Karolinska Institutet

Novel methods for dose–response meta–analysis

Half–time seminar

Alessio Crippa

12th May 2016

Acknowledgements

The members of the half–time committee

- Paul Lambert, University of Leicester
- Finn Rasmussen, Karolinska Institutet
- Martin Sköld, Stockholm University

My main supervisor

- Nicola Orsini, Karolinska Institutet

My co-supervisors

- Alicja Wolk, Karolinska Institutet
- Matteo Bottai, Karolinska Institutet
- Donna Spiegelman, Harvard School of Public Health

My mentor and co-authors

The audience

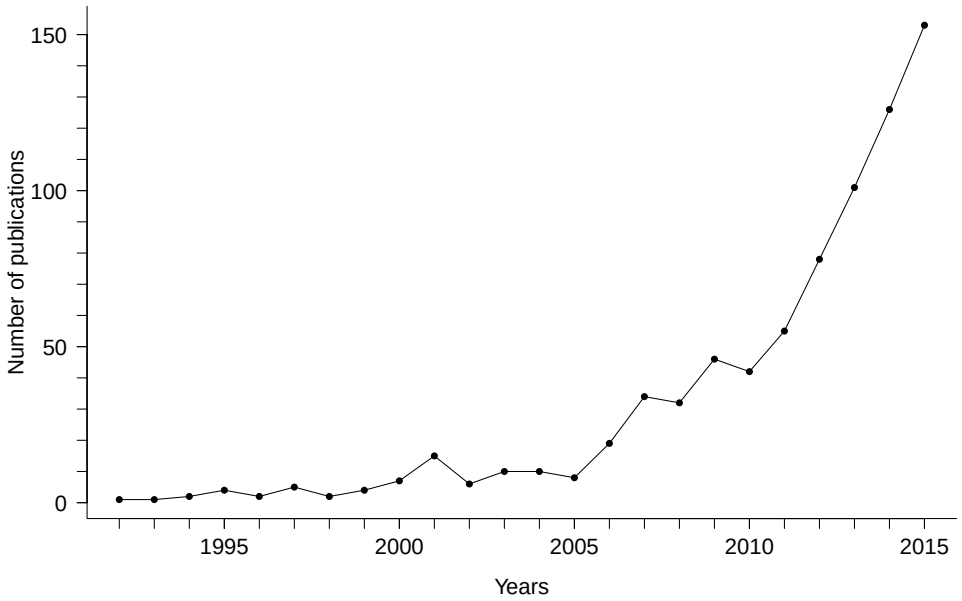
Dose–response meta–analysis

Summarize results on the relation between a quantitive exposure and the occurence of a health outcome

Research questions

- Is there any association between the quantitative exposure and the outcome? What is the shape of the association?
- What are the exposure values associated with the best or worst outcome?
- What are the factors that can influence the dose–response shape?

Increasing number of dose–response meta–analysis



Data source: Web of Science

- Several fields of application
- Many leading medical and epidemiological journals
- Global health organizations and foundations
- Measures of public health impact

Aggregated data

An example form a case-control data on alcohol consumption and breast cancer risk

g/day	dose	case	n	RR	95% CI
Ref.	0	165	337	1.00	—
<2.5	2	74	167	0.80	0.51, 1.27
2.5-9.3	6	90	186	1.16	0.73, 1.85
>9.3	11	122	212	1.57	0.99, 2.51

The RRs are not independent

RR = 1 for the referent category

Two stage dose–response meta–analysis

First stage

Define and estimate a common dose–response model in each study ($i = 1, \ldots, K$)

Second stage

Combine study–specific regression coefficients

Dose–response analysis

Log–linear model

$$\mathbf{y}_i = \mathbf{X}_i \beta_i + \varepsilon_i \tag{1}$$

$\mathbf{y}_i$ vector of non-referent log RRs in the i-th study

$\mathbf{X}_i$ contains the assigned doses (and/or transformations)

- Model without intercept
- $\mathrm{Cov}(\varepsilon_i) = \mathbf{\Sigma}_i$ can be approximated

$$\begin{aligned} \hat{\beta}_i =& (\mathbf{X}_i^\top \mathbf{\Sigma}_i^{-1} \mathbf{X}_i)^{-1} \mathbf{X}_i^\top \mathbf{\Sigma}_i^{-1} \mathbf{y}_i \\ \mathbf{V}_i =& \mathrm{Cov}(\hat{\beta}_i) = (\mathbf{X}_i^\top \mathbf{\Sigma}_i^{-1} \mathbf{X}_i)^{-1} \end{aligned} \tag{2}$$

Meta–analysis

Pooling of $\hat{\beta} = \left[\hat{\beta}_1, \ldots, \hat{\beta}_K\right]$ and $\boldsymbol{V} = [\boldsymbol{V}_1, \ldots, \boldsymbol{V}_K]$

Multivariate random–effect meta–analysis

$$\hat{\beta}_i \sim N_p(\beta, \mathbf{V_i} + \psi) \tag{3}$$

ψ is the between–study covariance matrix

Cochran Q test and measures of heterogeneity

Previous methodological papers

- Random–effects and meta–regression
- Multivariate meta–analysis
- Approximating covariance matrices
- Flexible modeling
- Non–zero reference category
- Evaluation of sources of bias and sensitivity analyses

Open questions

- Assessment of goodness of fit of dose–response meta–analytic models has not yet been discussed (*Paper I*)
- Little emphasis is placed on the assumptions underneath the common measures of heterogeneity (*Paper II*)
- The effect of differential shape and exposure distribution is hard to be addressed in a two-stage approach (*Paper III*)
- Dose–response and meta-regression models may be affected by small number of data points in some of the studies (*Paper IV*)

Paper I

Goodness of fit tools for dose–response meta–analysis of binary outcomes. *Res Synth Meth*, 2015

Specific aim

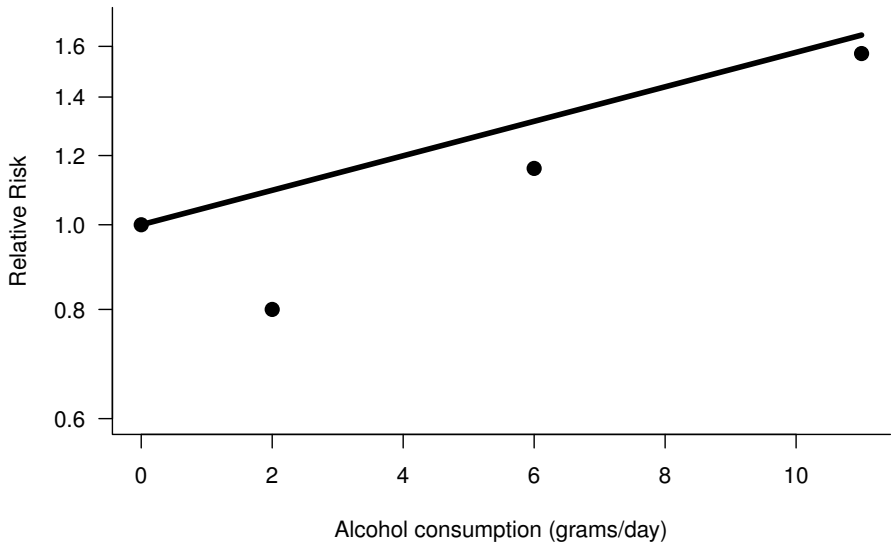
- To present and discuss different tools to evaluate the goodness of fit of dose–response meta–analysis of binary outcomes

Paper I

Does the pooled curve adequately summarize the aggregate data?

This question is typically ignored in published meta–analyses

Those that address this question ignore the correlation among the RRs



Paper I – Goodness of fit tools

Deviance (D)

- Total absolute distance between fitted and reported RRs
- Test for model specification

Coefficient of determination (R^2)

- Descriptive measure of agreement
- Dimensionless measure bounded between 0 and 1

Plot of decorrelated residuals versus exposure

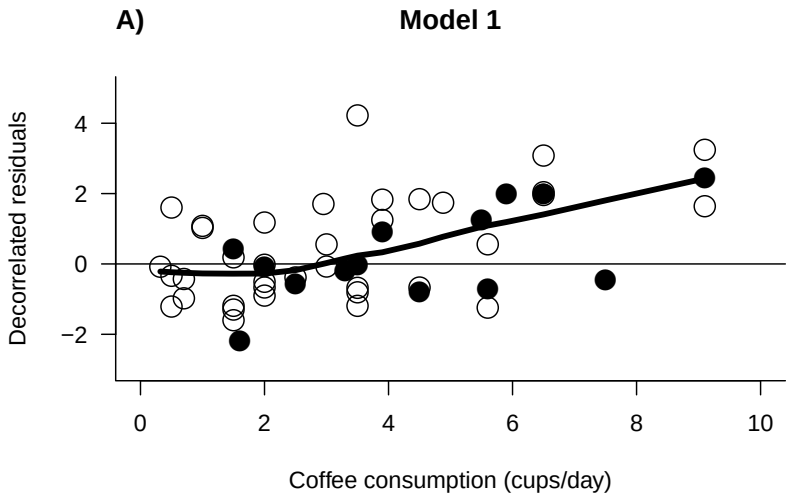
- Visual assessment of the goodness of fit
- Evaluate how the pooled dose–response curve fits the data by exposure levels

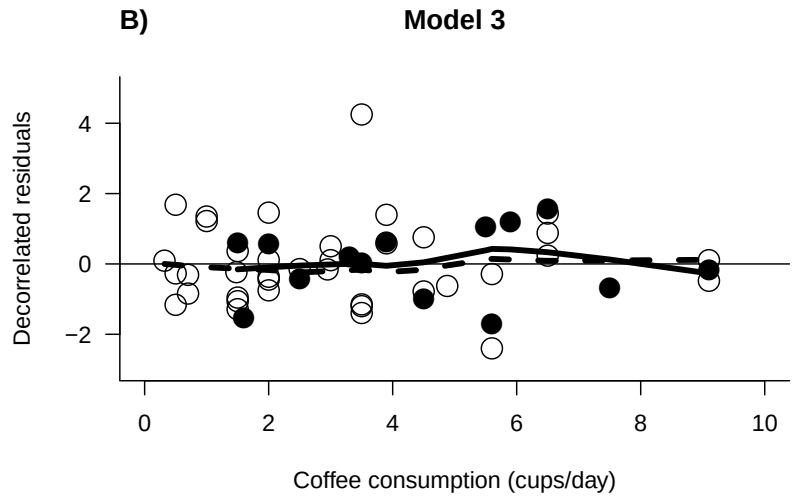
All these tools take into account the correlation between the RRs

Paper I

Is the fit of the dose–response curve coffee and risk of stroke adequate?

Model	Deviance	df	*p*-value	R^2	R^2_{adj}
1) Linear	140	51	<0.0001	41%	39%
2) RCS with 3 knots	75	50	0.01	68%	67%
3) RCS with 3 knots + interaction	64	48	0.06	73%	70%





Paper II

A new measure of between–studies heterogeneity in meta–analysis. *Stat. Med*, 2016

Specific aims

- To propose a new measure of heterogeneity
- Compare the performances of the new estimator through simulations studies

Paper II

Heterogeneity measures, I^2 and R_I, relates the heterogeneity, τ^2, to the total variance, $\tau^2 + \sigma^2$

σ^2 is a summary measure of the observed within-study variance, v_i

Homogeneity of within-studies variances is unlikely to hold

Analysis	**within-study variances**	$\sigma^2(I^2)$	$\sigma^2(R_I)$
A	[6, 6.1, 6.2, 5.9, 6, 5.9, 6.1, 5.8, 6, 6.2]	6.018	6.017
B	[5, 19, 3, 15, 6, 23, 4, 17, 2, 8.8]	6.017	5.602

A measure that relaxes this assumption is desirable

Paper II - R_b a new measure of heterogeneity

The new measure quantifies the contribution of τ^2 relative to the variance of the pooled random effects estimate

$$R_b = \frac{\tau^2}{K\mathrm{Var}\left(\hat{\beta}_{re}\right)} = \frac{1}{K}\sum_{i=1}^{K}\frac{\tau^2}{v_i + \tau^2} \tag{4}$$

R_b satisfied the properties for a measure of heterogeneity

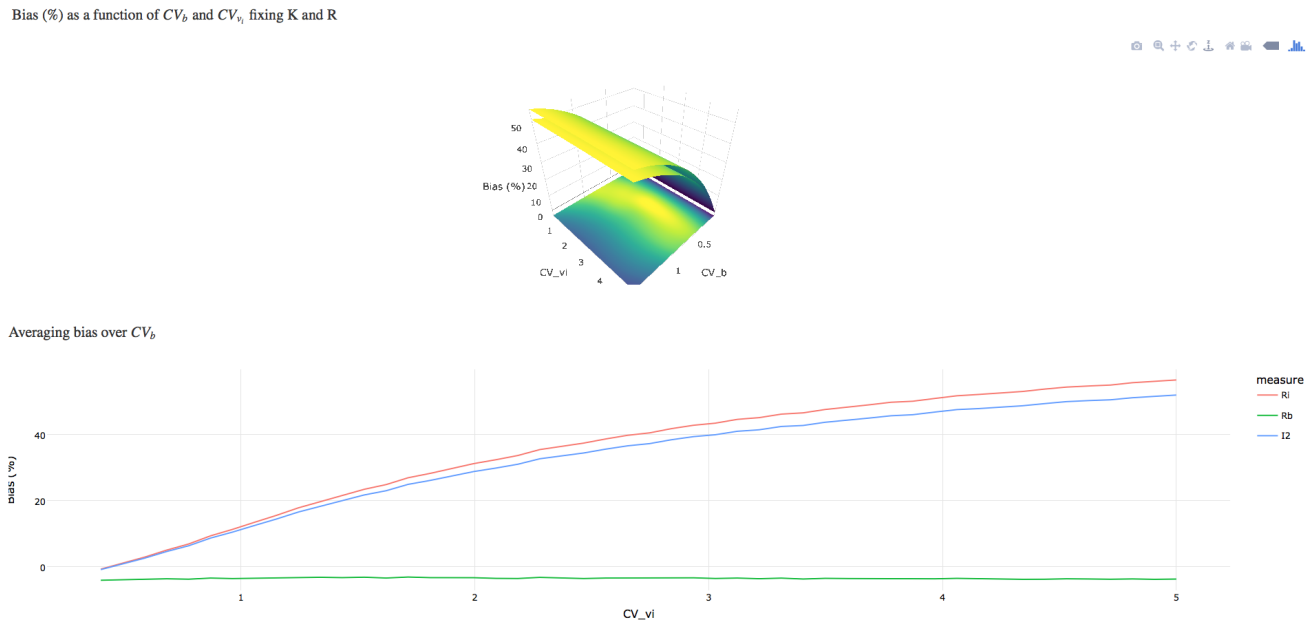
R_b is a consistent and asymptotically normal distributed estimator (Wald-type confidence intervals)

It coincides with I^2 and R_I when $v_i = \sigma^2 \; \forall i = 1, \ldots, K$

Paper II - Simulation study

Different scenario simulations ($R_b = 0.1, 0.5, 0.7$; $CV_{v_i} = 0.5, 1, 2$; $CV_B = 0.5, 1, 3$; $K = 5, 20, 50, 100$)

`https://alecri.shinyapps.io/bias/`



Paper II - Simulation results

No specific pattern in the bias for R_b according to CV_{v_i} and CV_B values

I^2 and R_I overestimated the impact of heterogeneity

The coverage was good for confidence intervals based upon R_b

Bias and coverage for I^2 and R_I worsened as CV_{v_i} increased

Paper III

A Pointwise Approach to Dose–Response Meta–Analysis of Aggregated Data

Specific aims

- To introduce more flexibility in the dose-response analysis
- To allow each study to contribute to the overall curve based on the observed exposure distribution

Paper III

General limitations of a two-stage approach

- Common study-specific functional relationship (1st stage)
- Information on study-specific exposure range is not considered (2nd stage)

Consequences

- Poor fit in some of the study-specific dose-response analyses
- Risk of extrapolating predicted relative risks

A point-wise average approach may overcome those limitations

Paper III - Point-wise average approach

It consists of

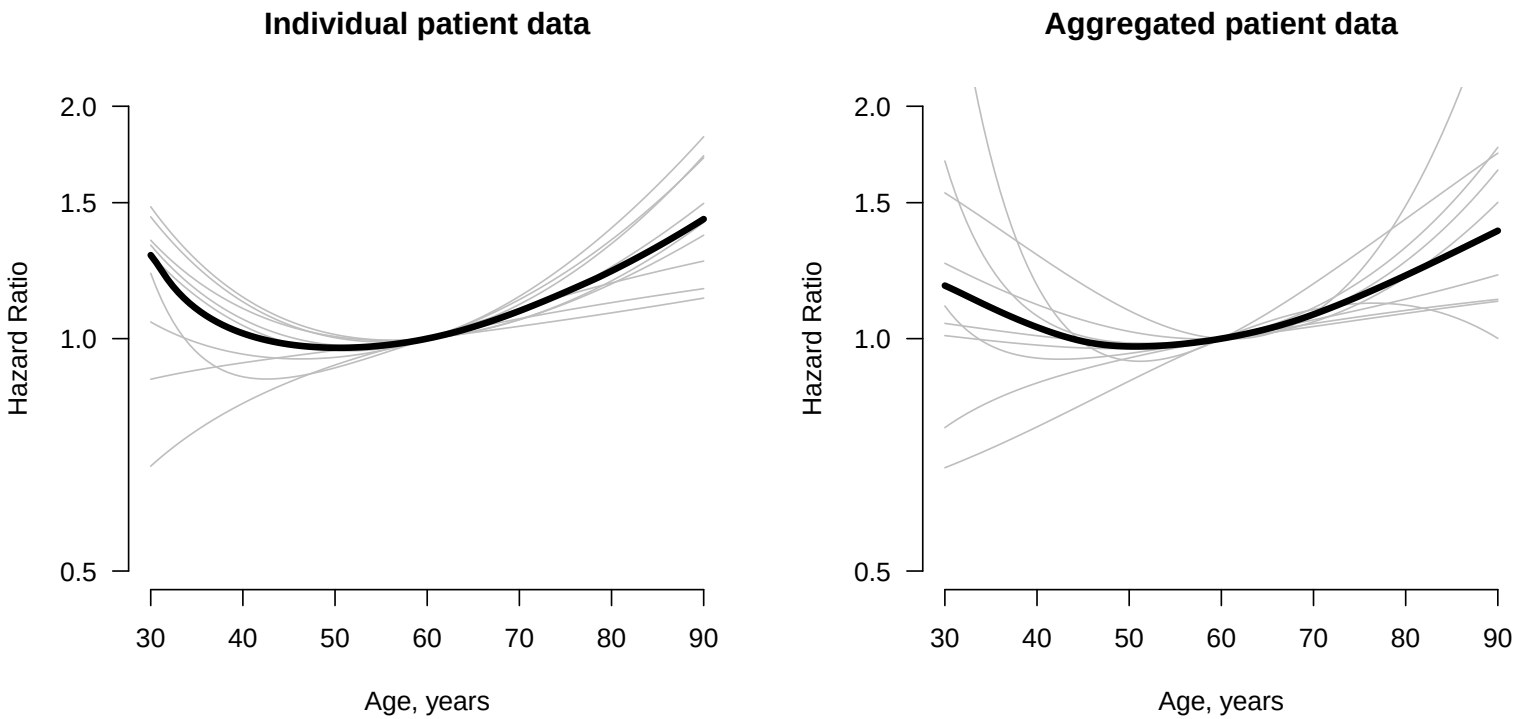
- Estimating study-specific dose-response curves
- Predicting study-specific effects (RRs) for a grid of exposure values
- Combining study-specific effects

Advantages

- The dose-response analyses may vary across studies
- RR predictions can be limited to study-specific exposure ranges

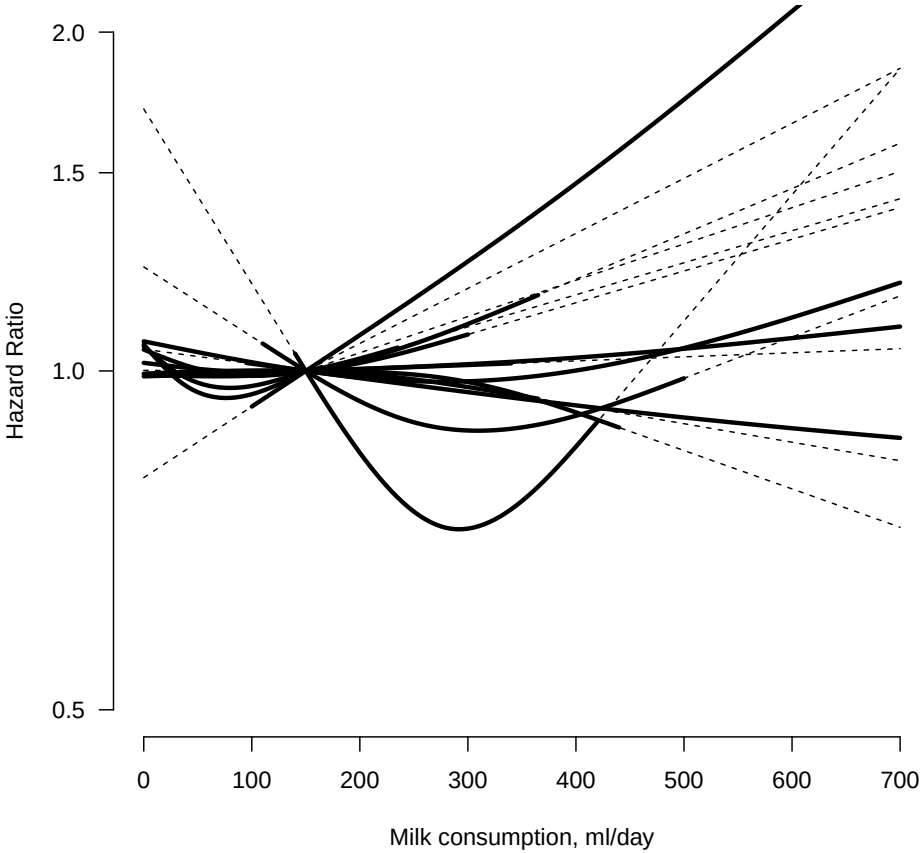
Paper III - Comparison with IPD meta-analysis

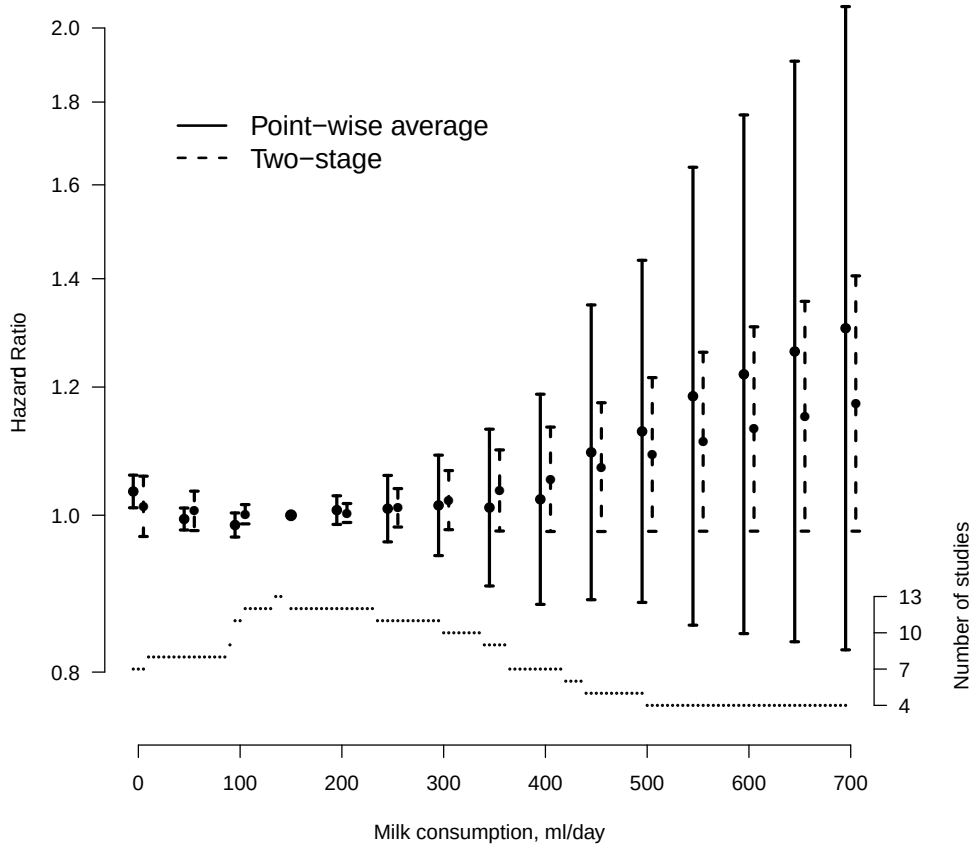
Based on breast cancer patients in the SEER program (`http://seer.cancer.gov/`)



Paper III - Comparison with two-stage meta-analysis

Re-analysis of a dose-response meta-analysis between milk and mortality





Paper IV

One-Stage Dose-Response Meta–Analysis of Aggregated Data

Specific aim

- To describe and implement a one-stage approach for dose–response meta–analysis of aggregated data

Paper IV

A one-stage procedure for random–effects meta–analysis of aggregated dose–response data

- Conceptually easier
- Avoid exclusion of studies with small observations
- More complex curves
- Interaction analysis

Paper IV – One-stage approach

Conditional meta-regression model

$$\underset{n\times 1}{\mathbf{y}_i} = \underset{(n\times p)(p\times 1)}{\mathbf{X}_i\beta} + \underset{(n\times p)(1\times q)(p\cdot q\times 1)}{(\mathbf{X}_i \otimes \mathbf{Z}_i)\ \gamma} + \underset{(n\times p)(p\times 1)}{\mathbf{X}_i\eta_i} + \underset{n\times 1}{\varepsilon_i} \qquad (5)$$

Distributional assumptions

$$\varepsilon_i \sim N_n(\mathbf{0}, \mathbf{\Sigma}_i)$$

$$\eta_i \sim N_p(\mathbf{0}, \mathbf{\Psi})$$

Marginal model

$$\mathbf{Y}_i \sim N_n\left(\mathbf{X}_i\beta + (\mathbf{X}_i \otimes \mathbf{Z}_i)\,\gamma,\ \ \mathbf{\Sigma}_i + \mathbf{X}_i\mathbf{\Psi}\mathbf{X}_i^\top\right) \qquad (6)$$

Paper IV

Software implementation is almost complete (`https://github.com/alecri/dosresmeta`)

If the study-specific dose-response models are identifiable, the one- and two-stage approaches are equivalent

Advantages and limitations will be explored re-analyzing meta-analyses (presenting heterogeneity and meta–regression)

Summary

- Use of the goodness of fit tools can improve practice of quantitative reviews
- The proposed measure of heterogeneity, R_b, can facilitate quantification of the impact of heterogeneity
- The point-wise approach is a flexible tool to evaluate the impact of heterogeneous exposure distributions
- A one-stage meta-analysis will avoid exclusion of studies with limited number of RRs and allow more flexibility in meta-regression models

References I

[1] Sander Greenland and Matthew P Longnecker.
Methods for trend estimation from summarized dose-response data, with applications to meta-analysis.
American journal of epidemiology, 135(11):1301–1309, 1992.

[2] Nicola Orsini, Ruifeng Li, Alicja Wolk, Polyna Khudyakov, and Donna Spiegelman.
Meta-analysis for linear and nonlinear dose-response relations: Examples, an evaluation of approximations, and software.
American journal of epidemiology, 175(1):66–73, 2012.

[3] Jesse A Berlin, Matthew P Longnecker, and Sander Greenland.
Meta-analysis of epidemiologic dose-response data.
Epidemiology, 4(3):218–228, 1993.

[4] A. Gasparrini, B. Armstrong, and M. G. Kenward.
Multivariate meta-analysis for non-linear and other multi-parameter associations.
Statistics in Medicine, 31(29):3821–3839, 2012.

[5] Vincenzo Bagnardi, Antonella Zambon, Piero Quatto, and Giovanni Corrao.
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, 2004.

References II

[6] Qin Liu, Nancy R Cook, Anna Bergström, and Chung-Cheng Hsieh.
A two-stage hierarchical regression model for meta-analysis of epidemiologic nonlinear dose-response data.
Computational Statistics & Data Analysis, 53(12):4157–4167, 2009.

[7] Jian Qing Shi and JB Copas.
Meta-analysis for trend estimation.
Statistics in medicine, 23(1):3–19, 2004.

[8] Willi Sauerbrei and Patrick Royston.
A new strategy for meta-analysis of continuous covariates in observational studies.
Statistics in medicine, 30(28):3341–3360, 2011.