Exploring heterogeneity in dose-response meta-analysis

Nicola Orsini, PhD

Department of Global Public Health
Karolinska Institutet

12th Northern European Stata Conference 2021

September 3, 2021

Outline

- Examples and challenges
- Weighted mixed-effects model using `drmeta`
- Quantiles of marginal and conditional dose-response
- New post-estimation command `drmeta_het`
- Interactive visualizations using Plotly Python
- Applications to simulated and real data
- Final remarks

Blood Pressure Effects of Sodium Reduction: Dose–Response Meta-Analysis of Experimental Studies. Circulation 2021

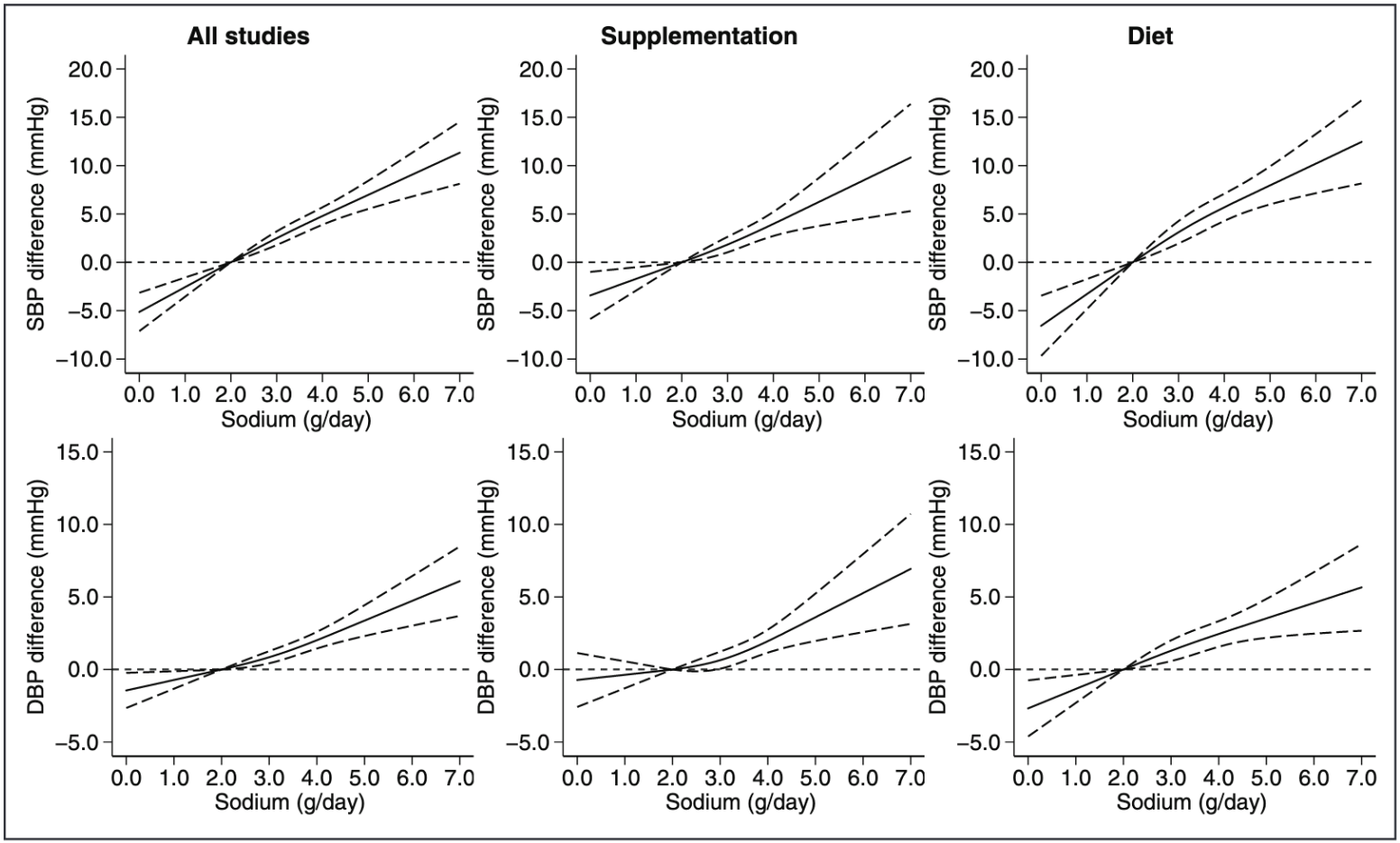
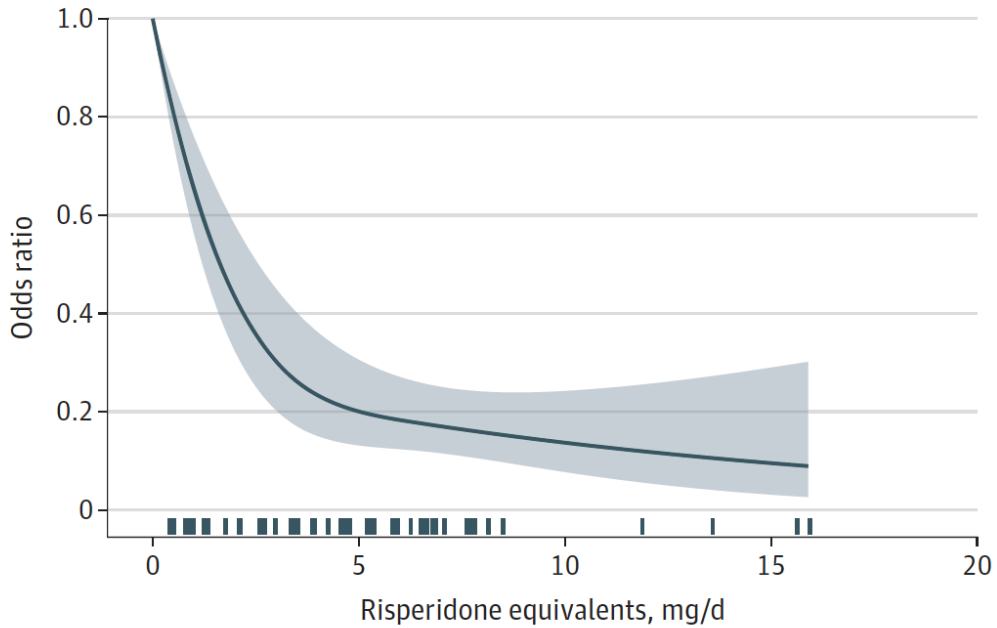


Figure 2. Dose–response meta-analysis of changes in SBP and DBP levels (mm Hg) according to achieved sodium excretion in the treatment and control groups at the end of the trials (all studies) and by type of intervention (supplementation or diet).
The average curve (solid line) with 95% confidence limits (dashed lines) was estimated with a 1-stage random-effects restricted cubic spline model, using 2 g/d as referent. DBP indicates diastolic blood pressure; and SBP, systolic blood pressure.

Examination of Dosing of Antipsychotic Drugs for Relapse Prevention in Patients With Stable Schizophrenia. JAMA Psychiatry 2021

Figure 1. Relapse



The dose-response curve for the primary outcome relapse after pooling all drugs using the primary scientific dose-equivalence method (the maximum effective dose method). The marks on the x-axis indicate for which doses data from study arms were available. A total of 26 studies with 71 individual dose arms including 4749 patients were included (1 publication reported on 2 studies).[27,32-55] The shaded areas indicate 95% CIs for the primary outcome.

Weighted Mixed Effects Model

A one-stage approach for meta-analysis of summarized dose-response data has been proposed in the general framework of linear mixed effects model (*Stat Meth Med Res*, 2019).

$$\hat{\gamma}_i = \boldsymbol{X}_i\boldsymbol{\beta} + \boldsymbol{Z}_i \mathrm{b}_i + \boldsymbol{\epsilon}_i$$

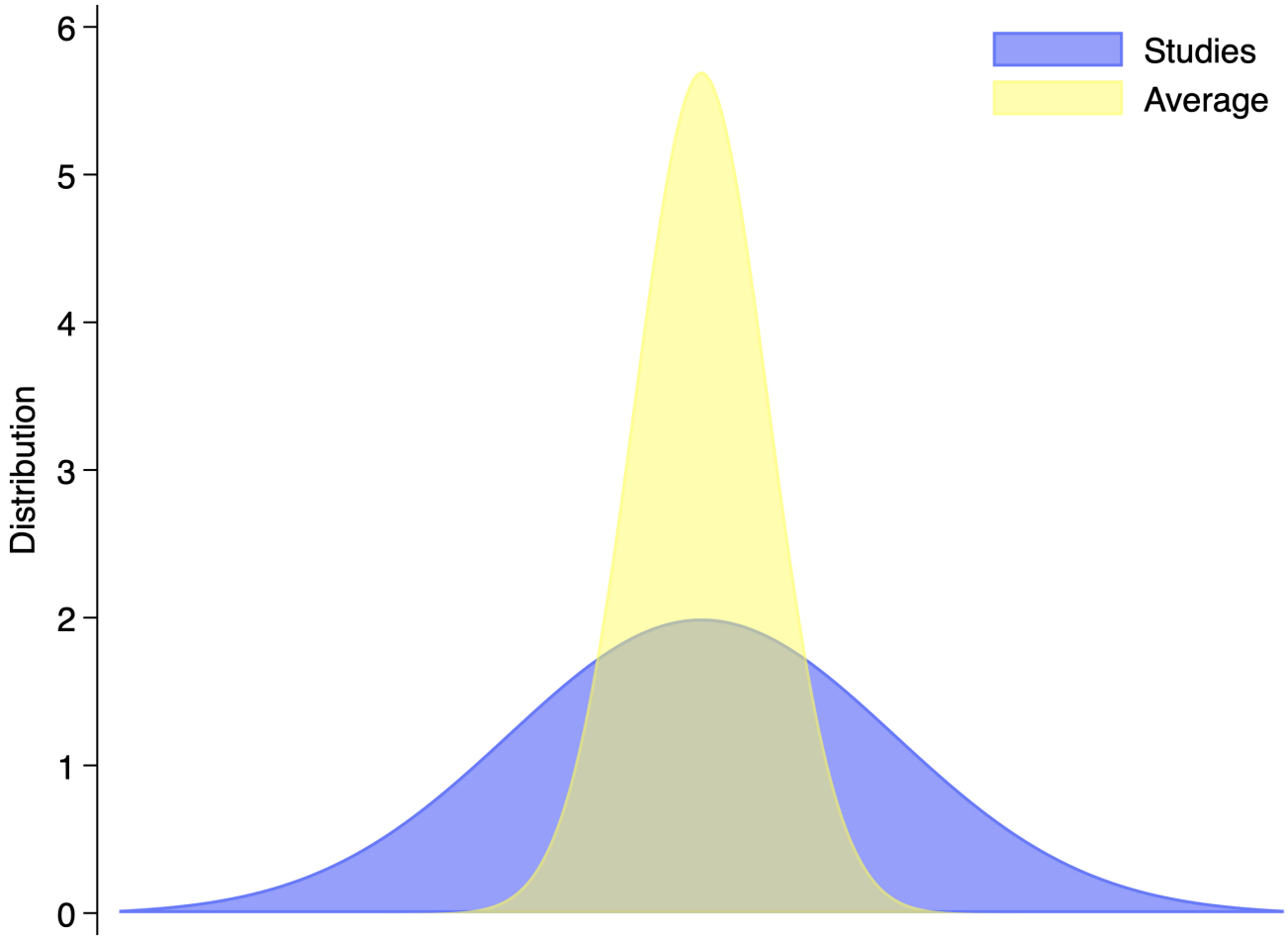
$$\mathrm{b}_i \sim \mathcal{N}(0, \Psi)$$
$$\epsilon_i \sim \mathcal{N}(0, \mathrm{S}_i)$$

$\hat{\gamma}_i$ is the vector of empirical constrasts (i.e. mean differences, log odds ratios, log hazard ratios) estimated relative to a common referent in the i-th study

It is implemented in the `drmeta` command (*Stata J*, 2021).

A graphical intuition: Average study vs Individual studies

$$\beta_i \sim N(\beta, \tau(\beta_i))$$



Heterogeneity of a specific contrast

$$\beta_i(x - x_0) \sim N(\beta(x - x_0), \tau(\beta_i(x - x_0)))$$

Consider a meta-analysis of I studies of the same size n, equal dose std deviation σ_{X_i}, and equal conditional outcome std deviation σ_{Y_i}

$$\tau(\beta_i(x - x_0)) = \sqrt{(x - x_0)^2(\widehat{SE}(\hat{\beta})^2 + \hat{\tau}^2)}$$

$$\widehat{SE}(\hat{\beta}) = 1/\sqrt{1/(\widehat{SE}(\hat{\beta}_i)^2 + \hat{\tau}^2)I}$$

$$\widehat{SE}(\hat{\beta}_i) = \sigma_{Y_i}/(\sigma_{X_i}\sqrt{n-1})$$

Quantiles of marginal and conditional dose-response

$$Q_p^C(\beta_i(x - x_0)) = \beta(x - x_0) + \phi^{-1}(p)\sqrt{(x - x_0)^2(\widehat{SE}(\hat{\beta})^2 + \hat{\tau}^2)}$$

$$Q_p^M(\beta_i(x - x_0)) = \beta(x - x_0) + \phi^{-1}(p)\sqrt{(x - x_0)^2\widehat{SE}(\hat{\beta})^2}$$

$\phi^{-1}(p)$ is the p-quantile of a standard normal distribution

$Q_{0.5}^M = Q_{0.5}^C$ because $\phi^{-1}(0.5) = 0$

If $\tau^2 > 0$, then $|Q_p^C| > |Q_p^M|$

Interpretation of Q_p^M and Q_p^C, however, is always different.

The post-estimation

$$Q_p^C(\beta_i(x - x_0)) = \beta(x - x_0) + \phi^{-1}(p)\sqrt{(x - x_0)^2(\widehat{SE}(\hat{\beta})^2 + \hat{\tau}^2)}$$

$$Q_p^M(\beta_i(x - x_0)) = \beta(x - x_0) + \phi^{-1}(p)\sqrt{(x - x_0)^2\widehat{SE}(\hat{\beta})^2}$$

$\phi^{-1}(p)$ is the p-quantile of a standard normal distribution

$Q_{0.5}^M = Q_{0.5}^C$ because $\phi^{-1}(0.5) = 0$

If $\tau^2 > 0$, then $|Q_p^C| > |Q_p^M|$

Interpretation of Q_p^M and Q_p^C, however, is always different.

Analytical example

The random-effect linear dose-response mechanism is $\beta_i \sim N(0.5, 0.2)$. Consider $I = 10$ studies of the same size $n = 1000$, equal dose distribution $X \sim \chi 2(5)$, and equal conditional outcome std deviation $\sigma_{Y_i} = 10$. Using a dose of 5 units as referent we have that

$$\beta_i(x-5) \sim N(0.5(x-5), 0.2(x-5)))$$

the standard error of the slope in any similar study would be

$$\widehat{SE}(\hat{\beta}_i) = 10/(\sqrt{5(2)}\sqrt{1000-1}) = 0.1$$

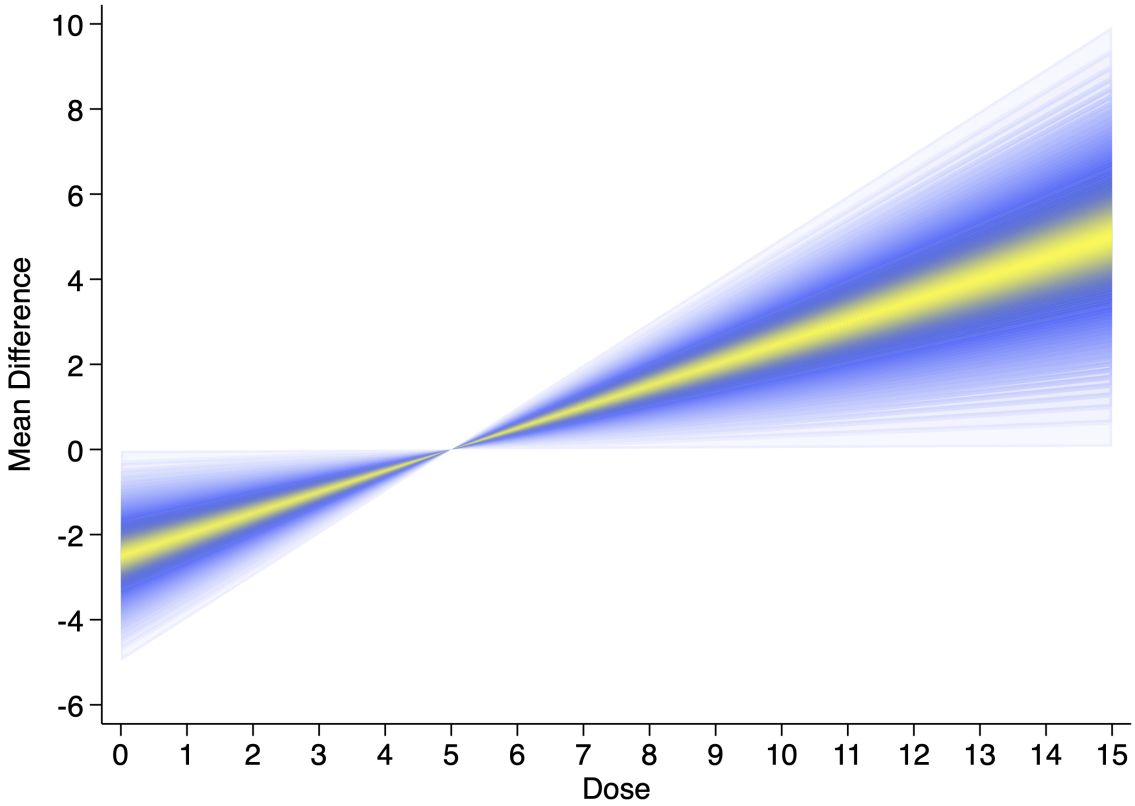
and the standard error of the slope for the average study would be

$$\widehat{SE}(\hat{\beta}) = 1/\sqrt{1/(0.1^2+0.2^2)10} = 0.07$$

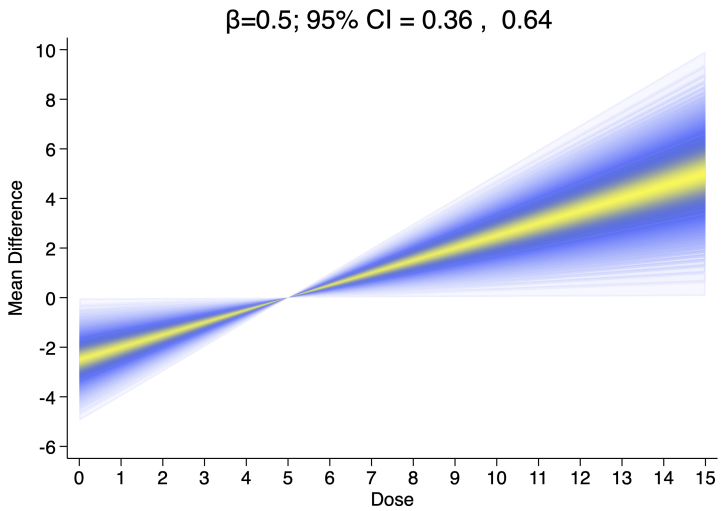
Marginal vs Conditional Quantiles

$$Q_p^C(\beta_i(x-5)) = 0.5(x-5) + \phi^{-1}(p)\sqrt{(0.07^2 + 0.2^2)(x-5)^2}$$

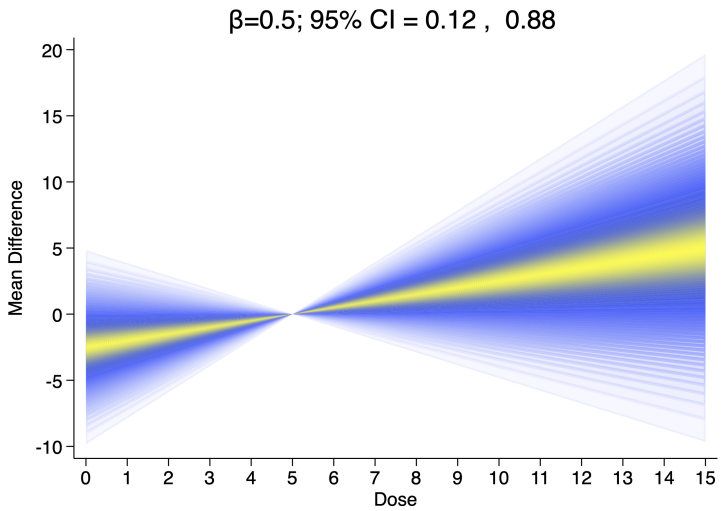
$$Q_p^M(\beta(x-5)) = 0.5(x-5) + \phi^{-1}(p)\sqrt{(0.07^2)(x-5)^2}$$



And so what?



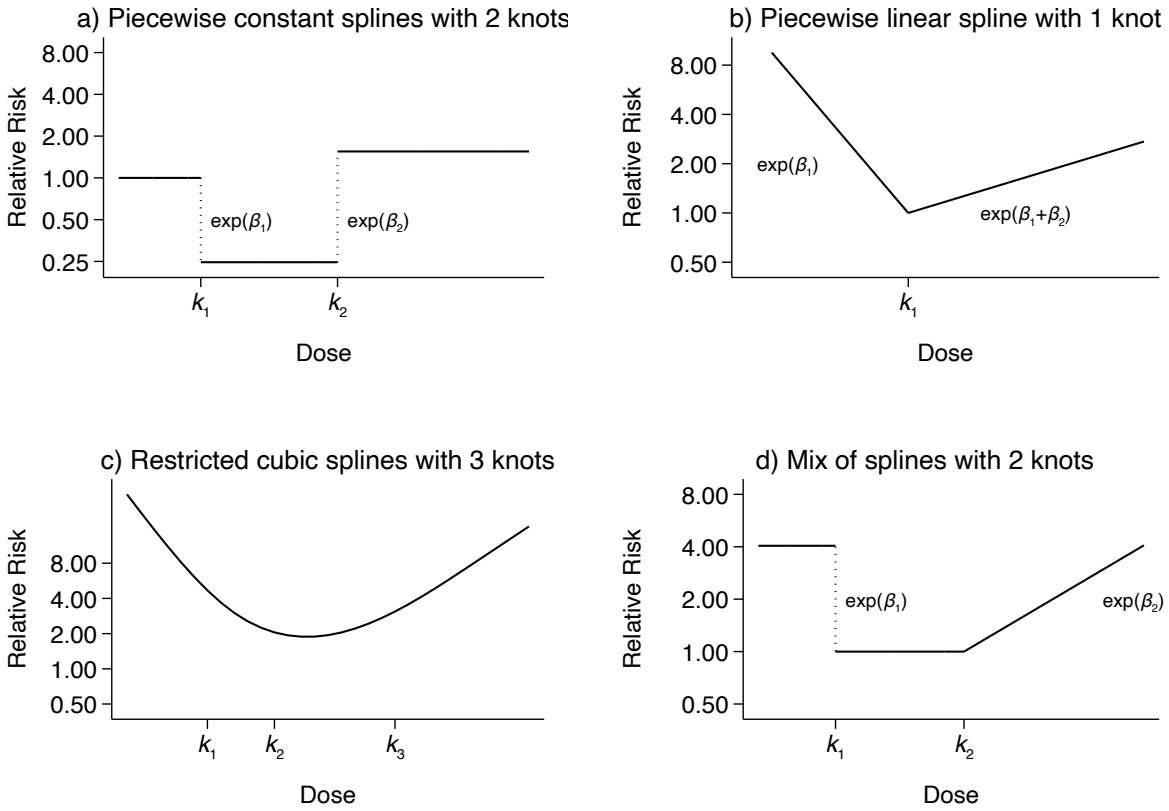
(a) $\tau = 0.2$



(b) $\tau = 0.6$

Figure: An interesting case is when the inference based on conditional quantiles is in opposite direction of the inference based on marginal quantiles

Moving beyond linear dose-response relationships



Orsini N, and Spiegelman D. *Meta-Analysis of Dose-Response Relationships*. Chapter 18. Handbook of Meta-Analysis. Ed. Schmid CH, Stijnen T, White, I. 2020. CRC Press.

Extend the reasoning

Let's consider two transformations (i.e. splines, fractional polynomials), saying $f_1(x)$ and $f_2(x)$, of the original dose.

$$\beta_{1i}(f_1(x) - f_2(x_0)) + \beta_{2i}(f_2(x) - f_2(x_0))$$

$$\begin{pmatrix} \beta_{1i} \\ \beta_{2i} \end{pmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}, \begin{bmatrix} \xi_1 & \\ \xi_3 & \xi_2 \end{bmatrix} \right)$$

The model is in terms of 2 fixed-effects plus 2 variances and 1 covariance of the 2 random-effects.

At this point, it helps to use a compact matrix notation

$$\boldsymbol{\beta_i} \sim \mathcal{N}(\boldsymbol{\beta}, \Psi)$$

Quantiles for the marginal and conditional dose-response relationship

Marginal

$$Q_p^M = (\boldsymbol{X}^* - \boldsymbol{x}_0^*)\hat{\boldsymbol{\beta}} + \phi^{-1}(p)\mathrm{diag}[(\boldsymbol{X}^* - \boldsymbol{x}_0^*)V(\hat{\boldsymbol{\beta}})(\boldsymbol{X}^* - \boldsymbol{x}_0^*)']^{1/2}$$

Conditional

$$Q_p^C = (\boldsymbol{X}^* - \boldsymbol{x}_0^*)\hat{\boldsymbol{\beta}} + \phi^{-1}(p)\mathrm{diag}[(\boldsymbol{X}^* - \boldsymbol{x}_0^*)(V(\hat{\boldsymbol{\beta}}) + \hat{\Psi})(\boldsymbol{X}^* - \boldsymbol{x}_0^*)']^{1/2}$$

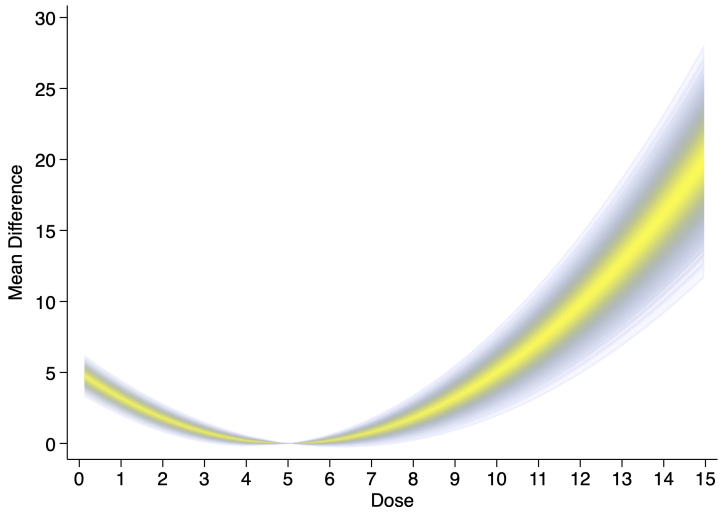
where

$\boldsymbol{X}^*$ indicates a matrix of user specified transformations

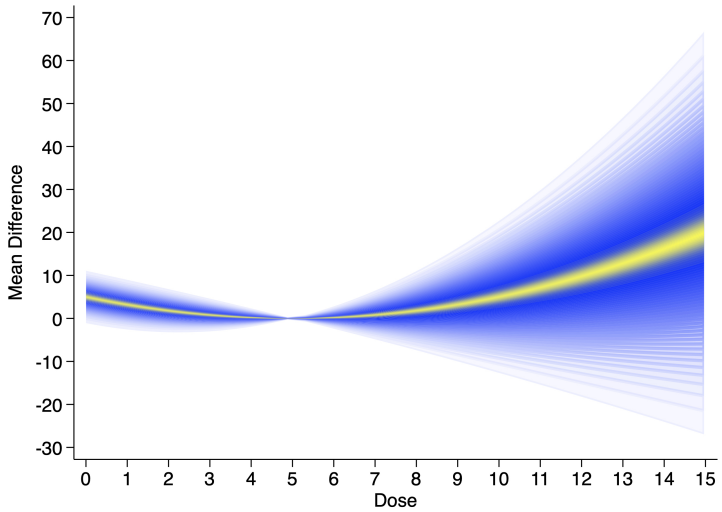
$\boldsymbol{x}_0^*$ indicates a matrix of reference values

Small and large non-linear heterogeneity

$$MD = -2(x-5) + 0.2(x^2 - 5^2)$$



(a) $\xi_1 = 0.0001$, $\xi_2 = 0.0001$, $\xi_3 = 0$



(b) $\xi_1 = 0.01$, $\xi_2 = 0.01$, $\xi_3 = 0$

Figure: Apparently small differences in variance components can lead to large heterogeneity for extreme comparisons

New post-estimation command `drmeta_het`

Ideally, it would be great to have a post-estimation command that

- works with a variety of dose transformations and outcome measures
- allows the user to choose between quantile of the conditional, marginal, or both
- allows the user to overlay the study-specific BLUPs
- easily provides both static and interactive visualizations

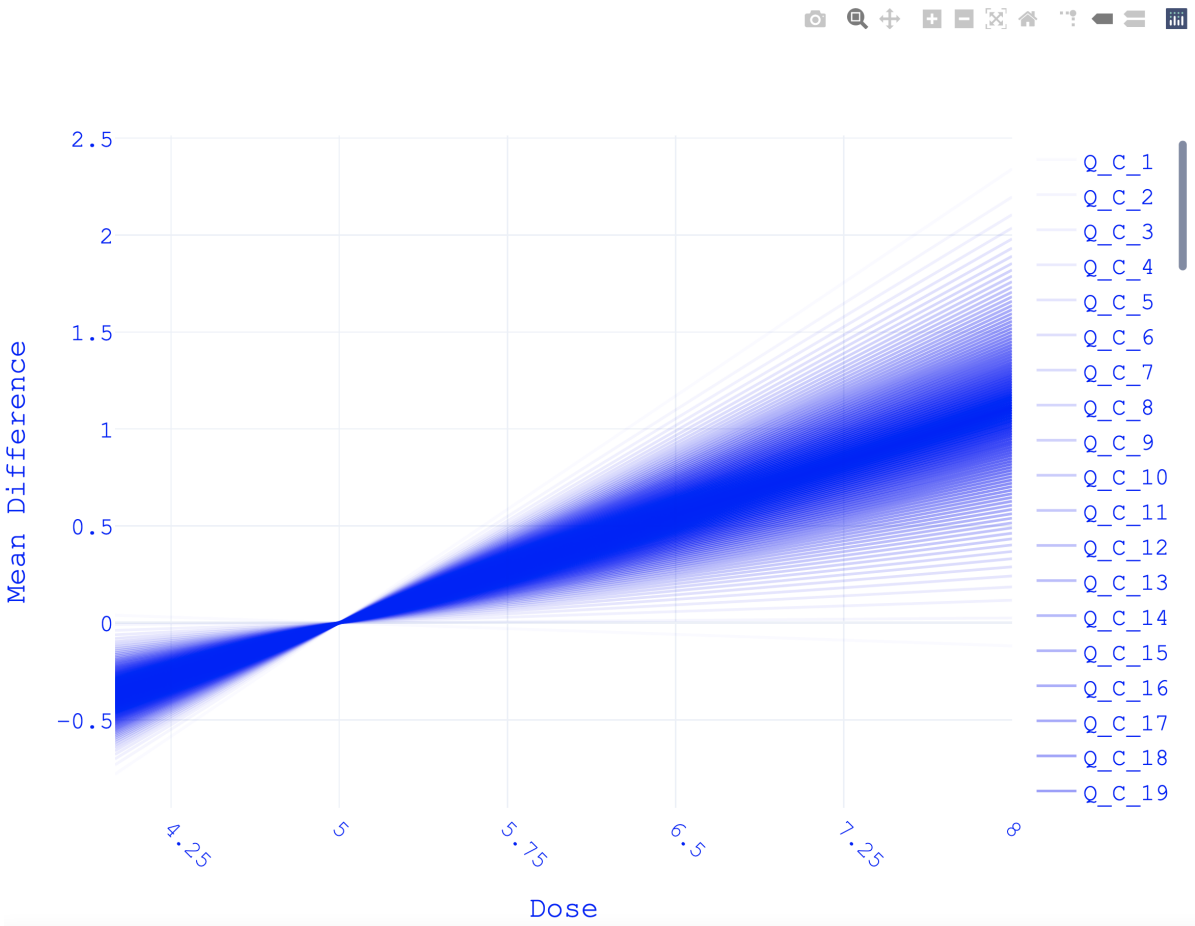
So I wrote `drmeta_het` using Plotly Python Graphing Library taking advantage of the recent Stata/Python integration.

Simulated Example

	id	md	dose	semd	n	sd
1.	1	0	2.020252	0	334	9.458647
2.	1	-.6599554	4.426369	.7682979	333	10.36173
3.	1	.9591057	8.731555	.7440263	333	9.754083
4.	2	-3.207253	1.784821	.8826443	250	9.898886
5.	2	-1.717937	3.514592	.8673572	250	9.555094
6.	2	0	5.421488	0	250	9.837545
7.	2	1.940711	9.397936	.8920666	250	10.10784
8.	3	0	1.801217	0	250	9.401782
9.	3	.627276	3.494548	.8673193	250	9.983344
10.	3	2.376793	5.359765	.862201	250	9.871885
11.	3	2.563129	9.262744	.8850933	250	10.366
12.	4	-.3530243	2.018016	.7970039	334	9.713749
13.	4	0	4.214327	0	333	10.83753
14.	4	.1312752	8.515213	.8070071	333	9.970876

Syntax of drmeta_het #1

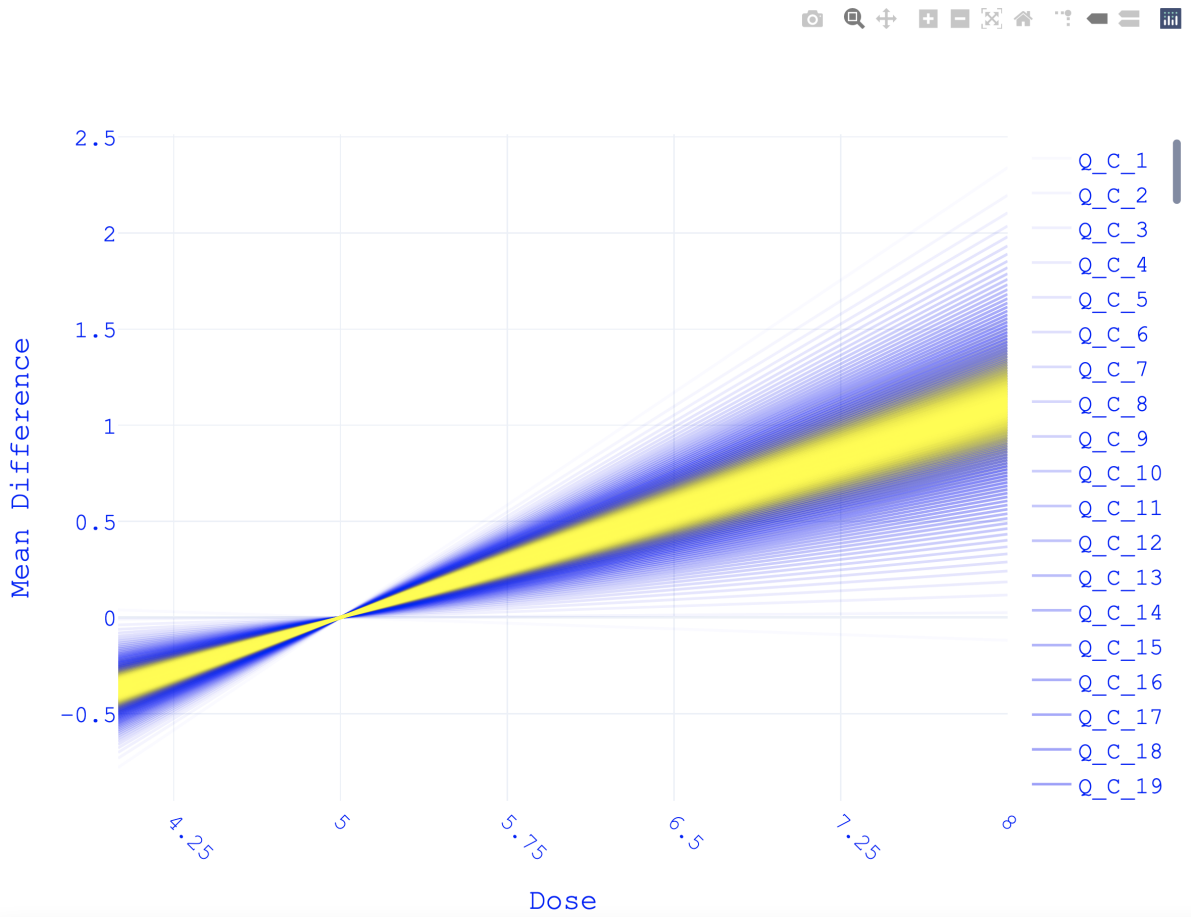
```
drmeta_het , dose(4(.5)8) ref(5) eq(d) iqc
```



Click Here

Syntax of drmeta_het #2

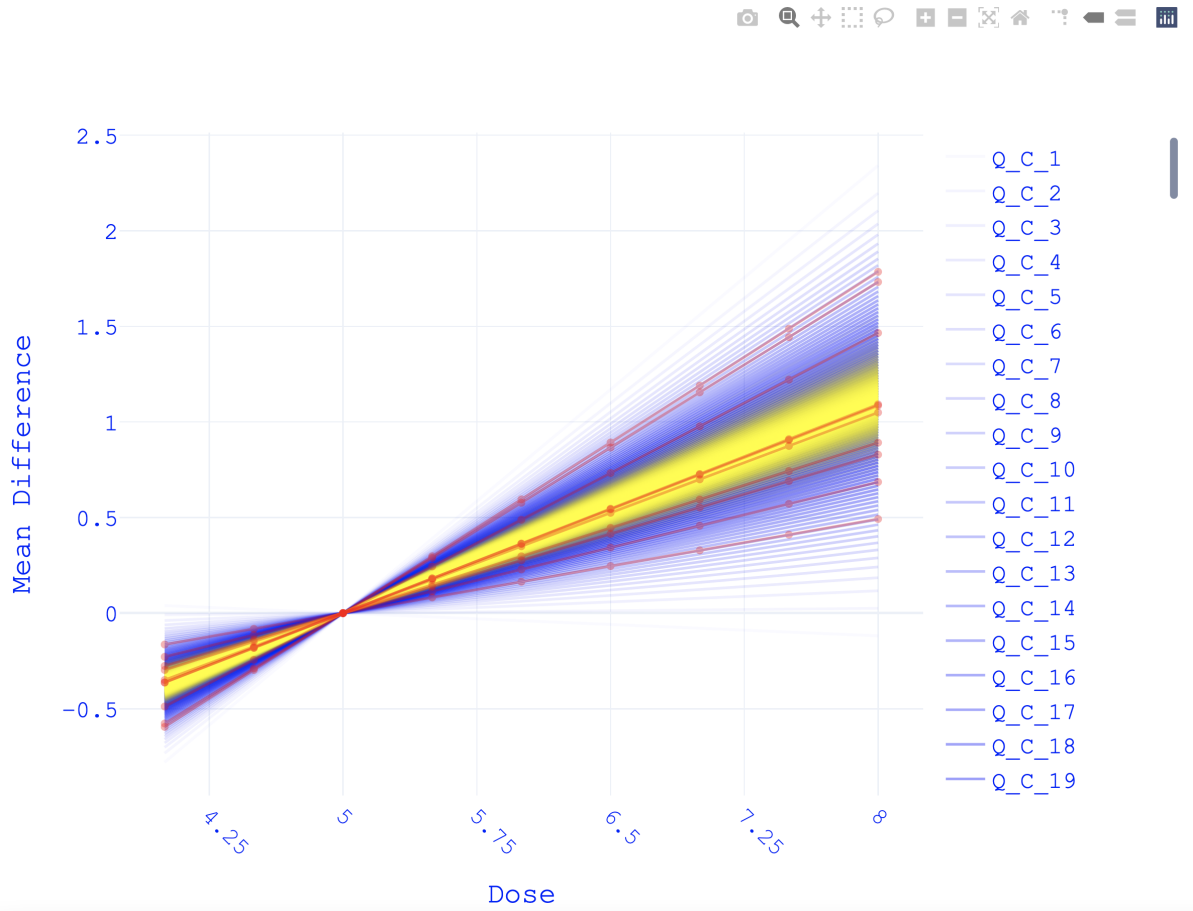
```
drmeta_het , dose(4(.5)8) ref(5) eq(d) iqc iqm
```



Click Here

Syntax of drmeta_het #3

```
drmeta_het , dose(4(.5)8) ref(5) eq(d) iqc iqm iqcb
```



Click Here

Simulated Example: Walking and mortality

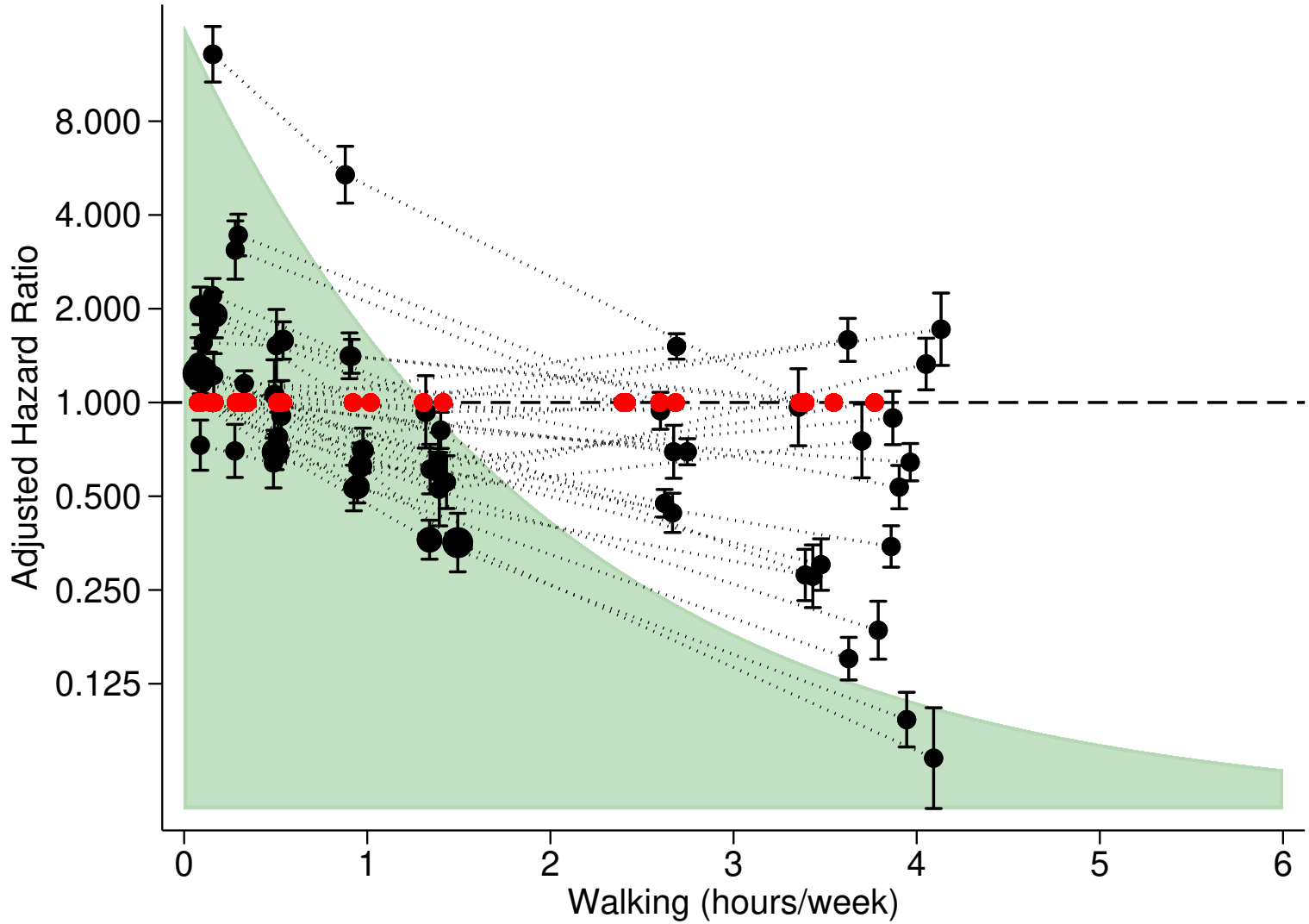
- Consider 30 prospective cohort studies investigating the association between baseline walking, measured in hours/week, and time until death, or end of follow-up (10 years), whichever came first.
- Age is inversely associated with walking levels and positively associated with higher mortality rates independently of walking levels.
- The true summary age-adjusted mortality hazard ratio is decreasing with higher walking levels with a threshold effect at 2 hours per week

$$HR = e^{-0.5(x-2)+0.5(x>2)(x-2)}$$

Snapshot of the aggregated data

id	walk	b	seb	case	py
1	0.3	1.13	0.11	229	777
1	2.4	0.00	0.00	137	1704
20	0.1	0.21	0.10	239	674
20	0.5	0.00	0.00	216	946
20	1.5	-1.04	0.11	133	1773
20	4.1	-2.63	0.19	32	2318
23	0.2	0.65	0.09	311	973
23	0.9	0.00	0.00	247	1765
23	3.4	-1.28	0.12	101	2752

Plotting the empirical contrasts



Piecewise linear weighted mixed-effects model

We specify a dose-response model with constant change for the age-adjusted log mortality hazard ratio associated with every 1 hour per week increase in walking before and after the knot at 2 hours per week.

$$\hat{\gamma}_{ij} = (\beta_1 + b_{1i})x_{ij} + (\beta_2 + b_{2i})I(x_{ij} > 2)(x_{ij} - 2) + \epsilon_{ij}$$

Stata output

```
. drmeta b walk walkplus, se(seb) data(py case) type(type)  id(id) ml

One-stage random-effects dose-response model     Number of studies =        30
Optimization   = ml                                  Number of obs =        61
           AIC = 37.55                               Model chi2(2) =    110.27
Log likelihood = -13.773298                            Prob > chi2 =    0.0000
```

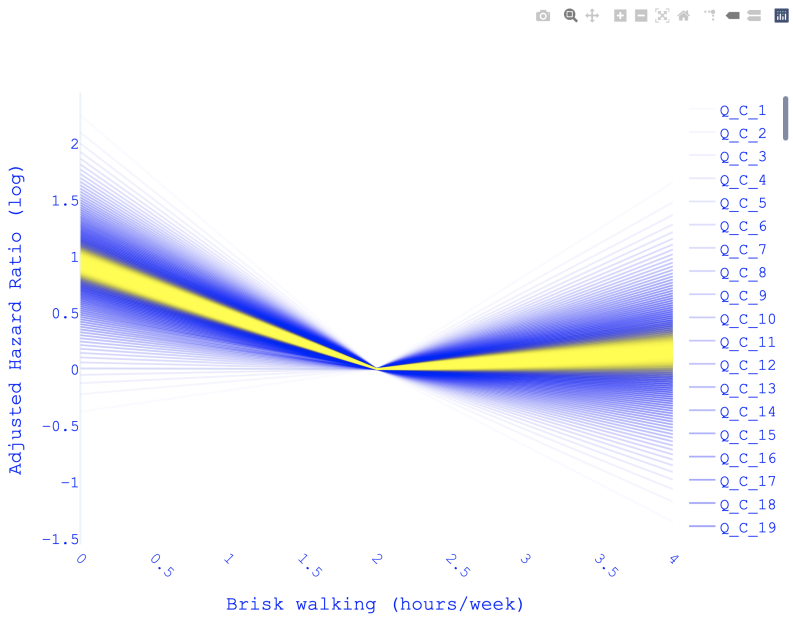
b	Coefficient	Std. err.	z	P>|z|	[95% conf.	interval]
walk	-.4678671	.0536744	-8.72	0.000	-.573067	-.3626673
walkplus	.5432787	.0626324	8.67	0.000	.4205213	.666036

Random-effects parameters	Estimate
var(walk,walk)	.0766958
var(walkplus,walkplus)	.0507463
cov(walk,walkplus)	-.0136841

LR test vs. no random-effects model = **2713.6** Prob >= chi2(3) = **0.0000**

Syntax of drmeta_het

```
drmeta_het ,  eq(d  (d>2)*(d-2) ) dose(0(.1)4)  ///
    ref(2) ///
    yt("Adjusted Hazard Ratio (log)") ///
    xt("Brisk walking (hours/week)") ///
    iqm iqc  iqcbm
```



Click Here

Alcohol intake and colorectal cancer risk

We combine the dose-response relation between alcohol intake and colorectal cancer rate arising from 8 prospective cohort studies including 489,979 women and men participating in the Pooling Project of Prospective Studies of Diet and Cancer. A total of 3,646 cases and 2,511,424 person-years are included in this analysis.

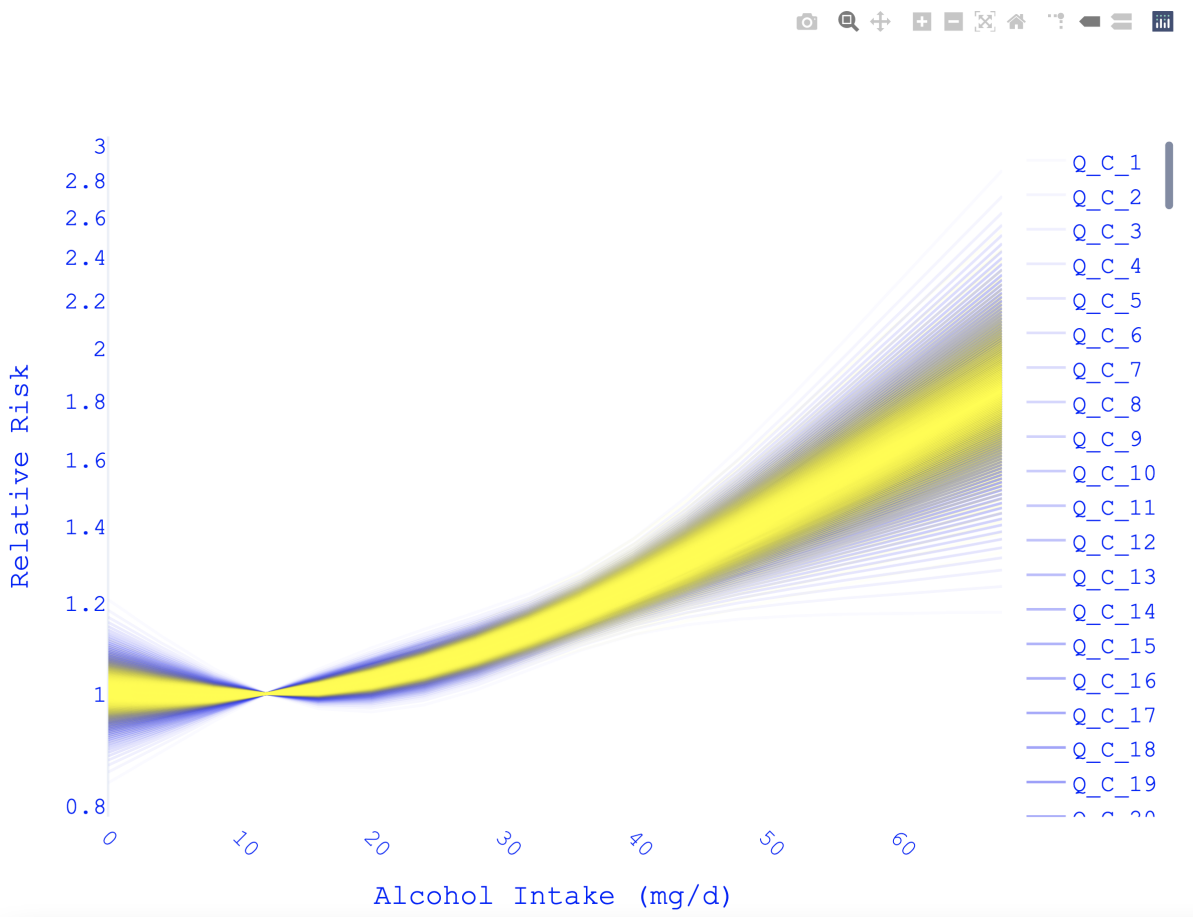
```
use ex_alcohol_crc.dta, clear
* Restricted cubic splines

mkspline doses = dose, nk(3) cubic
mat knots = r(knots)

drmeta logrr doses1 doses2 , data(peryears cases) ///
id(study) type(type) se(se) ml

drmeta_het , dose(0(4)70) ref(12) matk(knots) eform ///
yt("Relative Risk") xt("Alcohol Intake (mg/d)") iqc iqm
```

Alcohol intake and colorectal cancer risk



Click Here

Alcohol intake and colorectal cancer risk

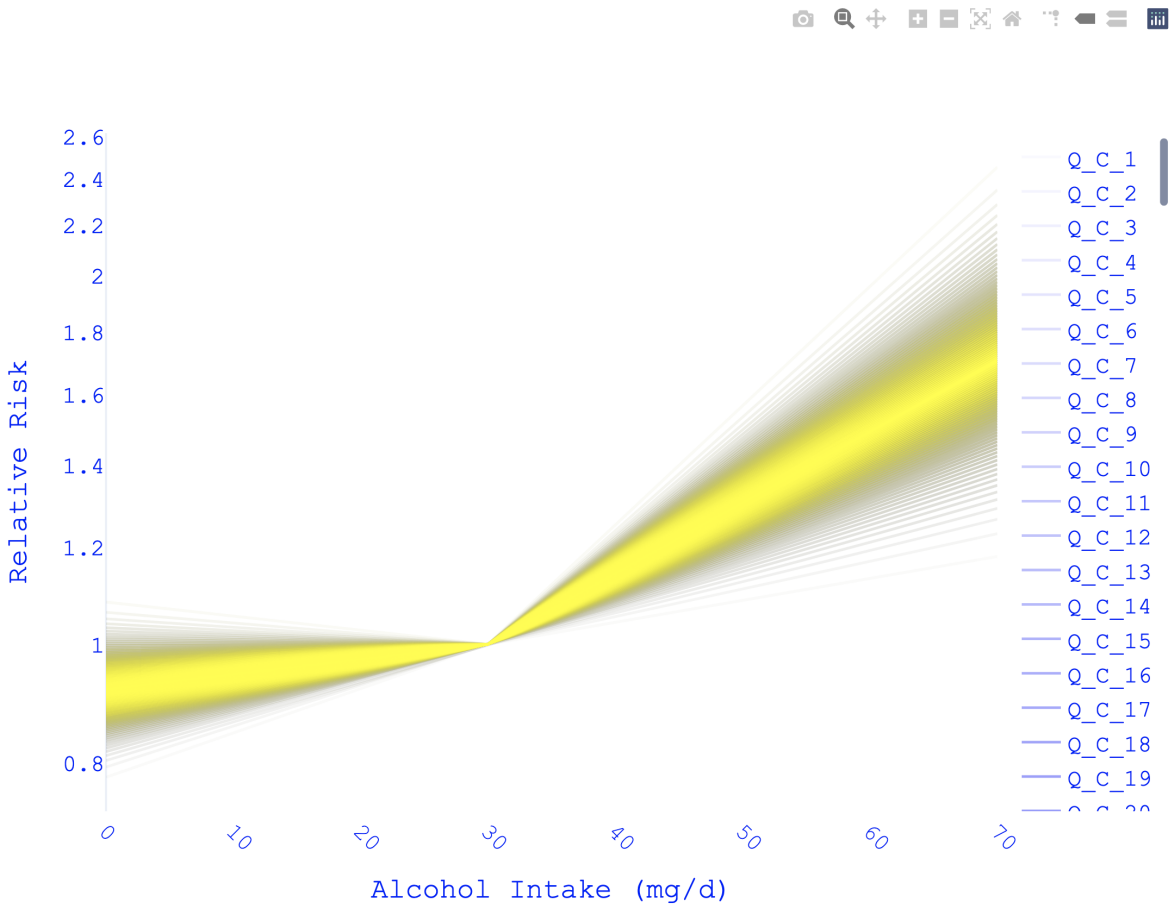
The drmeta_het command works even for piecewise linear splines.

```
gen dose_plus = (dose>30)*(dose-30)^1

drmeta logrr dose dose_plus,  se(se) ///
 data(peryears cases)  id(study) type(type) ml

drmeta_het , dose(0(1)70) ref(30) eform iqc iqm  ///
  eq(d (d>30)*(d-30)) ///
yt("Relative Risk") ///
xt("Alcohol Intake (mg/d)") iqc iqm
```

Alcohol intake and colorectal cancer risk

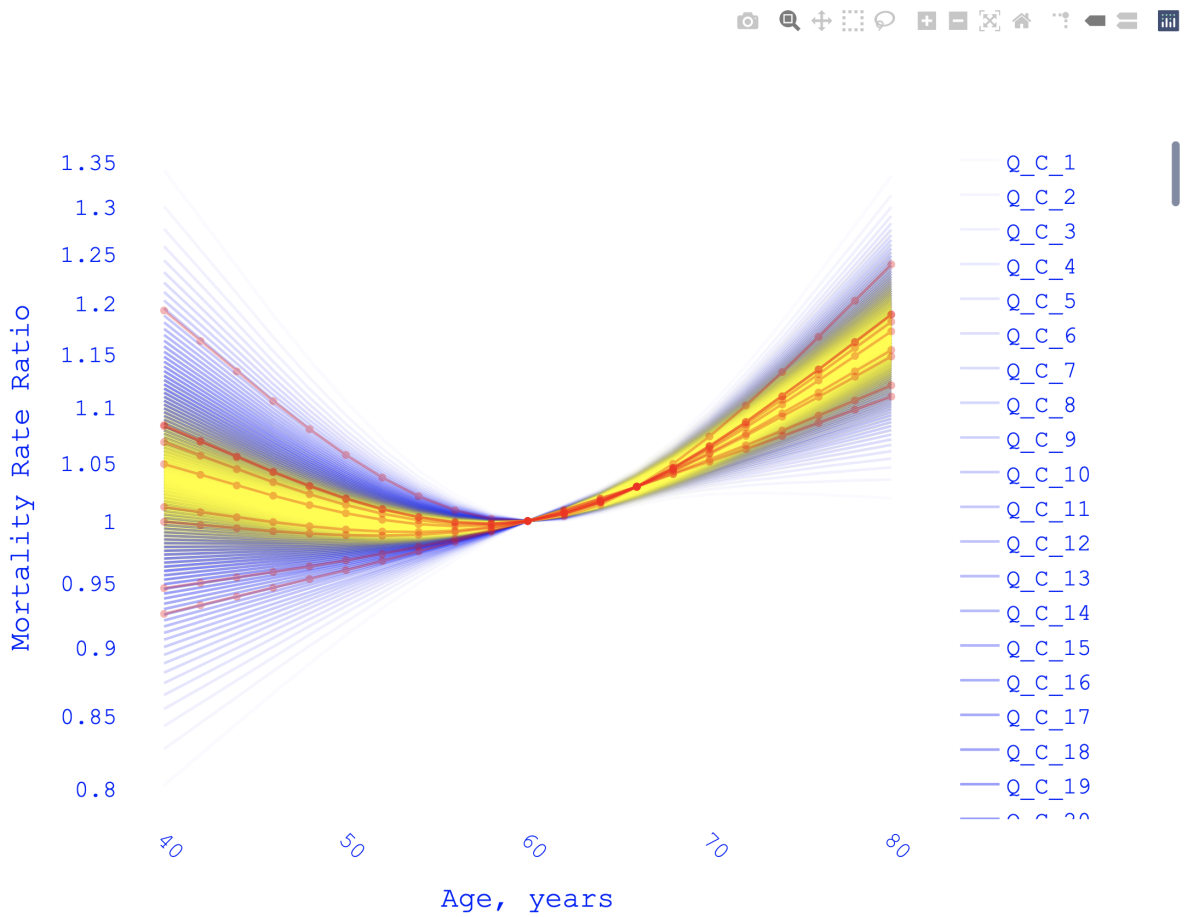


Click Here

Age and and breast cancer mortality

We use data from the Surveillance, Epidemiology, and End Results (SEER) program of the National Cancer Institute. The SEER program provides data about cancer statistics from several population-based registries in the USA (http://seer.cancer.gov) from San Francisco- Oakland, Connecticut, Metropolitan Detroit, Hawaii, Iowa, New Mexico, Seattle, Utah, and Metropolitan Atlanta that here are considered as different studies. Analysis are based on 9 studies on prognostic factors for breast cancer survival including a total of 84,404 women. During 554,812 person-years, 8,520 women died from breast cancer.

Age and and breast cancer mortality



Click Here

Final remarks

- development of `drmeta_het` is still on-going
- The extent of heterogeneity can be explored using quantiles of the conditional and marginal predicted dose-response
- The post-estimation command `drmeta_het` allows the user to explore all the quantiles using interactive visualizations
- A key assumption for deriving quantiles was a) a mixed model and b) the normal distribution of the random-effects

Selected References

- **Orsini N**. Weighted mixed-effects dose-response models for tables of correlated contrasts. *Stata Journal*. 2021. Vol 2.
- **Orsini N**, and Spiegelman D. Meta-Analysis of Dose-Response Relationships. Chapter 18. *Handbook of Meta-Analysis*. Ed. Schmid CH, Stijnen T, White, I. 2020. CRC Press.
- Crippa A, Discacciati A, Bottai M, Spiegelman D, **Orsini N**. One-stage dose-response meta-analysis for aggregated data. *Stat Methods Med Res*. 2019 May;28(5):1579-1596.