random-effects model for meta-analysis can be written as

$$\hat{\beta}_i = \beta + u_i + \varepsilon_i \tag{2.1}$$

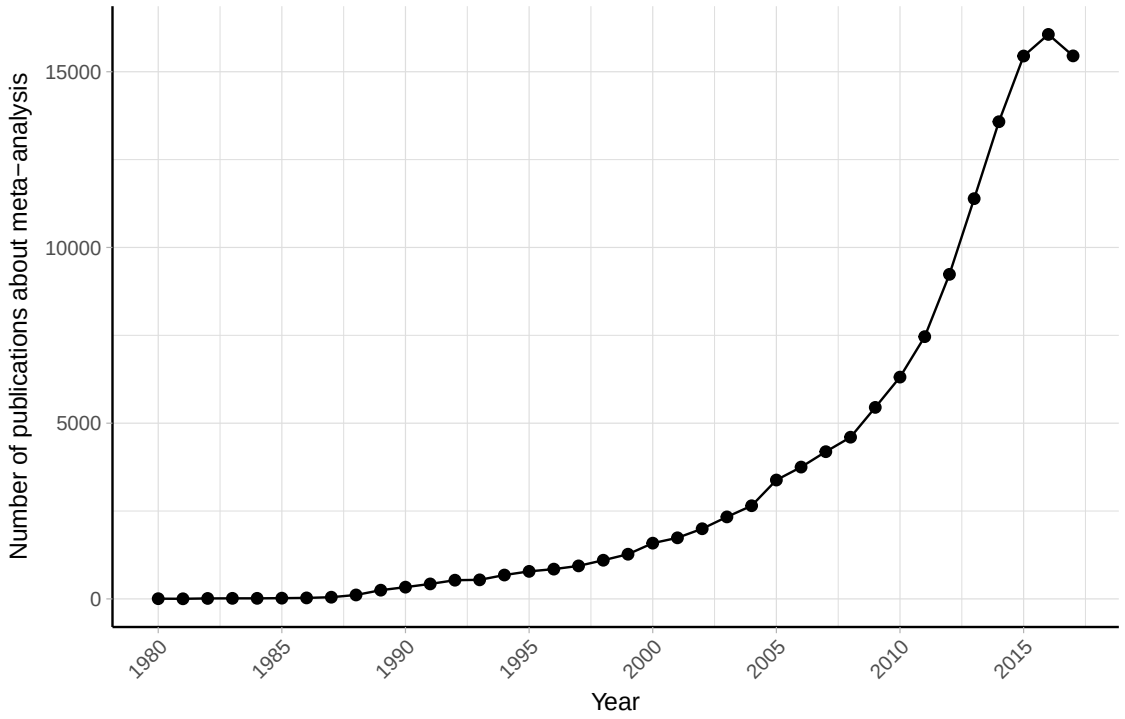

Figure 2.1: Number of publications about meta-analysis (results from Medline search using text "meta-analysis" until December 2017).

where β is the underlying mean effect, oftentimes the main parameter of interest. The random-effects u_i represent the study-specific deviations from the mean effect β allowing each study to estimate a similar parameter β_i defined as $\beta + u_i$. The random component follows a generic f distribution with mean 0 and variance equal to τ^2, the between-studies heterogeneity. The within-study error components ε_i have also mean 0 and variance equal to $\hat{v}_i$, an estimate of the sampling variance of $\hat{\beta}_i$. Because the sample size in the individual investigations is often large, the uncertainty around the estimates of the sampling variance is negligible. Therefore, $\hat{v}_i$ can be considered fixed and denoted as v_i. In addition, for the central limit theorem, $\varepsilon_i \sim \mathcal{N}(0, v_i)$, or alternatively, $\hat{\beta}_i | u_i \sim \mathcal{N}(\beta + u_i, v_i)$.

An inverse variance-weighted approach for meta-analysis estimates the mean effect β as a weighted average of the $\hat{\beta}_i$ (Whitehead and Whitehead, 1991; DerSimonian and Laird, 1986)

$$\hat{\beta} = \frac{\sum_{i=1}^{k} \hat{\beta}_i \hat{w}_i}{\sum_{i=1}^{k} \hat{w}_i} \tag{2.2}$$

$$\widehat{\mathrm{Var}}\left(\hat{\beta}\right) = \left(\sum_{i=1}^{k} \hat{w}_i\right)^{-1} \tag{2.3}$$

with weights $\hat{w}_i = \left(v_i + \hat{\tau}^2\right)^{-1}$ and $\hat{\tau}^2$ being an estimate of the between-study heterogeneity.

2.1.2 Test and estimates of heterogeneity

A second parameter of interest, often overlooked, is the between-study heterogeneity, τ^2. Focusing on the mean effect alone may provide only a limited piece of information, especially in case of heterogeneous effects (Borenstein *et al.*, 2010). Indeed, an evaluation of the extent of heterogeneity is a crucial step in determining the appropriateness of presenting a summary measure of the observed effect sizes.

Presence of heterogeneity is frequently defined as the excess in the variability of $\hat{\beta}_i$ above that expected alone by chance. A summary measure of the observed variability is represented by the Q statistic

$$Q = \sum_{i=1}^{K} \left(\hat{\beta}_i - \hat{\beta}_{\text{fe}}\right)^2 \tag{2.4}$$

where $\hat{\beta}_{\text{fe}} = \sum_{i=1}^{K} \hat{\beta}_i v_i^{-2} / \sum_{i=1}^{K} v_i^{-2}$ is the estimate of β in a fixed-effect model. Based on that, Cochrane developed a test for assessing hypothesis of homogeneity of the study-specific estimates (Cochran, 1954). Under the null hypothesis of no heterogeneity ($H_0 : \tau^2 = 0$) the Q statistic follows a χ^2 with $K-1$ degrees of freedom. A p value less than 0.10 is oftentimes used as evidence for presence of between-studies variability. It is known, however, that the test is sensible to the number of studies K failing to reject the null hypothesis even for high value of τ^2 when K is small, and contrary it is more likely to reject H_0 for negligible between-studies variation when K is big (Higgins and Thompson, 2002; Takkouche *et al.*, 1999). Therefore, failing to reject the null hypothesis does not provide evidence supporting homogeneity in the effect sizes (Biggerstaff and Tweedie, 1997). In addition, the dichotomization heterogeneous/homogeneous is not very informative, especially because heterogeneity is almost always present.

An estimate of τ^2, instead, directly provides information about the amount of heterogeneity and is thus the more natural measure of between-studies variability. Based on the expectation of Q, Dersimonian and Laird proposed the following estimator for τ^2 using the method of moments (DerSimonian and Laird, 1986)

$$\hat{\tau}^2_{\text{DL}} = \max\left\{0, \frac{Q-(K-1)}{\sum_{i=1}^{K} v_i^{-2} - \sum_{i=1}^{K} v_i^{-4} / \sum_{i=1}^{K} v_i^{-2}}\right\} \tag{2.5}$$

The moment-based estimator is one of the most popular estimators of τ^2 because it has a simple non-iterative formulation and does not require any distributional assumptions for the random-effects rather than having a finite first order moment. Other common non-iterative alternatives include an estimator based on the variance components (Hedges, 1983) and using methods for estimating the error variance in a weighted linear model (Sidik and Jonkman, 2005). Iterative methods based on maximizing the likelihood or restricted likelihood can also be used by specifying a distributional form for the random-effects. The more conventional choice is typically a normal distribution $u_i \sim \mathcal{N}\left(0, \tau^2\right)$, which implies $\beta_i \sim \mathcal{N}\left(\beta, \tau^2\right)$ and $\hat{\beta}_i \sim \mathcal{N}\left(\beta + u_i, \tau^2 + v_i\right)$.

Although τ^2 is the more natural and appropriate measure of between-study variability, the actual value is difficult to interpret because it depends on type of effect size (e.g. log relative risk, standardized mean difference) and has no upper limit. Therefore, both evaluation of the degree (or levels) and the comparison of heterogeneity in different meta-analyses can hardly be based on the estimate of τ^2.

2.1.3 Measures of heterogeneity

To complement the test based approach and the information provided by $\hat{t}au^2$, measures that quantify the impact of heterogeneity have been proposed (Higgins and Thompson, 2002). Higgins et al. presented several possibilities in the simpler case where all the sampling variances v_i are equal to a fixed and known value σ^2.

Two measures aim to estimate the ratio $\sigma^2/(\sigma^2/+\tau^2)$, namely the $H^2 = Q/(K-1)$ that represents the excess in Q statistic relative to its degrees of freedom, and $R^2 = \mathrm{Var}\,\hat{\beta}/\mathrm{Var}\,\hat{\beta}_{\mathrm{FE}}$ which describe the inflation in the variability of the mean effect in a random-effects model compared with a fixed-effect model. Other measures, instead, relate the between-studies heterogeneity, τ^2, to the marginal or unconditional variability τ^2+v_i, which is defined by the sum of within- and between-study components. These measures can be more easily interpreted as the percentage of the total variability due to heterogeneity, similar to the intraclass correlation coefficient defined for linear mixed-effects models. The ratio directly involves the within-terms v_i that again varies across the studies. Indeed, the most popular measures, namely the R_I (Takkouche *et al.*, 1999) and the I^2 (Higgins and Thompson, 2002), replaced v_i with a statistic that summarizes the observed distribution of v_i. Takkouche et al. chose

$$s_1^2 = \frac{K}{\sum_{i=1}^{K} v_i^{-2}} \tag{2.6}$$

that is the harmonic mean of the inverse of the sampling variances. Higgins et al., instead, described the "typical" within-study variance as

$$s_2^2 = \frac{(K-1)\sum_{i=1}^{K} v_i^{-2}}{\left(\sum_{i=1}^{K} v_i^{-2}\right)^2 - \sum_{i=1}^{K} v_i^{-4}} \tag{2.7}$$

that provided a direct relationship with the Q statistic: $I^2 = (Q-(K-1))/Q$ when τ^2 is estimated using the method of moments. Both statistics can be expressed as a percentage where 0% corresponds to no heterogeneity and increasing values imply higher levels of heterogeneity. It is known that these measures depend on precision of the study-specific estimates and tend to increase to 100% when the v_i are much smaller than the estimated τ^2. A complementary measure is the between-studies coefficient of variation, defined as $\tau^2/|\hat{\beta}|$, that does not directly depend on the within-study variances. However, it increases quickly as $\hat{\beta}$ becomes smaller, and is not defined for $\hat{\beta} = 0$.

While the limitations of the Q test approach are widely known, little emphasis is placed

on the assumptions underneath the definition of the established measures of heterogeneity, i.e. all the estimates being reported with the same precision, which is unlikely to be met in almost all the applications. A measure of the impact of heterogeneity that does not require such an assumption would be desirable.

2.2 Categorical models for dose–response analysis

2.2.1 Categorical analysis with individual participant data

2.2.2 Aggregated dose–response data

2.2.3 High versus low meta-analysis

2.3 Dose–response meta-analysis

2.3.1 First stage: study-specific curves

2.3.2 Second stage: multivariate meta-analysis

2.4 Software

Here an example of a figure (Figure 2.2).

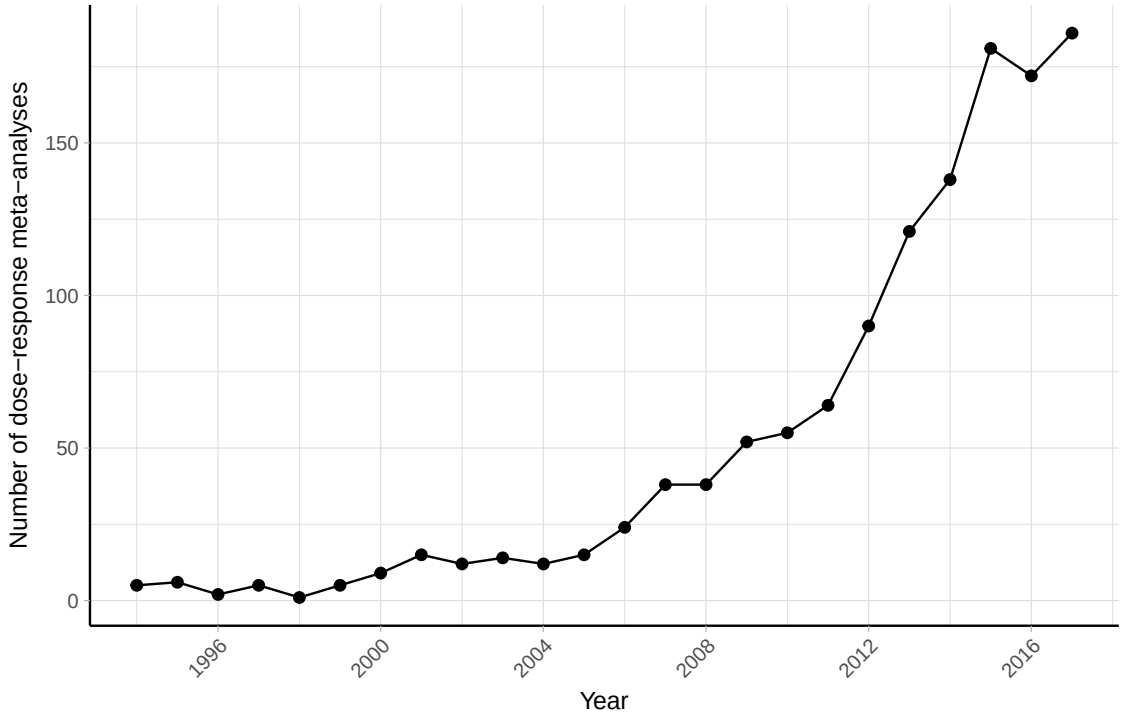

Figure 2.2: Number of citations of the paper by Greenland and Longnecker (1992) obtained from Google Scholar 1992-2017 (until December 2017).

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.
- To develop a new measure of between-study heterogeneity in the broader context of meta-analysis and assess its 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

Methods

4.1 The `dosresmeta` R package

4.1.1 Implementation

4.1.2 Exposure modeling

4.2 A new measure of heterogeneity

4.2.1 Statistical proporties

4.2.2 Simulation study

4.3 Goodness-of-fit

4.3.1 Deviance

4.3.2 Coefficient of determination

4.3.3 Visual tools

4.4 A point-wise approach

4.4.1 Prediction of study-specific (log) relative risks

4.4.2 Pool of dose–response prediction

4.5 A one-stage model

4.5.1 Model definition

4.5.2 Estimation and hypothesis testing

4.5.3 Prediction

4.5.4 Comparison with two-stage analysis

Chapter 5

Results

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 for applying dose–response meta-analyses and for addressing specific questions including how to evaluate the goodness-of-fit and how to measure the impact of the between-studies heterogeneity. Furthermore, this thesis describes 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 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. Contrary to the available measures of 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 therefore 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. Thhe main advantage is that flexible curves can be estimated regardless of the number of data-points in the individual analyses (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

Biggerstaff B, Tweedie R (1997). "Incorporating variability in estimates of heterogeneity in the random effects model in meta-analysis." *Statistics in medicine*, **16**(7), 753–768.

Borenstein M, Hedges LV, Higgins J, Rothstein HR (2009). *References*. Wiley Online Library.

Borenstein M, Hedges LV, Higgins J, Rothstein HR (2010). "A basic introduction to fixed-effect and random-effects models for meta-analysis." *Research synthesis methods*, **1**(2), 97–111.

Cochran WG (1954). "The combination of estimates from different experiments." *Biometrics*, **10**(1), 101–129.

Colditz GA, Burdick E, Mosteller F (1995). "Heterogeneity in meta-analysis of data from epidemiologic studies: a commentary." *American journal of epidemiology*, **142**(4), 371–382.

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*.

DerSimonian R, Laird N (1986). "Meta-analysis in clinical trials." *Controlled clinical trials*, **7**(3), 177–188.

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

Haidich AB (2010). "Meta-analysis in medical research." *Hippokratia*, **14**(Suppl 1), 29.

Hedges LV (1983). "A random effects model for effect sizes." *Psychological Bulletin*, **93**(2), 388.

Higgins J, Thompson SG (2002). "Quantifying heterogeneity in a meta-analysis." *Statistics in medicine*, **21**(11), 1539–1558.

Higgins J, Thompson SG, Spiegelhalter DJ (2009). "A re-evaluation of random-effects meta-analysis." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, **172**(1), 137–159.

Rice K, Higgins J, Lumley T (2017). "A re-evaluation of fixed effect (s) meta-analysis." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Sidik K, Jonkman JN (2005). "Simple heterogeneity variance estimation for meta-analysis." *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **54**(2), 367–384.

Takkouche B, Cadarso-Suárez C, Spiegelman D (1999). "Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis." *American journal of epidemiology*, **150**(2), 206–215.

Whitehead A, Whitehead J (1991). "A general parametric approach to the meta-analysis of randomized clinical trials." *Statistics in medicine*, **10**(11), 1665–1677.

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).