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

- Ehimen C. Aneni, Alessio Crippa, Chukwuemeka U. Osondu, Javier Valero-Elizondo, Adnan Younus, Khurram Nasir, and Emir Veledar
  Estimates of Mortality Benefit From Ideal Cardiovascular Health Metrics: A Dose Response Meta-Analysis
  Journal of the American Heart Association 2017 Dec 1;6(12):e006904.

- 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

- 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
CS	Cubic Splines
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
logRR	log–Relative Risk
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, 2008). 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 I studies indexed by $i = 1, \ldots, I$, 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

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

where β is the underlying mean effect, oftentimes the main parameter of interest. The random-

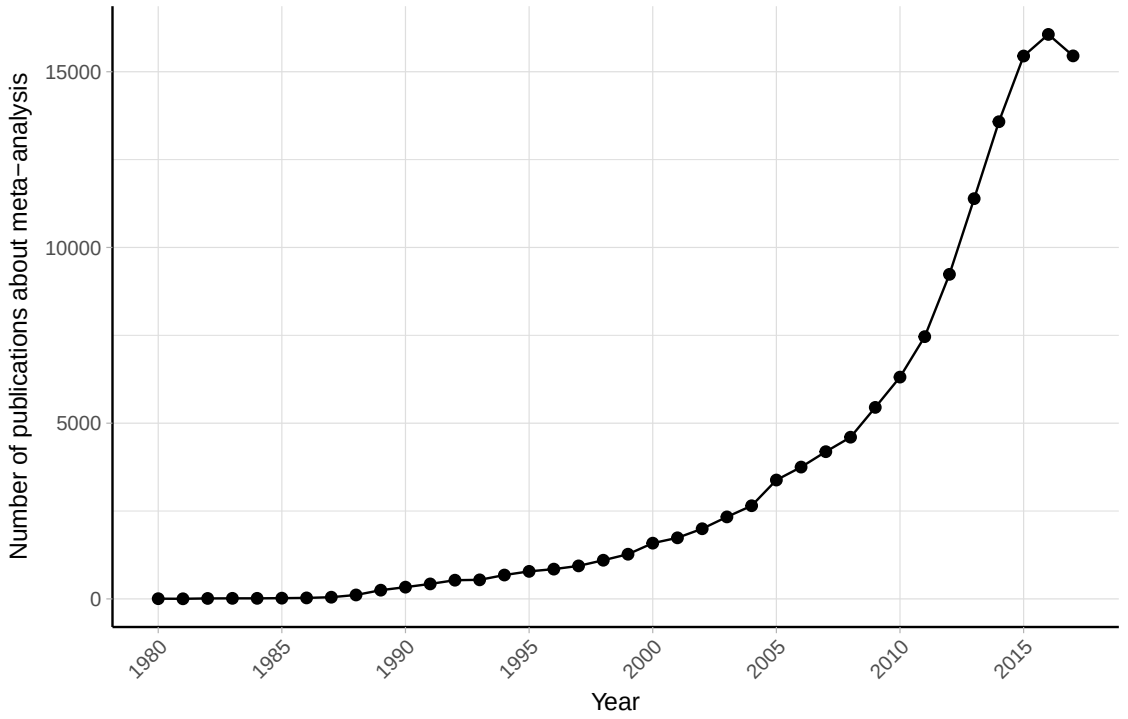

Figure 2.1: Number of publications about meta-analysis (results from Medline search using text "meta-analysis" until December 2017).

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 effects follow 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)

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

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

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

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

where $\hat{\beta}_{\text{fe}} = \sum_{i=1}^{I} \hat{\beta}_i v_i^{-2} / \sum_{i=1}^{I} v_i^{-2}$ is the estimate of β in a fixed-effect model. Based on this statistic, Cochrane developed a test for assessing the 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 distribution with $I - 1$ 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 I failing to reject the null hypothesis even for high value of τ^2 when K is small, and contrary, 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 (Higgins, 2008).

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)

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

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 assumption for the random-effects rather than having a finite first order moment. Other common non-iterative alternatives include estimators based on the variance components (Hedges, 1983) and on methods for estimating the error variance in weighted linear models (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 $\hat{\tau}^2$, 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 $H^2 = Q/(I-1)$ that represents the excess in Q statistic relative to its degrees of freedom, and $R^2 = \mathrm{Var}\left(\hat{\beta}\right) / \mathrm{Var}\left(\hat{\beta}_{\mathrm{FE}}\right)$ which describes 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 $\tau^2 + v_i$, 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

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

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

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

that provided a direct relationship with the Q statistic: $I^2 = (Q-(I-1))/Q$ 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 $\hat{\beta} = 0$.

2.2 Categorical models for dose–response analysis

Epidemiological studies often assess the strength and direction of the association between protective or risk factors (generally referred to as exposures) and the occurrence of a health outcome. When the exposure of interest is measured on a continuous scale, the additional information on the shape of the relationship is mostly of interest. Including the continuous variable simply as covariate in the appropriate statistical model assumes that the outcome linearly depends on the covariate. Associations between variables, however, are rarely linear. If the real dose-response is in fact non-linear, estimating a linear trend will have important consequences in detecting an association (Harrell Jr, 2015).

One common approach to relax the linearity assumption is to divide the quantitative exposure in categories. This categorical approach has been frequently criticized because of severe limitations (Royston *et al.*, 2006; Greenland, 1995) including loss of information and thus power, assuming an unrealistic step function, and subjective choice in selecting cut-points. Instead, many articles presented and illustrated alternative solutions such as the use of fractional polynomials and regression splines for easily modelling non-linear relationships.

Nevertheless, a recent survey among top medical and epidemiological journals estimated that categorization occurred 86% of the times (Turner *et al.*, 2010). One possible reason is that a categorical approach facilitates the interpretation of the estimated regression coefficients and simplifies the presentation of the results in a tabular format (Orsini *et al.*, 2011a).

2.2.1 Aggregated dose–response data

In a categorical approach the quantitative exposure is divided in $J+1$ categories. The corresponding indicator or dummy variables index by $j = 1, \ldots, J$ are included in the model in place of the exposure variable. The results from such a categorical dose–response analysis are expressed as relative measures using one category (corresponding to the omitted dummy variable) as referent. Depending on the study-design and on the statistical model, the results consist of estimated odds ratios, rate ratios, or risk ratios (generally referred to as relative risks (RRs)) for the different exposure categories, possibly adjusted for potential confounders. The corresponding 95% confidence intervals $\widehat{\text{RR}}_L, \widehat{\text{RR}}_U$ provide information on the uncertainty related to the estimated regression coefficients. Additional information about the assigned dose (mean or median within exposure intervals), the number of cases and the total number of subjects or person-time usually complements the reported results. The general structure and notation for aggregated or summarized dose–response data are presented for a generic i-th study in table 2.1. The i pedix in J_i highlights that independent studies may categorized the continuous exposure using different number of categories.

The statistical model relate the effect of the exposure categories on a transformation of the mean outcome. Typically, these transformations involve the natural logarithm such as the log odds, log risk, or log rate. The estimated regression coefficients are then exponentiated for ease of interpretation but the inference is actually performed on the modelling scale. Therefore, the

Table 2.1: Aggregated results from a categorical dose–response analysis.

Exposure level	Assigned dose	Cases	n[a]	$\widehat{\text{RR}}$	95% CI
0	x_{i0}	c_{i0}	n_{i0}	1	—
1	x_{i1}	c_{i1}	n_{i1}	$\widehat{\text{RR}}_{i1}$	$\widehat{\text{RR}}_{Li1}$, $\widehat{\text{RR}}_{Ui1}$
⋮	⋮	⋮	⋮	⋮	⋮
J_i	x_{iJ_i}	c_{iJ_i}	J_{iJ_i}	$\widehat{\text{RR}}_{iJ_i}$	$\widehat{\text{RR}}_{LiJ_i}$, $\widehat{\text{RR}}_{UiJ_i}$

[a] Depending on the study design, this column reports either total number of subjects or amount of person-time.

effect sizes considered in a meta-analysis of multiple aggregated dose–response data consist of the estimated log $\widehat{\text{RR}}$s and the corresponding standard errors that can be easily derived from the data available in table 2.1

$$\widehat{\text{SE}}\left(\log \widehat{\text{RR}}\right) = \frac{\log\left(\widehat{\text{RR}}_U\right) - \log\left(\widehat{\text{RR}}_L\right)}{2\, z_{1-\alpha/2}} \tag{2.8}$$

where $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of a standard normal distribution, usually approximated to 1.96 for the common $\alpha = 5\%$ level.

A distinctive feature of aggregated dose–response data is the correlation among the (log) $\widehat{\text{RR}}$s, which arises from the fact that they are estimated using a common reference group. Each $\widehat{\text{RR}}$ has the same baseline risk as denominator that works as comparator. If the observed baseline risk happens to be high or low just by chance, the estimated $\widehat{\text{RR}}$s will be also higher or lower than expected. This adds complications in guessing a trend from a categorical dose-response analysis or in directly comparing results based on different baseline categories.

2.2.2 High vs. low, categorical, and meta-regression models

A common approach for synthetizing the information from multiple aggregated dose-response data is to limit the analysis on a small portion of the available results. In particular, a high-versus-low meta-analysis focuses on the results for the highest exposure categories. By selecting only the last raw of the aggregated dose–response data, the meta-analytic models discussed in section 2.1.1 are used for combining and contrasting the results, with $\hat{\beta}_i = \widehat{\text{RR}}_{iJ_i}$. A major limitation of a high- versus-low approach is that both the highest and the lowest category may be associate to a different exposure value. To limit the impact of heterogeneous category definitions, practitioners should carefully plan the analysis by selecting the $\widehat{\text{RR}}$s for exposure categories whose definition is more consistent across studies. If also the choice of baseline category substantially differs, the $\widehat{\text{RR}}$s can be re-expressed using an alternative reference category implementing dedicated methodologies Hamling *et al.* (2008).

The major limitation, however, is that only a subset of the data is analyzed, while the remaining information about intermediate exposure categories is excluded from the analysis. As a consequence, much of the information about the shape of the dose-response is lost and the

power of detecting an association may dramatically decrease (e.g. in case of a U-shape relationship). A possible remedy, although less common, is to conduct a categorical meta-analysis, which consists of separate univariate meta-analyses pooling the results from comparable exposure categories. A dose-response association is then deducted from observing the combined $\widehat{\text{RR}}$s for increasing dose levels. A part from evident difficulties in identifying $\widehat{\text{RR}}$s for homogenous exposure intervals in applied works, this approach does not take into account the correlations across set of log $\widehat{\text{RR}}$s and suffers from the same problem of guessing a trend from a categorical dose–response analysis.

An additional alternative may be the use meta-regression models Berkey *et al.* (1995), where the dummy variables for the exposure categories or transformation of the dose are included as covariates in model 2.1. The rational would be to estimate a pooled RRs for different dose levels and to reduce the heterogeneity across study finding. Despite the quantitative exposure can be modelled using flexible tools, a meta-regression model treats the continuous predictor as a confounder. In addition, inference may be severely biased because the described approach fails to handle the hierarchical structure of the data, that is dose levels nested within studies.

2.3 Dose–response meta-analysis

The aim of a dose–response meta-analysis is to reconstruct the shape of the association from multiple aggregated dose–response data. As compared to the previous strategies, it has the advantages of using the whole information available and being more informative. By describing the variation of the outcome over the entire exposure range, a dose–response meta-analysis allows to answer the following questions

- Is there any association between increasing dose levels and the outcome? If that's the case, what is the shape of the relationship?
- Which exposure values are associated with the minimum or maximum outcome value?
- Is there any difference in the study-specific dose–response associations? Which factors can explain the observed heterogeneity?

The methodology for dose–response meta-analysis was first presented by Greenland and Longnecker in their seminal paper (Greenland and Longnecker, 1992), which quickly became a standard reference for applied works. Indeed, the number of published dose–response meta-analyses increased exponentially from 9 in 2000 to 172 in 2016 (figure 2.2). The most popular fields of application include oncology, environmental and public health, nutrition epidemiology, and general internal medicine. Dose-response meta-analyses are published in many leading medical and epidemiological journals, including JAMA, Lancet, Stroke, Gastroenterology, American J of Medicine, American J of Clinical Nutrition, American J Epidemiology, International J Epidemiology, Journal National Cancer Institute, International J of Cancer, Statistics in Medicine

and many others. The method is also used by the World Cancer Research Fund/American Institute for Cancer Research for reviewing the evidence on the relations between life-style factors (e.g. diet and physical activity) and cancer. Guidelines based on these quantitative reviews are central to promote the overall health and prevent many chronic diseases.

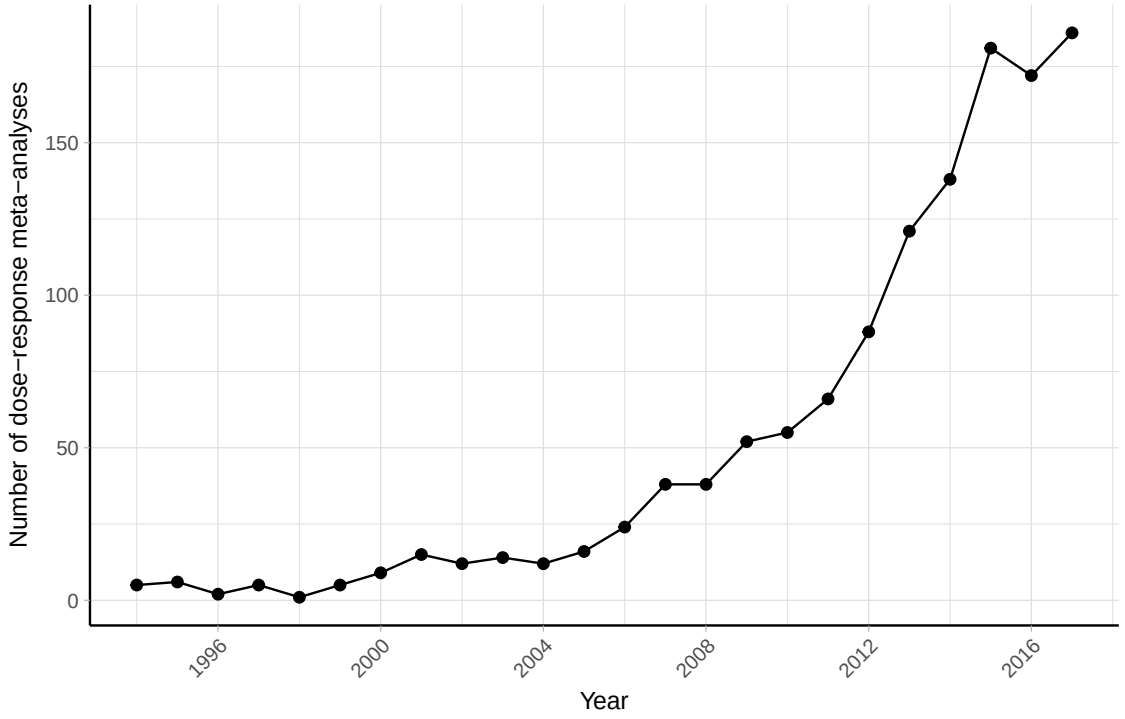

Figure 2.2: Number of citations of the paper by Greenland and Longnecker (1992) obtained from Google Scholar 1992-2017 (until December 2017).

The common approach for dose–response meta-analysis consists of a two-stage procedure, where the regression coefficients for the study-specific trends are first estimated separately within each study, and then combined using meta-analysis. In the next sections we cover the main methodological aspects related to each stage of the analysis.

2.3.1 First stage: study-specific trends

If we had access to the individual patient data, the dose-response model for a simple linear trend could be written as

$$\log(\lambda(x,\mathbf{z})) = \beta_0 + \beta_1 x + \boldsymbol{\gamma}^\top \mathbf{z} \tag{2.9}$$

with x the quantitative exposure and z the set of possible confounders. The outcome variable is the log of a transformation of the mean outcome (e.g. odds, risk, or rate). Transformations of the exposure variable can be included to relax the linearity assumption, such as a quadratic term

$$\log(\lambda(x,\mathbf{z})) = \beta_0 + \beta_1 x + \beta_2 x^2 + \boldsymbol{\gamma}^\top \mathbf{z} \tag{2.10}$$

However, we have rarely access to the individual patient data and our inference is limited to a summary of the initial data. In particular, aggregated data from a categorical analysis can be often retrieved from published articles. The aim of the first stage of a dose–response meta-analysis is to estimate the β coefficients in equation 2.9 and 2.10 using aggregated dose–response data. We consider the notation presented in table 2.1 with $i = 1, \ldots, I$ indexing the studies and $j = 1, \ldots, J_i$ the non-referent dose levels of a generic i-th study. The corresponding two models can be written as

$$\log\left(\widehat{\mathrm{RR}}_{ij}\right) = \log\left(\hat{\lambda}\left(x = x_{ij}\right)\right) - \log\left(\hat{\lambda}\left(x = x_{i0}\right)\right) = \beta_1\left(x_{ij} - x_{i0}\right) \tag{2.11}$$

$$\log\left(\widehat{\mathrm{RR}}_{ij}\right) = \log\left(\hat{\lambda}\left(x = x_{ij}\right)\right) - \log\left(\hat{\lambda}\left(x = x_{i0}\right)\right) = \beta_1\left(x_{ij} - x_{i0}\right) + \beta_2\left(x_{ij}^2 - x_{i0}^2\right) \tag{2.12}$$

More generally, the i-th dose-response model is defined as

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i \tag{2.13}$$

The outcome $\mathbf{y}_i$ is the J_i length vector of the non-referent log $\widehat{\mathrm{RR}}$s while $\mathbf{X}_i$ the $J_i \times p$ design matrix containing the p transformations of the assigned dose used to model the dose–response association

$$\mathbf{X}_i = \begin{bmatrix} g_1(x_{i1}) - g_1(x_{i0}) & \ldots & g_p(x_{ip}) - g_p(x_{i0}) \\ \vdots & & \vdots \\ g_1(x_{iJ_i}) - g_1(x_{i0}) & \ldots & g_p(x_{iJ_i}) - g_p(x_{i0}) \end{bmatrix} \tag{2.14}$$

In the linear trend analysis (model 2.11), $\mathbf{X}_i$ includes only the dose levels, $p = 1$, $g_1(x) = x$ (identity function)

$$\mathbf{X}_i = \begin{bmatrix} x_{i1} - x_{i0} \\ \vdots \\ x_{iJ_i} - x_{i0} \end{bmatrix}$$

while $p = 2$ columns are needed in the quadratic model 2.12: $g_1(x) = x$ and $g_2(x) = x^2$

$$\mathbf{X}_i = \begin{bmatrix} x_{i1} - x_{i0} & x_{i1}^2 - x_{i0}^2 \\ \vdots & \vdots \\ x_{iJ_i} - x_{i0} & x_{iJ_i}^2 - x_{i0}^2 \end{bmatrix}$$

A feature of the models 2.13 is the absence of the intercept term. The reference line in table 2.1 is not actually used for the estimation of the regression coefficients but introduces the constrain on the predicted log $\widehat{\mathrm{RR}}$, which needs to be 0 ($\widehat{\mathrm{RR}} = 1$) for the reference dose value x_{i0}, as explicit in models 2.11 and 2.12.

Approximation of the covariance between log $\widehat{\mathrm{RR}}$

A particular characteristic of summarized dose–response data is that the log $\widehat{\mathrm{RR}}$s are reported with different precision and are constructed using the same baseline group. Thus, the error

terms $\boldsymbol{\epsilon}_i$ in equation 2.13 are heterogeneous and correlated, with a covariance matrix structured as

$$\text{Cov}(\boldsymbol{\epsilon}_i) = \mathbf{S}_i = \begin{bmatrix} \sigma_{i11} & & & & \\ \vdots & \ddots & & & \\ \sigma_{i1j} & & \sigma_{ijj} & & \\ \vdots & & & \ddots & \\ \sigma_{i1J_i} & \dots & \sigma_{iJ_ij} & \dots & \sigma_{iJ_iJ_i} \end{bmatrix} \tag{2.15}$$

with the variance of the log $\widehat{\text{RR}}$s on the diagonal (σ_{ijj}) and the pairwise covariances as non-diagonal elements ($\sigma_{ijj'}$).

Two methods have been proposed to approximate the covariances $\sigma_{ijj'}$ (Greenland and Longnecker, 1992; Hamling *et al.*, 2008). Greenland and Longnecker described an algorithm to construct a table of pseudo or effective counts (number of cases and participants or person-time) that would produce the adjusted log $\widehat{\text{RR}}$s as those published. A unique solution for the algorithm is ensured by keeping the margins of the pseudo-counts equal to the observed ones. Alternatively, Hamling et al. modified the previous algorithm in such a way that the pseudo-counts would also match the standard errors of the log $\widehat{\text{RR}}$s.

Estimation

The dose-response coefficients $\boldsymbol{\beta}_i$ can be efficienly estimated using generalized least squares estimator (GLS), which minimizes the quadratic loss function $\left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i\right)^\top \mathbf{S}_i^{-1}\left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i\right)$ with respect to $\boldsymbol{\beta}_i$ assuming the covariance matrix $\mathbf{S}_i$ known.

$$\begin{aligned} \hat{\boldsymbol{\beta}}_i &= (\mathbf{X}_i^\top \mathbf{S}_i^{-1}\mathbf{X}_i)^{-1}\mathbf{X}_i^\top \mathbf{S}_i^{-1}\mathbf{y}_i \\ \widehat{\text{Var}}\left(\hat{\boldsymbol{\beta}}_i\right) &= (\mathbf{X}_i^\top \mathbf{S}_i^{-1}\mathbf{X}_i)^{-1} \end{aligned} \tag{2.16}$$

The GLS estimates in equation 2.16 do not require any distributional assumption for the error terms. However, for the central limit theory, the error terms follow approximately a normal distribution $\boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_i)$. Using this additional assumption, the log-likelihood of model 2.13 is

$$\ell\left(\boldsymbol{\beta}_i\right) = -\frac{J_i}{2}\log(2\pi) - \frac{1}{2}|\mathbf{S}_i| - \frac{1}{2}\left[\left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i\right)^\top (\mathbf{S}_i)^{-1}\left(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}_i\right)\right] \tag{2.17}$$

Interestingly, the maximum likelihood estimates that maximize the log-likelihood 2.17 coincides with the GLS estimates in 2.16 (Orsini *et al.*, 2006). Introducing the normality distribution for the random errors facilitates the inference, i.e. test of hypothesis and confidence intervals, on the $\boldsymbol{\beta}_i$ coefficients. The estimates in 2.16 are a linear combination of normal distributions ($\mathbf{y}_i \sim \mathcal{N}\left(\mathbf{X}_i\boldsymbol{\beta}_i, \mathbf{S}_i\right)$) and therefore are also normally distributed $\hat{\boldsymbol{\beta}}_i \sim \mathcal{N}\left(\boldsymbol{\beta}_i, \text{Var}\left(\hat{\boldsymbol{\beta}}_i\right)\right)$.

The ML and GLS estimators always give unbiased estimates of $\boldsymbol{\beta}_i$ regardless of the specification of $\mathbf{S}_i$ (Orsini *et al.*, 2006). As a consequence, also a weighted least square estimator (WLS) that assumes independence of the log $\widehat{\text{RR}}$s will produce unbiased estimates. However, taking into account the correlation will improve the statistical properties of the estimator, in particular the efficiency. We investigated the differences between the GLS and WLS estimators using a

simulation study of 5000 aggregated dose–response data where the true trends were linear (β_{TRUE} = -0.014). As expected, both the estimators were unbiased and consistent but the empirical distribution of the GLS estimator was more concentrated around the true β value 2.3. The WLS estimates of the standard errors of the $\hat{\beta}_i$ were lower than the corresponding GLS values. This had a direct effect on the inference for the estimated linear trend. For istance, it may be interesting to fit a quadratic curve as in 2.12 and test the hypothesis $H_0 : \beta_2 = 0$, i.e. departure from a linear trend. Using inference based on WLS estimators the null hypothesis were wrongly rejected 3.96% of the time, lower than the nominal level $\alpha = 5\%$. The corresponding number for the GLS estimator was instead closer (4.8%). We also implemented simulations assuming a quadratic curve with the true coefficients $\boldsymbol{\beta}_{\text{TRUE}}$ = (-0.092, 0.003). Similar results for the empirical bivariate distribution of $\hat{\boldsymbol{\beta}}_i$ and their standard errors are presented in figure 2.4.

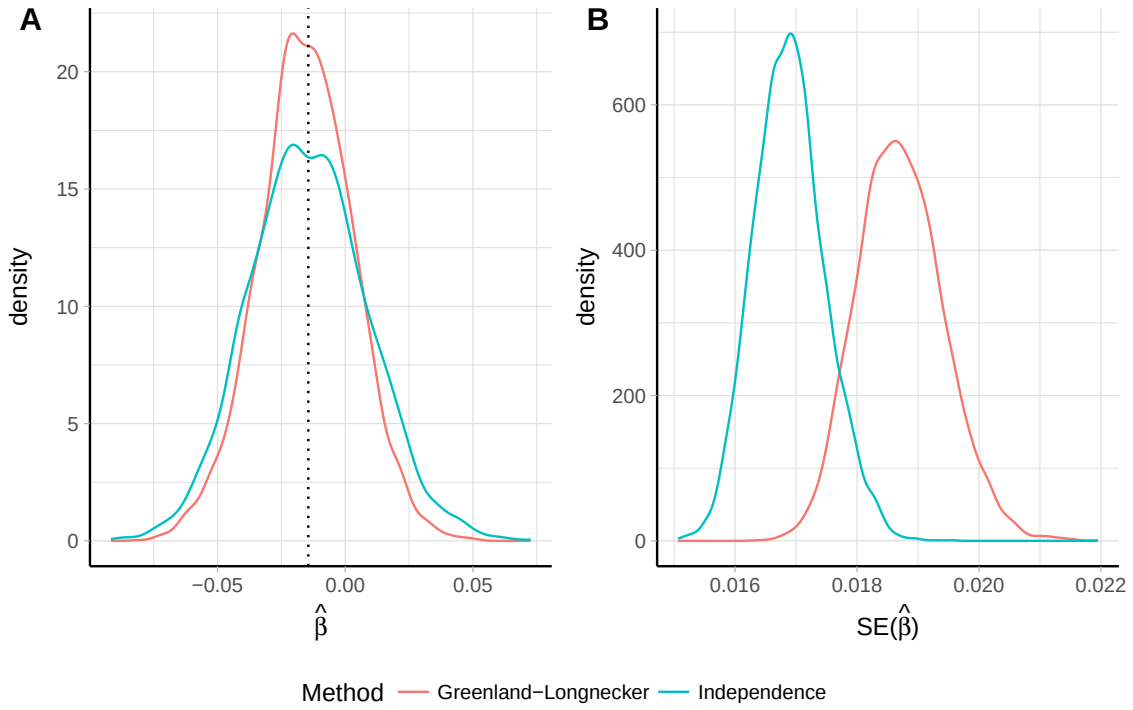

Figure 2.3: Empirical distribution of the $\hat{\beta}$ (panel A) and $\widehat{\text{Var}}\left(\hat{\beta}_i\right)$ (panel B) for a linear trend assuming independence of the log $\widehat{\text{RR}}$ and reconstructing the covariances using the Greenland and Longnecker's method. Results are based on simulations with 5000 replications and a true linear trend β =-0.014.

2.3.2 Second stage: multivariate meta-analysis

The study-specific dose-response curves are defined by the p transformations, $g_1(x), \ldots, g_p(x)$, and the estimated regression coefficients $\hat{\boldsymbol{\beta}}_i$. A pooled dose–response can be obtained by combining the $\hat{\boldsymbol{\beta}}_i$ coefficients. For that purpose, the same functional relationship needs to be defined across the studies. Therefore, the transformations of the exposure were not subscripted by the study index i.

The p length vector of the $\hat{\boldsymbol{\beta}}_i$ parameters and the accompanying $p \times p$ covariances matrices

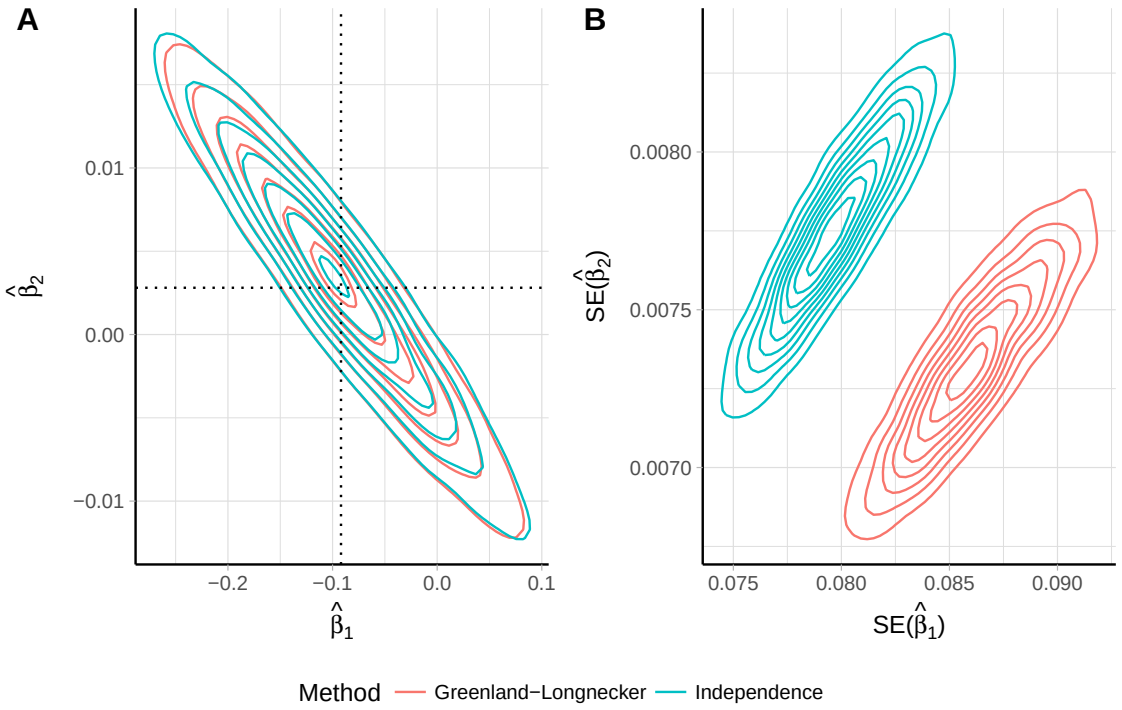

Figure 2.4: Empirical bivariate distribution of the beta coefficients (panel A) and their standard errors (panel B) for a quadratic trend assuming independence of the log $\widehat{\text{RR}}$ and reconstructing the covariances using the Greenland and Longnecker's method. Results are based on simulations with 5000 replications and a true quadratic trend β_1 =-0.092, β_2 =0.003.

$\widehat{\text{Var}}\left(\hat{\boldsymbol{\beta}}_i\right)$ serve as outcome in the meta-analytic model. We consider the setting with $p \geq 2$ and relate the univariate case as a simpler instance of the more general multivariate case. Since the dimension of the outcome is no longer univariate, extensions of models 2.1 to the multivariate settings can be implemented for accommodating the synthesis of correlated estimates **??**.

Model definition

A multivariate random-effects model has a similar formulation as in the univariate case

$$\hat{\boldsymbol{\beta}}_i = \boldsymbol{\beta} + \mathbf{u}_i + \boldsymbol{\varepsilon}_i \tag{2.18}$$

The unobserved random effects $\mathbf{u}_i$ are now of dimension p, still representing study-specific deviation from the mean $\boldsymbol{\beta}$ parameter. As before, they have zero mean $\text{E}[\mathbf{u}_i] = \mathbf{0}$ and $\text{Var}[\mathbf{u}_i] = \boldsymbol{\Psi}$, the $p \times p$ between-study variance matrix. Specification of a parametric distribution for the random-effects may facilitate the inference (especially confidence intervals) and improve the prediction of marginal and conditional dose-response associations. Typically a multivariate normal distribution is adopted $\mathbf{u}_i \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Psi})$. Hence, we can write the marginal model of 2.18 as

$$\hat{\boldsymbol{\beta}}_i \sim \mathcal{N}(\boldsymbol{\beta}, \boldsymbol{\Sigma}_i) \tag{2.19}$$

where the marginal variance $\boldsymbol{\Sigma}_i = \widehat{\text{Var}}\left(\hat{\boldsymbol{\beta}}_i\right) + \boldsymbol{\Psi}$ is defined by the sum of the within-study and between-studies variance components. The model 2.19 implies a two-stage sampling procedure where the study-specific $\boldsymbol{\beta}_i$ parameters are assumed to be sampled from a multivariate normal distribution centered around the population average parameter $\boldsymbol{\beta}$. The study-specific estimates $\hat{\boldsymbol{\beta}}_i$ are themselves sampled from a multivariate distribution with zero mean and error variance assumed known.

The multivariate random-effects model 2.19 can be extended to meta-regression models by including study-levels covariates that might change the shape of the dose-response relationship. The dose-response coefficients are then modeled as a linear combination of the m study-level covariates $\mathbf{z}_i = (z_{i1}, \ldots, z_{im})$, with $z_{i1} = 1$ indicating the intercept term

$$\hat{\boldsymbol{\beta}}_i \sim \mathcal{N}\left(\mathbf{Z}_i\boldsymbol{\beta}, \boldsymbol{\Sigma}_i\right) \tag{2.20}$$

The $p \times pm$ design matrix $\mathbf{Z}_i$ is constructed taking the Kronecker product between the $\mathbf{z}_i$ and the identity matrix of dimension p ($\mathbf{I}_{(p)}$)

$$\mathbf{Z}_i = \mathbf{I}_{(p)} \otimes \mathbf{z}_i^\top = \begin{bmatrix} 1 & z_{i2} & \cdots & z_{im} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & & & \ddots & & & & \\ 0 & 0 & \cdots & 0 & \cdots & 1 & z_{i2} & \cdots & z_{im} \end{bmatrix} \tag{2.21}$$

For example, the $\mathbf{Z}_i$ matrix relating the effect of a binary variable z_i to the dose–response coefficients for a quadratic trend is

$$\mathbf{Z}_i = \mathbf{I}_{(2)} \otimes \mathbf{z}_i^\top = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes (1, z_i) = \begin{bmatrix} 1 & z_i & 0 & 0 \\ 0 & 0 & 1 & z_i \end{bmatrix}$$

The dimension of $\hat{\boldsymbol{\beta}}$ is now $m \times p$. The coefficients related to the intercept terms are interpreted as the mean dose-response coefficient when all the study-level covariates $\mathbf{z}$ are equal to zero. The remaining coefficients indicate how the mean dose-response asspciation vary with respect to the corresponding study-level covariate.

Estimation

Several methods are available for estimating the parameters of interest, namely the $p \times m$ dose–response coefficients in $\boldsymbol{\beta}$ and the $p(p+1)/2$ length vector $\boldsymbol{\xi}$ containing the elements of the between-studies covariance $\boldsymbol{\Psi}$. There is generally no reason to assume a specific covariance structure (White *et al.*, 2011). We consider here likelihood-based estimators (Verbeke, 1997; Pinheiro and Bates, 2010). In particular, ML estimators estimate simultaneously $\boldsymbol{\beta}$ and $\boldsymbol{\xi}$ by maximizing the log-likelihood of the marginal model 2.20

$$\ell\left(\boldsymbol{\beta}, \boldsymbol{\xi}\right) = -\frac{1}{2} Ip \log(\pi) - \frac{1}{2}\sum_{i=1}^{I} \log|\boldsymbol{\Sigma}_i| - \frac{1}{2}\sum_{i=1}^{I}\left[\left(\hat{\boldsymbol{\beta}}_i - \mathbf{Z}_i\boldsymbol{\beta}\right)^\top \boldsymbol{\Sigma}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \mathbf{Z}_i\boldsymbol{\beta}\right)\right] \tag{2.22}$$

ML estimators, however, don't take into account the loss of degrees of freedom due to the $\boldsymbol{\beta}$ estimation Harville (1977). Alternatively, restricted maximum likelihood methods (REML) maximizes a set of contrasts defined as a function of the only covariance parameters

$$\ell_R(\xi) = -\frac{1}{2}(Ip - pm) - \frac{1}{2}\sum_{i=1}^{I}\log|\boldsymbol{\Sigma}_i| - \frac{1}{2}\sum_{i=1}^{I}\log\left|\mathbf{Z}_i^\top \boldsymbol{\Sigma}_i \mathbf{Z}_i\right| + \tag{2.23}$$

$$-\frac{1}{2}\sum_{i=1}^{I}\left[\left(\hat{\boldsymbol{\beta}}_i - \mathbf{Z}_i\boldsymbol{\beta}\right)^\top \boldsymbol{\Sigma}_i^{-1}\left(\hat{\boldsymbol{\beta}}_i - \mathbf{Z}_i\boldsymbol{\beta}\right)\right] \tag{2.24}$$

Both estimation methods require iterative algorithms, where conditional estimates of $\hat{\boldsymbol{\beta}}$ are plugged in either 2.3.2 or 2.3.2, regarded as function of ξ only, until convergence. More details on the implementation of iterative methods for maximizing equations and are described by Gasparrini *et al.* (2012).

Hypothesis testing and heterogeneity

There are two main domains of interest for making inference that relate either to the fixed-effects $\boldsymbol{\beta}$ or the variance components in $\boldsymbol{\Psi}$. Using the normality assumption for the random-effects, inference is based on the approximated normal distribution for $\hat{\boldsymbol{\beta}}$, with mean and covariance matrix defined similarly as in equation 2.16.
Since the mean dose-response association is defined by the $\boldsymbol{\beta}$, the hypothesis of no association can be evaluated by testing $H_0 : \boldsymbol{\beta} = \mathbf{0}$. Alternatively, a subset or linear combinations of $\boldsymbol{\beta}$ may be of interest. For example, in a quadratic trend the non-linearity is introduced by the quadratic term x^2. Thus, testing $H_0 : \beta_2 = 0$ is a possible way for evaluating departure from a linear dose-response relationship.

As previously presented in section 2.1.2, the coefficients defining $\boldsymbol{\Psi}$ are not nuisance parameters rather than useful for quantifying the variation of the study-specific associations $\boldsymbol{\beta}_i$. Similar measures for testing and quantifying the impact of heterogeneity have been extended to the multivariate setting (Berkey *et al.*, 1996). In particular, the Q statistic

$$Q = \sum_{i=1}^{I}\left(\hat{\boldsymbol{\beta}}_i - \mathbf{Z}_i\hat{\boldsymbol{\beta}}_{\text{fe}}\right)^\top \widehat{\text{Var}}\left(\hat{\boldsymbol{\beta}}_i\right)^{-1}\left(\hat{\boldsymbol{\beta}}_i - \mathbf{Z}_i\hat{\boldsymbol{\beta}}_{\text{fe}}\right) \tag{2.25}$$

with $\hat{\boldsymbol{\beta}}_{\text{fe}}$ estimated under a fixed-effect model, is used to test $H_0 : \boldsymbol{\Psi} = \mathbf{0}$. Under the null hypothesis, the Q statistic follow a χ^2 with $Ip - pm$ degrees of freedom. When $p = 1$ the formulations 2.4 and 2.25 coincide. The multivariate extension of the I^2 was derived relating the Q statistics to its degrees of freedom $I^2 = \max\left\{0, \frac{Q-(Ip-pm)}{Ip-pm}\right\}$ (Jackson *et al.*, 2012).

Prediction

Oftentimes the estimated mean coefficients $\hat{\boldsymbol{\beta}}$ are not directly interpretable (an exception is the estimate for a linear trend). The dose-response results are thus communicated as predicted

(log) relative risk for selected exposure levels using one value as referent. Obtaining predictions either in a graphical or tabular presentation is thus an important step of the analysis and yet often poorly implemented. Based on the model 2.18, the predicted $\log RR$ for a dose level x_v using x_{ref} as referent can be calculated as

$$\log \widehat{RR}(x = x_v) = \mathbf{X}_v \hat{\boldsymbol{\beta}} \tag{2.26}$$

$$\text{Var}\left(\log \widehat{RR}(x = x_v)\right) = \mathbf{X}_v \widehat{\text{Var}}\left(\hat{\boldsymbol{\beta}}\right) \mathbf{X}_v^\top \tag{2.27}$$

where $\mathbf{X}_v$ is the design matrix defined in the first-stage analysis (equation 2.14). For example, the predicted $\log RR$ for the quadratic model 2.12 comparing x_v versus x_{ref} is

$$\log \widehat{RR}(x = x_v) = \hat{\beta}_1 (x_v - x_{\text{ref}}) + \hat{\beta}_2 \left(x_v^2 - x_{\text{ref}}^2\right)$$

Of note, the referent dose x_{ref} is an arbitrary value and thus does not need to correspond to any of the study-specific reference values x_{i0}.

A confidence interval for the predicted $\log \widehat{RR}(x = x_v)$ is based on the normal distribution of $\hat{\boldsymbol{\beta}}$

$$\log \widehat{RR}(x = x_v) \mp z_{1-\alpha/2} \text{Var}\left(\log \widehat{RR}(x = x_v)\right)^{\frac{1}{2}}$$

Formulas 2.26 and 2.27 can be extended to meta-regression models. The predicted $\log RR$ conditional on a specific study-level covariate pattern $\mathbf{z} = \mathbf{z}_v$ is

$$\log \widehat{RR}\,(x = x_v, \mathbf{z} = \mathbf{z}_v) = \mathbf{X}_v \left(\mathbf{I}_{(p)} \otimes \mathbf{Z}_v^\top\right) \hat{\boldsymbol{\beta}} \tag{2.28}$$

$$\text{Var}\left(\log \widehat{RR}\,(x = x_v, \mathbf{z} = \mathbf{z}_v)\right) = (\mathbf{X}_v \mathbf{Z}_v) \widehat{\text{Var}}\left(\hat{\boldsymbol{\beta}}\right) (\mathbf{X}_v \mathbf{Z}_v)^\top \tag{2.29}$$

2.3.3 Methodological research

Dose–response meta-analysis has received attention not only in applied works but also in theoretical articles that covered different aspects of the methodology. Greenland and Longnecker originally presented the two-stage approach for efficiently estimating a linear trend in a fixed-effect analysis. An alternative model for estimating curvilinear model, referred to as “pool first”, was also presented. The technique consists of a one-stage approach where the aggregated data are considered altogether and a single model as in 2.13 is fitted. By first combining the data, more flexible curve such as polynomials or splines can be estimated. The study-specific dose-response analyses are limited by the minimum number of non-referent log RRs across the studies. For example, if the aggregated data for a study consists of only one non referent log RR, only a univariate model with $p = 1$ can be estimated. The authors refined the methodology by extending the two-stage approach to allow for heterogeneity limited to a linear trend analysis (Berlin *et al.*, 1993).

Flexible dose–response models

The primary interest of the methodological reserach was in presenting alternative strategies for estimating non-linear curves. Bagnardi et al. described the use of fractional polynomials and restricted cubic splines using aggregated dose–response data (Bagnardi *et al.*, 2004). Based on a practical example on the association between alcohol consumption and all-cause mortality, the authors showed how implementation of these flexible techniques may prevent misleading results from conventional polynomials (e.g. quadratic) curves. Fractional polynomials of order two (FP2) consist of a large family of curves defined in the general form of

$$\text{FP2}(x) = \beta_1 x^{p_1} + \beta_2 x^{p_2} \tag{2.30}$$

where p_1 and p_2 are chosen in the set of power coefficients $\{-2, -1, -0.5, 0, 0.5, 1, 2, 3\}$ (Royston and Altman, 1994). When $p = 0$, x^p becomes $\log(x)$, while if $p_1 = p_2$ the second transformation of x becomes $x^{p_2}\log(x)$. The advantage of FP2 models is that different shapes, including U- and J-shapes, can be estimated by only two coefficients chosen using different combinations for the power terms (p_1, p_2). Typically, the best fitting fractional polynomial is chosen in such a way that the (p_1, p_2) corresponds to the model with the highest likelihood, or equivalently, lowest deviance.

A popular alternative for flexibly model the dose–response association is represented by the use of splines (De Boor *et al.*, 1978), largely presented by Orsini *et al.* (2011b) using data from the Pooling Project of Prospective Studies of Diet and Cancer. Splines functions consist of consecutive polynomials connected at specific points of the exposure range called knots. Choosing $\mathbf{k} = (k_1, \ldots, k_K)$ knots and third order polynomials, the model, also known as cubic splines (CS), is defined as

$$\text{CS}(x) = \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{l=1}^{K-1} \beta_{l+3}(x - k_l)_+^3 \tag{2.31}$$

where the '+' notation has been used ($u_+ = u$ if $u \geq 0$ and $u_+ = 0$ otherwise). To avoid strange behaviors at the extremes of the exposure range, the model 2.31 is constraint to be linear before and after the first and last knots, respectively. For example, using three knots a restricted cubic spline model (RCS) can be specified in terms of two coefficients

$$\text{RCS}(x) = \beta_1 x + \beta_2 \left[(x - k_1)_+^3 - \frac{k_3 - k_1}{k_3 - k_2}(x - k_2)_+^3 + \frac{k_2 - k_1}{k_3 - k_2}(x - k_3)_+^3 \right] \tag{2.32}$$

The second transformation is generally divided by $(k_3 - k_1)^2$ to improve the numerical behavior and to put the spline transformations on the same scale (Harrell Jr, 2015).

The previous strategies have been often presented for exposures where 0 was the natural reference categories (e.g. alcohol consumption). Liu *et al.* (2009) extended the methodology for handling non-zero reference categories, as in the case of Body Mass Index. In particular, they first described how to construct the design matrix in terms of contrasts, as clarified in

equation 2.14. Misspecification of the design matrix (i.e. $\log \text{RR} = \beta_1(x_1 - x_0) + \beta_2(x_1 - x_0)^2$ instead of $\log \text{RR} = \beta_1(x_1 - x_0) + \beta_2(x_1^2 - x_0^2)$) may increase the risk of generating artifacts and misleading conclusions. Note that for zero exposure categories the problem is generally not relevant since many functions return zero for $x = 0$ ($g(0) = 0$).

Multivariate meta-analysis

The major contribution for estimating non-linear curves in a random-effects analysis came with the extension of univariate meta-analytic models to the multivariate case. The formalization and implementation of multivariate meta-analysis enabled the extension of a two-stage dose–response meta-analysis to the more complex case of multiple parameters association (Gasparrini *et al.*, 2012). The multivariate framework can accommodate the synthesis of correlated outcomes, or estimates, derived in the first stage dose–response analyses. The application of the strategies presented in 2.30 and 2.32 in a random-effects setting has been easily facilitated by the implementation of dedicated packages for multivariate meta-analysis (White *et al.*, 2011; Jackson *et al.*, 2011).
More generally, the methodological advancements for meta-analysis were diverse and numerous (see Sutton and Higgins (2008) for an overview). Important improvements that directly affected how results are presented related mainly to the quantification and assessment of heterogenity, with the definition of the measures presented in section 2.1.2. In addition, many other articles enriched the set of tools for pooling study-specific effects, with a direct application to the second stage of a dose–response meta-analysis. Among the many, it is worth to mention the implementation of several estimation methods for the between-study variability (see Langan *et al.* (2017) for a comparison based on simulation studies); advancement in performing meta-regression (Van Houwelingen *et al.*, 2002); proposal of sequential approaches (Pogue and Yusuf, 1997) and statistical power (Sutton *et al.*, 2007); and introduction of Bayesian methods (Sutton and Abrams, 2001).

Covariance and sensitivity analysis

Orsini *et al.* (2006) refined the initial formulas presented by Greenland and Longnecker for approximating the study-specific covariance matrices depending on the study-design. **?**berrington2003generalized) described an alternative method that avoids the reconstruction of the covariance matrices. Instead, upper and lower bounds for the covariance matrix are used in a sensitivity analysis of the dose-response coefficients, adopting a range of plausible values for the covariances. While the alternative algorithm proposed by Hamling *et al.* (2008) was presented in section 2.3.1, Easton *et al.* (1991) proposed the implementation of the floating absolute risks where the parameters and their standard errors can be estimated without specifying a baseline group and thus can be regarded as independent. Using individual patient data from the Pooling Project of Prospective Studies of Diet and Cancer (`http://www.hsph.harvard.edu/poolingproject`), negligible differences in the reconstructed covariance matrices were found comparing the three approaches Orsini *et al.* (2011b). Of

note, none of the methods would be needed if the authors of the original articles provided the covariance matrix along with the estimated coefficients, as it is usually done in consortia projects.

Berlin *et al.* (1993) presented alternatives for assigning the dose levels within exposure categories and illustrated the use of meta-regression models for investigating the possible effect of study-level characteristics on the estimated linear trends. Shi and Copas (2004) further discussed the issue of dose assignment in grouped measures allowing for arbitrary dose levels. In addition, they investigated the effect of heterogeneity and publication bias by means of sensitivity analyses. A similar problem of dose assignment was addressed by using a likelihood approach limited to a linear trend analysis (Takahashi and Tango, 2010). This idea has been further extended to the case of restricted cubic splines (Takahashi *et al.*, 2013).

2.3.4 Research questions

There are still many open research questions that need to be addressed to improve the synthesis of aggregated dose–response data. It was useful to observe the current practice in applied works in order to identify the more relevant questions. We search the PubMed database for articles published between January 1, 2013 and April 1, 2013 using the research query ("meta-analysis" [Title] and "dose-response" [Title]) and, after excluding irrelevant articles, found 42 applied dose–response meta-analyses. The authors of the select articles conducted a linear trend analysis most of the times (25 times, 60%) while only 17 articles considered non-linear associations by means of restricted cubic splines (15) and fractional polynomials (12). The papers modelling non-linear curves reported a graphical presentation of the pooled dose–response association.

Interestingly, none of the retrieved articles evaluated the goodness-of-fit of the selected dose–response model. The assessment of how the estimated curve fits the aggregated data should be a natural and important step in a dose–response analysis. In Paper II we will address this important issue by presenting relevant measures and graphical tools to help the assessment of goodness-of-fit.

The majority of the screened papers (39, 93%) quantify the impact of heterogeneity by reporting results for the Q test and the value of I^2. 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. In Paper III we overcome this limitation by proposing an alternative measure of heterogeneity and comparing the performance of the new and available measures.

None of the surveyed meta-analyses discussed the sensitivity of the overall dose-response relationship to differences in study-specific exposure distribution. This analysis can be very relevant in case of studies reporting results for heterogeneous exposure range and can enhance estimation by limiting the impact of extrapolation. The point-wise average approach presented

by Saurebrei and Royston in the context of individual patient data may represent an interesting alternative to the averaging of regression coefficients. In Paper IV we will evaluate the advantages of this strategy based on aggregated dose-response data.
In all the meta-analyses assessing departure from of linearity, the authors excluded those studies reporting less than two non-referent RRs. Indeed, a two-stage dose-response meta-analysis requires that all the models in the dose-response analysis are identifiable. A one stage approach would avoid that requirement. Such an approach is conceptually easier to understand, and more elegant from a statistical point of view. In addition, it allows investigation of much more flexible dose-response curves, that are not possible within the context of a traditional two-stage analysis. Aim of Paper V is to describe implementation and advantages of a one-stage random-effects dose-response meta-analysis of aggregated data.

2.4 Software

Dissemination of new statistical methodologies is certainly facilitated by the development and implementation of statistical software components. Many theoretical papers have not been considered in applied works because of lack of user-friendly software.
In 2006 Orsini et al. described the `glst` command in Stata, the first publicly available procedure dedicated for dose–response meta-analysis. The command implements both the one- and two-stage approaches limited, in case of a random-effects analysis, to a linear trend. A two-stage random-effects meta-analysis of non-linear relationships can be performed with the aid of the `mvmeta` command for multivariate meta-analysis. Several worked examples and codes are available at `http://www.imm.ki.se/biostatistics/glst/`. Later on, Li and Spiegelman wrote the macro `%metadose`, a similar procedure for SAS users.

The majority of the applied meta-analyses retrieved in our survey were performed using the `glst` procedure in Stata (36, 87%), while 2 used the `metadose` macro in SAS, 2 functions in RevMan. No dedicated package was available in the free software programming language R. Therefore, in 2013, we released the first version of the **dosresmeta** package on CRAN (`https://CRAN.R-project.org/package=dosresmeta`), a package specifically designed for dose-response meta-analysis in R, with specific functions that greatly facilitates the application in practical works.

The **dosresmeta** package is now available in the updated version 2.0.1 and new features are being implemented in the version under development on GitHub (`https://github.com/alecri/dosresmeta`). Currently, the **dosresmeta** package is downloaded and used worldwide, with a median number of 260 downloads/month (Figure 2.5). The countries where it has been downloaded most are Great Britain (4005), United States (3905), and China (1605) (Figure 2.6). Working examples, codes, and data are available at `http://alecri.github.io/software` for fully reproduce figures and numbers presented in both applied and theoretical papers.

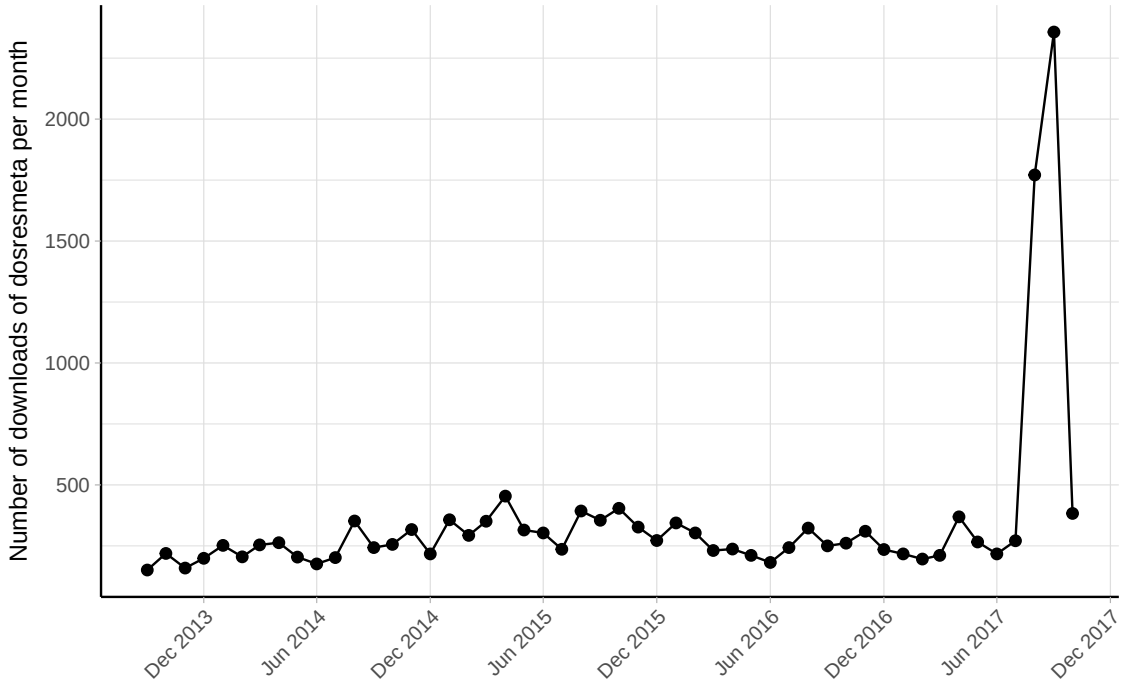

Figure 2.5: Monthly number of downloads of the dosresmeta R package from the RStudio CRAN mirror September 2013 - December 2017.

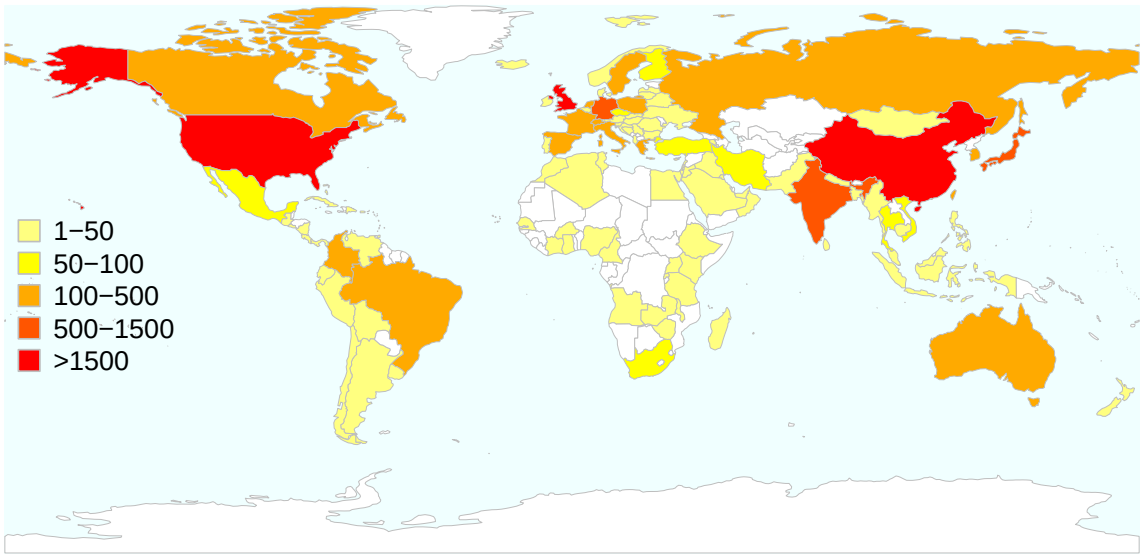

Figure 2.6: Total number of downloads of the doseresmeta R package worldwide from the RStudio CRAN mirror September 2013 - December 2017.

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

Bagnardi V, Zambon A, Quatto P, Corrao G (2004). "Flexible meta-regression functions for modeling aggregate dose-response data, with an application to alcohol and mortality." *American journal of epidemiology*, **159**(11), 1077–1086.

Berkey C, Anderson J, Hoaglin D (1996). "Multiple-outcome meta-analysis of clinical trials." *Statistics in medicine*, **15**(5), 537–557.

Berkey CS, Hoaglin DC, Mosteller F, Colditz GA (1995). "A random-effects regression model for meta-analysis." *Statistics in medicine*, **14**(4), 395–411.

Berlin JA, Longnecker MP, Greenland S (1993). "Meta-analysis of epidemiologic dose-response data." *Epidemiology*, **4**(3), 218–228.

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*.

De Boor C, De Boor C, Mathématicien EU, De Boor C, De Boor C (1978). *A practical guide to splines*, volume 27. Springer-Verlag New York.

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*.

Easton DF, Peto J, Babiker AG (1991). "Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group." *Statistics in medicine*, **10**(7), 1025–1035.

Gasparrini A, Armstrong B, Kenward M (2012). "Multivariate meta-analysis for non-linear and other multi-parameter associations." *Statistics in medicine*, **31**(29), 3821–3839.

Greenland S (1995). "Dose-response and trend analysis in epidemiology: alternatives to categorical analysis." *Epidemiology*, pp. 356–365.

Greenland S, Longnecker MP (1992). "Methods for trend estimation from summarized dose-response data, with applications to meta-analysis." *American journal of epidemiology*, **135**(11), 1301–1309.

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

Hamling J, Lee P, Weitkunat R, Ambühl M (2008). "Facilitating meta-analyses by deriving relative effect and precision estimates for alternative comparisons from a set of estimates presented by exposure level or disease category." *Statistics in medicine*, **27**(7), 954–970.

Harrell Jr FE (2015). *Regression modeling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*. Springer.

Harville DA (1977). "Maximum likelihood approaches to variance component estimation and to related problems." *Journal of the American Statistical Association*, **72**(358), 320–338.

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.

Higgins JP (2008). "Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified." *International journal of epidemiology*, **37**(5), 1158–1160.

Jackson D, Riley R, White IR (2011). "Multivariate meta-analysis: Potential and promise." *Statistics in Medicine*, **30**(20), 2481–2498.

Jackson D, White IR, Riley RD (2012). "Quantifying the impact of between-study heterogeneity in multivariate meta-analyses." *Statistics in medicine*, **31**(29), 3805–3820.

Langan D, Higgins J, Simmonds M (2017). "Comparative performance of heterogeneity variance estimators in meta-analysis: a review of simulation studies." *Research synthesis methods*, **8**(2), 181–198.

Liu Q, Cook NR, Bergström A, Hsieh CC (2009). "A two-stage hierarchical regression model for meta-analysis of epidemiologic nonlinear dose–response data." *Computational Statistics & Data Analysis*, **53**(12), 4157–4167.

Orsini N, Bellocco R, Greenland S, *et al.* (2006). "Generalized least squares for trend estimation of summarized dose-response data." *Stata Journal*, **6**(1), 40.

Orsini N, Greenland S, *et al.* (2011a). "A procedure to tabulate and plot results after flexible modeling of a quantitative covariate." *Stata Journal*, **11**(1), 1.

Orsini N, Li R, Wolk A, Khudyakov P, Spiegelman D (2011b). "Meta-analysis for linear and nonlinear dose-response relations: examples, an evaluation of approximations, and software." *American journal of epidemiology*, **175**(1), 66–73.

Pinheiro J, Bates D (2010). *Mixed-Effects Models in S and S-PLUS*. Springer Science & Business Media. ISBN 978-1-4419-0318-1.

Pogue JM, Yusuf S (1997). "Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis." *Controlled clinical trials*, **18**(6), 580–593.

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)*.

Royston P, Altman DG (1994). "Regression using fractional polynomials of continuous covariates: parsimonious parametric modelling." *Applied statistics*, pp. 429–467.

Royston P, Altman DG, Sauerbrei W (2006). "Dichotomizing continuous predictors in multiple regression: a bad idea." *Statistics in medicine*, **25**(1), 127–141.

Shi JQ, Copas J (2004). "Meta-analysis for trend estimation." *Statistics in medicine*, **23**(1), 3–19.

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.

Sutton AJ, Abrams KR (2001). "Bayesian methods in meta-analysis and evidence synthesis." *Statistical methods in medical research*, **10**(4), 277–303.

Sutton AJ, Cooper NJ, Jones DR, Lambert PC, Thompson JR, Abrams KR (2007). "Evidence-based sample size calculations based upon updated meta-analysis." *Statistics in medicine*, **26**(12), 2479–2500.

Sutton AJ, Higgins J (2008). "Recent developments in meta-analysis." *Statistics in medicine*, **27**(5), 625–650.

Takahashi K, Nakao H, Hattori S (2013). "Cubic spline regression of J-shaped dose-response curves with likelihood-based assignments of grouped exposure levels." *J. Biom. Biostat*, **4**, 1–6.

Takahashi K, Tango T (2010). "Assignment of grouped exposure levels for trend estimation in a regression analysis of summarized data." *Statistics in medicine*, **29**(25), 2605–2616.

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.

Turner EL, Dobson JE, Pocock SJ (2010). "Categorisation of continuous risk factors in epidemiological publications: a survey of current practice." *Epidemiologic Perspectives & Innovations*, **7**(1), 9.

Van Houwelingen HC, Arends LR, Stijnen T (2002). "Advanced methods in meta-analysis: multivariate approach and meta-regression." *Statistics in medicine*, **21**(4), 589–624.

Verbeke G (1997). *Linear mixed models for longitudinal data*. Springer.

White IR, *et al.* (2011). "Multivariate random-effects meta-regression: updates to mvmeta." *Stata Journal*, **11**(2), 255.

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).