Dose-Response Meta-Analysis of Continuous Outcomes from Aggregated Data

Alessio Crippa[1,2] Nicola Orsini[1,2]

http://www.imm.ki.se/biostatistics/
alessio.crippa@ki.se

[1]Unit of Nutritional Epidemiology, Karolinska Institutet
[2]Unit of Biostatistics, Karolinska Institutet

5th Nordic-Baltic Biometric Conference, Reykjavik

Acknowledgments

I would like to thank my co-authors:

- Stefan Leucht
- John Davis
- Nicola Orsini

Outline

Dose-response analysis

- How does a (continuous) variable vary as a function of an exposure?
- Dose-specific means and standard deviations
- Improving efficiency in (single investigation) findings
- Inconclusive results
- Meta-analytical methods

- No established methodology
- Davis and Chen, 2004
- Thomas et al, E_{max} model

Aims

- Propose a new method to pool aggregated dose-response data
- How to flexibly model the dose/exposure
- Illustrate using a real example

Aggregated dose-response data

Table 1: Notation for aggregated data in the i-th study $i = 1, \ldots, I$.

dose	$\bar{Y}$	$sd(Y)$	n
0	$\bar{Y}_{i0}$	sd_{i0}	n_{i0}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{ij}	$\bar{Y}_{ij}$	sd_{ij}	n_{ij}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{iJ_i}	$\bar{Y}_{iJ_i}$	sd_{iJ_i}	n_{iJ_i}

Effect size computation

Mean differences

$$d_{ij} = \bar{Y}_{ij} - \bar{Y}_{i0} \quad (1)$$

$$\text{Var}\left(d_{ij}\right) = \frac{n_{ij} + n_{i0}}{n_{ij}\, n_{i0}} s^2_{p_i} \quad (2)$$

$$\text{Cov}\left(d_{ij}, d_{ij'}\right) = \text{Var}\left(\bar{Y}_{j0}\right) = \frac{s^2_{i0}}{n_{i0}}, \quad j \neq j'$$

dose index $j = 1, \dots, J_i$

$s^2_{p_i} = \sum_{j=0}^{J_i} \left(n_{ij} - 1\right) sd^2_{ij} / \sum_{j=0}^{J_i} \left(n_{ij} - 1\right)$

Effect size computation

Mean differences

$$d_{ij} = \bar{Y}_{ij} - \bar{Y}_{i0} \tag{1}$$

$$\mathrm{Var}\left(d_{ij}\right) = \frac{n_{ij} + n_{i0}}{n_{ij} n_{i0}} s^2_{p_i} \tag{2}$$

$$\mathrm{Cov}\left(d_{ij}, d_{ij'}\right) = \mathrm{Var}\left(\bar{Y}_{j0}\right) = \frac{s^2_{i0}}{n_{i0}}, \quad j \neq j'$$

dose index $j = 1, \ldots, J_i$

$s^2_{p_i} = \sum_{j=0}^{J_i} \left(n_{ij} - 1\right) sd^2_{ij} / \sum_{j=0}^{J_i} \left(n_{ij} - 1\right)$

Dose-response analysis (1st stage)

$$\boldsymbol{d}_i = f\left(\boldsymbol{x}_i, \boldsymbol{\theta}_i\right) + \varepsilon_i, \quad \varepsilon_i \sim N\left(\mathbf{0}, \hat{\boldsymbol{\Sigma}}_i\right), \quad i = 1, \ldots, I \tag{3}$$

- E_{max} model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i^{\theta_{3i}} / \left(\theta_{2i} + \boldsymbol{x}_i^{\theta_{3i}}\right)$
- Quadratic model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i + \theta_{2i}\boldsymbol{x}_i^2$
- Splines regression $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_{1i} + \theta_{2i}\boldsymbol{x}_{2i}$

Generalized least square estimation

$$\hat{\theta}_i = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{d}_i$$

$$\hat{\boldsymbol{V}}_i = \text{Var}\left(\hat{\theta}_i\right) = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \tag{4}$$

Dose-response analysis (1st stage)

$$\boldsymbol{d}_i = f\left(\boldsymbol{x}_i, \boldsymbol{\theta}_i\right) + \varepsilon_i, \quad \varepsilon_i \sim N\left(\mathbf{0}, \hat{\boldsymbol{\Sigma}}_i\right), \quad i = 1, \ldots, I \tag{3}$$

- $E_{\max}$ model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i^{\theta_{3i}} / \left(\theta_{2i} + \boldsymbol{x}_i^{\theta_{3i}}\right)$
- Quadratic model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i + \theta_{2i}\boldsymbol{x}_i^2$
- Splines regression $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_{1i} + \theta_{2i}\boldsymbol{x}_{2i}$

Generalized least square estimation

$$\hat{\theta}_i = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{d}_i$$

$$\hat{\boldsymbol{V}}_i = \text{Var}\left(\hat{\theta}_i\right) = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \tag{4}$$

Dose-response analysis (1st stage)

$$\boldsymbol{d}_i = f\left(\boldsymbol{x}_i, \boldsymbol{\theta}_i\right) + \varepsilon_i, \quad \varepsilon_i \sim N\left(\mathbf{0}, \hat{\boldsymbol{\Sigma}}_i\right), \quad i = 1, \ldots, I \tag{3}$$

- $E_{\max}$ model $\qquad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i^{\theta_{3i}} / \left(\theta_{2i} + \boldsymbol{x}_i^{\theta_{3i}}\right)$
- Quadratic model $\qquad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i + \theta_{2i}\boldsymbol{x}_i^2$
- Splines regression $\qquad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_{1i} + \theta_{2i}\boldsymbol{x}_{2i}$

Generalized least square estimation

$$\hat{\theta}_i = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{d}_i$$

$$\hat{\boldsymbol{V}}_i = \text{Var}\left(\hat{\theta}_i\right) = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \tag{4}$$

Dose-response analysis (1st stage)

$$\boldsymbol{d}_i = f\left(\boldsymbol{x}_i, \boldsymbol{\theta}_i\right) + \varepsilon_i, \quad \varepsilon_i \sim N\left(\boldsymbol{0}, \hat{\boldsymbol{\Sigma}}_i\right), \quad i = 1, \ldots, I \tag{3}$$

- E_{max} model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i^{\theta_{3i}} / \left(\theta_{2i} + \boldsymbol{x}_i^{\theta_{3i}}\right)$
- Quadratic model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i + \theta_{2i}\boldsymbol{x}_i^2$
- Splines regression $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_{1i} + \theta_{2i}\boldsymbol{x}_{2i}$

Generalized least square estimation

$$\hat{\theta}_i = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{d}_i$$

$$\hat{\boldsymbol{V}}_i = \text{Var}\left(\hat{\theta}_i\right) = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \tag{4}$$

Dose-response analysis (1st stage)

$$\boldsymbol{d}_i = f\left(\boldsymbol{x}_i, \boldsymbol{\theta}_i\right) + \varepsilon_i, \quad \varepsilon_i \sim N\left(\boldsymbol{0}, \hat{\boldsymbol{\Sigma}}_i\right), \quad i = 1, \dots, I \tag{3}$$

- E_{max} model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i^{\theta_{3i}} / \left(\theta_{2i} + \boldsymbol{x}_i^{\theta_{3i}}\right)$
- Quadratic model $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_i + \theta_{2i}\boldsymbol{x}_i^2$
- Splines regression $\quad \boldsymbol{d}_i = \theta_{1i}\boldsymbol{x}_{1i} + \theta_{2i}\boldsymbol{x}_{2i}$

Generalized least square estimation

$$\hat{\boldsymbol{\theta}}_i = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{d}_i$$

$$\hat{\boldsymbol{V}}_i = \text{Var}\left(\hat{\boldsymbol{\theta}}_i\right) = \left(\boldsymbol{X}_i^T \hat{\boldsymbol{\Sigma}}_i^{-1} \boldsymbol{X}_i\right)^{-1} \tag{4}$$

Multivariate meta-analysis (2nd stage)

Regardless the chosen f

$$\hat{\boldsymbol{\theta}}_i \sim N\left(\boldsymbol{\theta}, \hat{\boldsymbol{V}}_i + \boldsymbol{\Psi}\right) \tag{5}$$

- Fixed-effects model ($\boldsymbol{\Psi} = \boldsymbol{0}$)
- Cochran's Q test and I^2
- Random-effects model ($\boldsymbol{\theta}_i$ vary across studies)
- REML estimation $\hat{\boldsymbol{\theta}}$

Dose-findings

$$\hat{d} = f\left(\boldsymbol{X}, \hat{\theta}\right) \tag{6}$$

$$\hat{d} \mp z_{\frac{\alpha}{2}} \sqrt{\text{diag}\left(f\left(\boldsymbol{X}, \hat{\theta}\right)^T \text{Cov}\left(\hat{\theta}\right) f\left(\boldsymbol{X}, \hat{\theta}\right)\right)} \tag{7}$$

Target doses (ED_γ) which produces γ% of the maximum predicted response

$$\widehat{\mathrm{ED}}_\gamma = \underset{x \in (0, x_{\max}]}{\text{argmax}} \left\{ \frac{\widehat{md}_x}{\widehat{md}_{x_{\max}}} \geq \gamma \right\} \tag{8}$$

Confidence intervals by parametric bootstrap resampling

Dose-findings

$$\hat{d} = f\left(\boldsymbol{X}, \hat{\boldsymbol{\theta}}\right) \tag{6}$$

$$\hat{d} \mp z_{\frac{\alpha}{2}} \sqrt{\text{diag}\left(f\left(\boldsymbol{X}, \hat{\boldsymbol{\theta}}\right)^T \text{Cov}\left(\hat{\boldsymbol{\theta}}\right) f\left(\boldsymbol{X}, \hat{\boldsymbol{\theta}}\right)\right)} \tag{7}$$

Target doses (ED_γ) which produces γ% of the maximum predicted response

$$\widehat{\mathrm{ED}}_\gamma = \underset{x \in (0, x_{\max}]}{\text{argmax}} \left\{ \frac{\widehat{md}_x}{\widehat{md}_{x_{\max}}} \geq \gamma \right\} \tag{8}$$

Confidence intervals by parametric bootstrap resampling

Working example

Table 2: Dose-response data for three of five clinical trials investigating effectiveness of aripiprazole (PANNS) in schizoaffective patients.

ID	Author, Year	dose	$\bar{Y}$	sd(Y)	n	d	Var(d)
1	Cutler, 2006	0	5.300	18.310	85	0.000	0.000
		2	8.230	18.320	92	2.930	7.593
		5	10.600	18.310	89	5.300	7.715
		10	11.300	18.320	94	6.000	7.515
2	McEvoy, 2007	0	2.330	26.100	107	0.000	0.000
		10	15.040	27.600	103	12.710	13.344
		15	11.730	26.200	103	9.400	13.344
		20	14.440	25.900	97	12.110	13.764
3	Kane, 2002	0	2.900	24.280	102	0.000	0.000
		15	15.500	26.490	99	12.600	12.038
		30	11.400	22.900	100	8.500	11.977

Study-specific analysis

Mean differences

$$\boldsymbol{d_1} = (2.93, 5.30, 6.00)^\top$$

$$\text{Cov}(\boldsymbol{d_1}) = \begin{bmatrix} 7.59 & & \\ 3.94 & 7.72 & \\ 3.94 & 3.94 & 7.52 \end{bmatrix}$$

f = restricted cubic spline model with knots located at 0, 10, and 30 mg/day

$$\hat{\theta}_1 = (1.215, -5.738)^T$$

$$\hat{\boldsymbol{V}}_1 = \text{Cov}\left(\hat{\theta}_1\right) = \begin{bmatrix} 0.49 & \\ -3.65 & 31.64 \end{bmatrix}$$

Study-specific analysis

Mean differences

$$\boldsymbol{d_1} = (2.93, 5.30, 6.00)^\top$$

$$\text{Cov}\left(\boldsymbol{d_1}\right) = \begin{bmatrix} 7.59 & & \\ 3.94 & 7.72 & \\ 3.94 & 3.94 & 7.52 \end{bmatrix}$$

f = restricted cubic spline model with knots located at 0, 10, and 30 mg/day

$$\hat{\boldsymbol{\theta}}_1 = (1.215, -5.738)^T$$

$$\hat{\boldsymbol{V}}_1 = \text{Cov}\left(\hat{\boldsymbol{\theta}}_1\right) = \begin{bmatrix} 0.49 & \\ -3.65 & 31.64 \end{bmatrix}$$

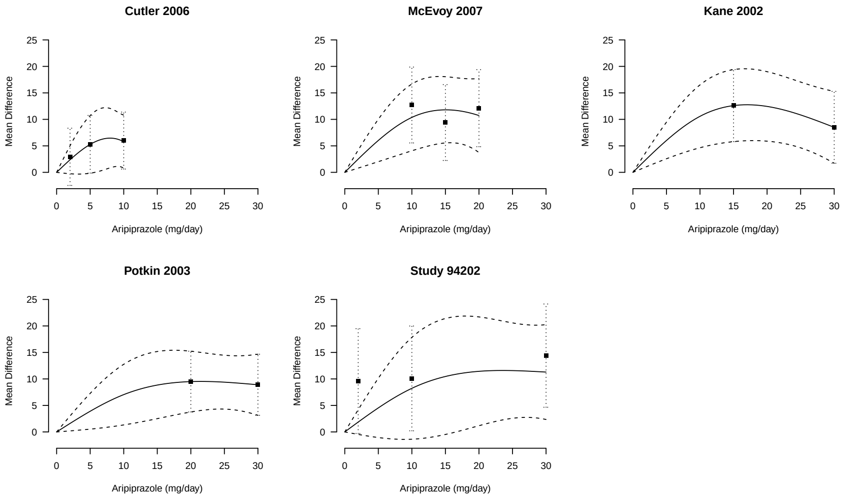

Figure 1: Study-specific predicted dose-response curves and modeled data.

Meta-analysis

Table 3: Study-specific dose-response coefficients and corresponding covariances.

ID	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{Var}\left(\hat{\theta}_1\right)$	$\text{Cov}\left(\hat{\theta}_1, \hat{\theta}_2\right)$	$\text{Var}\left(\hat{\theta}_2\right)$
1	1.22	-5.74	0.49	-3.65	31.64
2	1.26	-2.00	0.19	-0.48	1.43
3	1.25	-1.74	0.14	-0.22	0.40
4	0.81	-0.92	0.13	-0.23	0.44
5	0.93	-1.01	0.36	-0.58	1.02

REML pooled estimates

$$\hat{\theta} = (0.937, -1.156)^T$$

$$\widehat{\text{Cov}}\left(\hat{\theta}\right) = \begin{bmatrix} 0.03 & \\ -0.05 & 0.10 \end{bmatrix}$$

Meta-analysis

Table 3: Study-specific dose-response coefficients and corresponding covariances.

ID	$\hat{\theta}_1$	$\hat{\theta}_2$	$\text{Var}\left(\hat{\theta}_1\right)$	$\text{Cov}\left(\hat{\theta}_1, \hat{\theta}_2\right)$	$\text{Var}\left(\hat{\theta}_2\right)$
1	1.22	-5.74	0.49	-3.65	31.64
2	1.26	-2.00	0.19	-0.48	1.43
3	1.25	-1.74	0.14	-0.22	0.40
4	0.81	-0.92	0.13	-0.23	0.44
5	0.93	-1.01	0.36	-0.58	1.02

REML pooled estimates

$$\hat{\boldsymbol{\theta}} = (0.937, -1.156)^T$$

$$\widehat{\text{Cov}}\left(\hat{\boldsymbol{\theta}}\right) = \begin{bmatrix} 0.03 & \\ -0.05 & 0.10 \end{bmatrix}$$

Pooled dose-response results

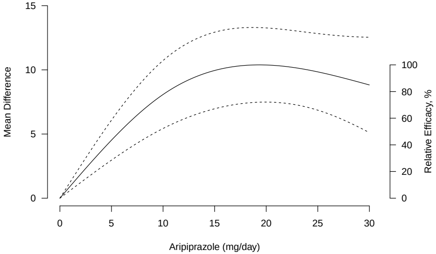

Figure 2: Pooled predicted dose-response curve.

- non-linear association ($p \leq 0.001$)
- E_{max} 10.39 (95% CI 7.48, 13.30) at $x_{\text{max}} = 19.32$ mg/day
- no change after 20 mg/day
- $\widehat{\text{ED}}_{80} = 10.43$ mg/day (95% CI: 9.02, 16.73)
- no heterogeneity ($Q = 3.505$, $p = 0.9$)

Pooled dose-response results

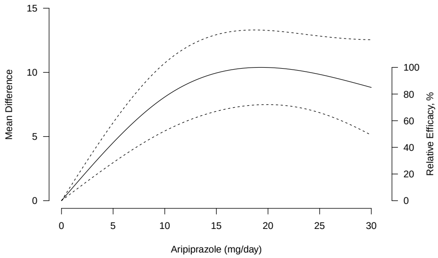

Figure 2: Pooled predicted dose-response curve.

- non-linear association ($p \leq 0.001$)
- E_{max} 10.39 (95% CI 7.48, 13.30) at $x_{max} = 19.32$ mg/day
- no change after 20 mg/day
- $\widehat{\mathrm{ED}}_{80} = 10.43$ mg/day (95% CI: 9.02, 16.73)
- no heterogeneity ($Q = 3.505, p = 0.9$)

Pooled dose-response results

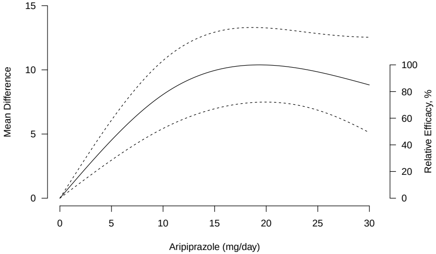

Figure 2: Pooled predicted dose-response curve.

- non-linear association ($p \leq 0.001$)
- $E_{\max}$ 10.39 (95% CI 7.48, 13.30) at $x_{\max} = 19.32$ mg/day
- no change after 20 mg/day
- $\widehat{\mathrm{ED}}_{80} = 10.43$ mg/day (95% CI: 9.02, 16.73)
- no heterogeneity ($Q = 3.505, p = 0.9$)

Pooled dose-response results

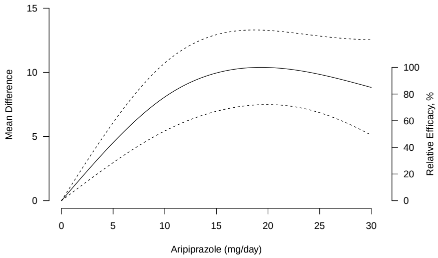

Figure 2: Pooled predicted dose-response curve.

- non-linear association ($p \leq 0.001$)
- E_{max} 10.39 (95% CI 7.48, 13.30) at $x_{max} = 19.32$ mg/day
- no change after 20 mg/day
- $\widehat{\mathrm{ED}}_{80} = 10.43$ mg/day (95% CI: 9.02, 16.73)
- no heterogeneity ($Q = 3.505, p = 0.9$)

Sensitivity analysis

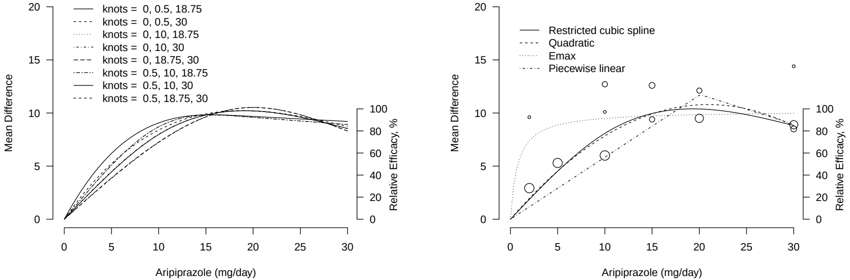

Figure 3: Graphical sensitivity analysis.
Left panel: different location of the three knots.
Right panel: alternative models.

Conclusions

- Two-stage procedure
- Variety of dose-response models
- Spline regression
- Extension to observational studies
- Method implemented in `dosresmeta` R package

References

- Bretz F. et al "Combining Multiple Comparisons and Modeling Techniques in Dose-Response Studies." Biometrics 2005
- Davis et al. "Dose response and dose equivalence of antipsychotics." Journal of clinical psychopharmacology 2004
- Desquilbet L. et al. "Dose-response analyses using restricted cubic spline functions in public health research." Statistics in medicine 2010
- Gasparrini, A. et al. "Multivariate meta-analysis for non-linear and other multi-parameter associations." Statistics in Medicine 2012.