Why meta-analytic evidence is anything else than ´rock solid´:

The case of environmental tobacco smoke

Michael Bošnjak
Free University of Bozen-Bolzano, School of Economics and Management

Holger Pahl & Robin Würfel
University of Mannheim, Department of Psychology II

Slides > http://invited-presentations.bosnjak.eu

Agenda

1. Meta-analytic principles and sketch of two common methods (in Psychology)
2. The bone of contention:
   Meta-analysis on the effects of ETS at workplace on lung cancer
   (Stayner et al., 2007)
3. Are meta-analyses rock solid?
   Understanding subjectivity in meta-analytic research processes: The ´crossroads model´ applied to the Stayner data.

Evidence unclear?

	r	*N*	Sig.?	95% CI	
				-	+
Study 1	0,25	30	n.s.	-0,13	0,63
Study 2	-0,18	40	n.s.	-0,50	0,14
Study 3	0,41	50	*	0,12	0,70
Study 4	0,09	60	n.s.	-0,17	0,35
Study 5	0,28	70	*	0,04	0,52
Study 6	0,32	80	*	0,10	0,54
Study 7	0,11	90	n.s.	-0,10	0,32
Study 8	0,31	100	*	0,11	0,51

***r*= .22** (.13/.30); *Q*= 11.45, *df*= 7, *p*= .13

Meta-analysis : Basics

- Generic term, encompassing a comprehensive set of specific research synthesis/analysis techniques and approaches.
- Overall process of systematiclly retrieving, synthesizing, and analyzing the results of thematically related studies.
- Effect sizes are being synthesized and analysed, such as:
  - measures of association ("*r family*")
  - mean differences ("*d family*")
  - ratios and ratio differences
  - Effect sizes for dichotomous measures, such as relative risks, odds ratio and derivatives.

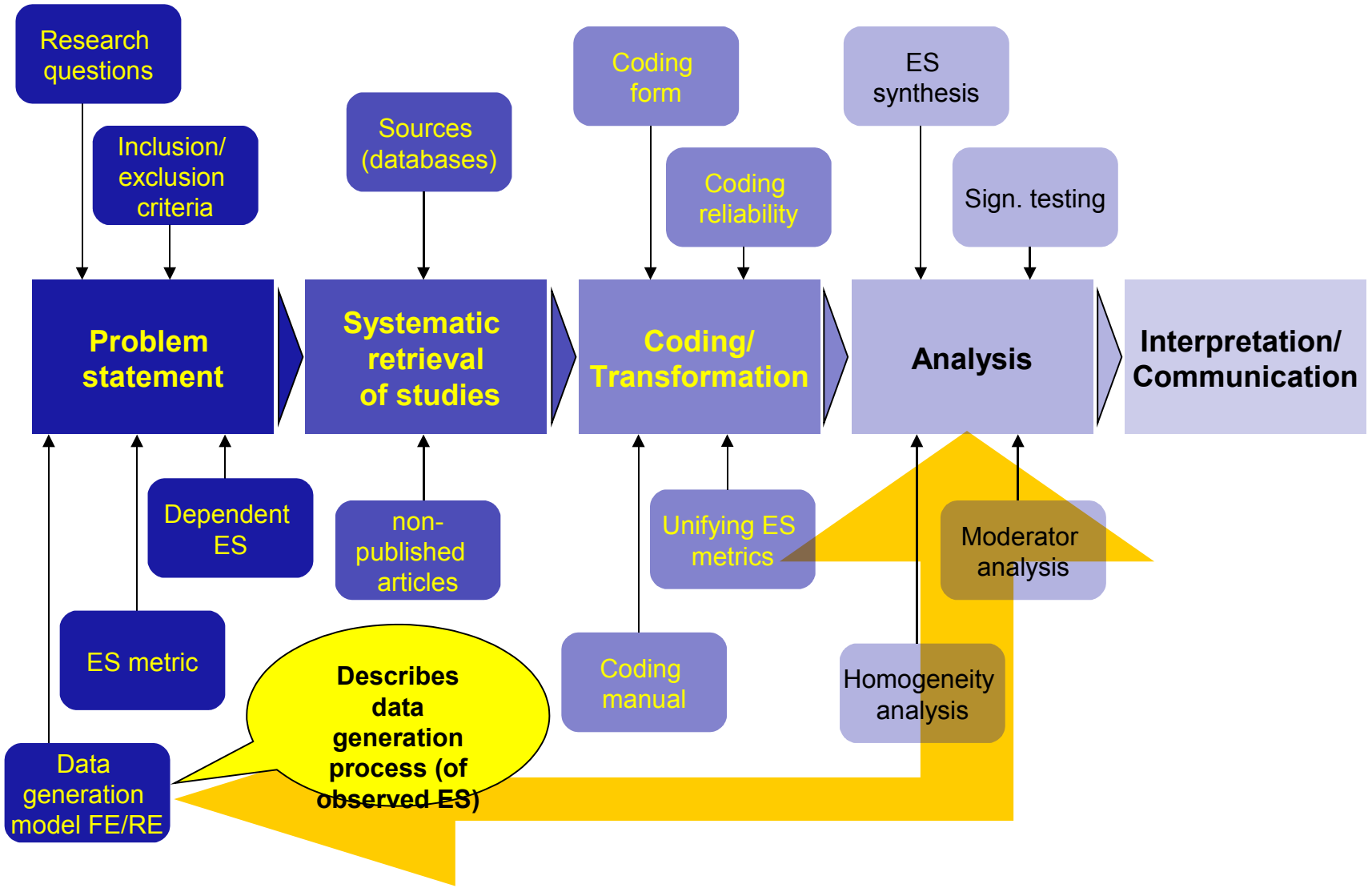
Meta-Analysis : Objectives

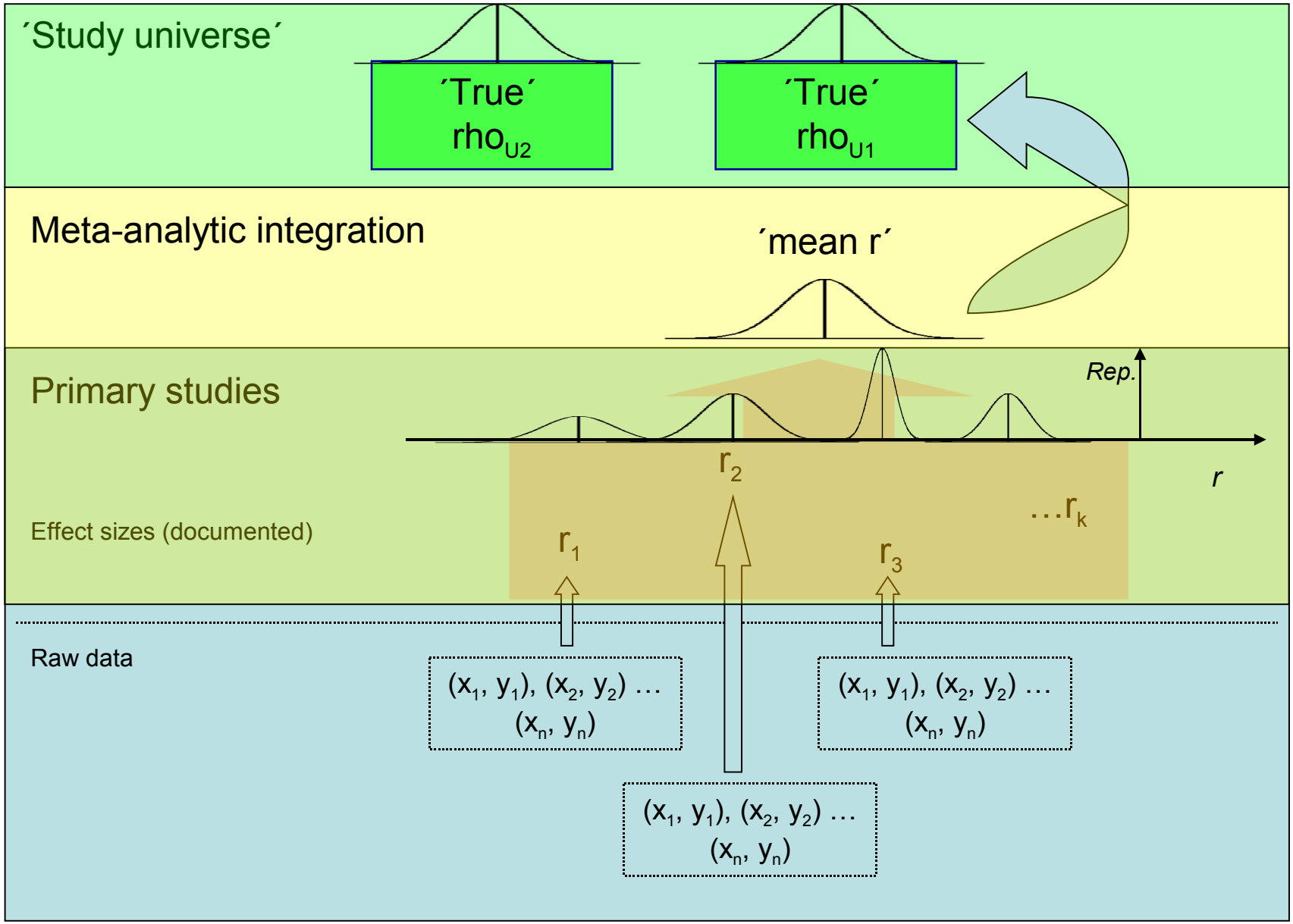
Estimation of a ´true´ effect, characterized by a higher precision and validity compared to primary studies.

Evidence-based ...

- ... description of a thematic area (also to identify gaps in research),
- ... theory testing,
- ... theory development/refinement,
- ... conclusions on the benefit / harm of an intervention.

Generic Procedure




´Study universe´
´True´
rho$_{U2}$
´True´
rho$_{U1}$
Meta-analytic integration
´mean r´
Primary studies
Rep.
r
r$_2$
…r$_k$
r$_1$
r$_3$
Effect sizes (documented)
Raw data
(x$_1$, y$_1$), (x$_2$, y$_2$) …
(x$_n$, y$_n$)
(x$_1$, y$_1$), (x$_2$, y$_2$) …
(x$_n$, y$_n$)
(x$_1$, y$_1$), (x$_2$, y$_2$) …
(x$_n$, y$_n$)

Two Common Approaches (in Psychology)

- "HO"-Meta-Analysis (Hedges/Olkin)
  - Data generation model: True effects are confounded by (up to two different types of) sampling error
  - Approx. 75% of all meta-analyses published in *Psychological Bulletin are HO meta-analyses* (fixed effects model)

- Psychometric MA (Hunter/Schmidt)
  - Data generation model : ´True´effects are attenuated by sampling error and systematic artifacts
  - Approx. 80% of all meta-analyses in IO Psychology

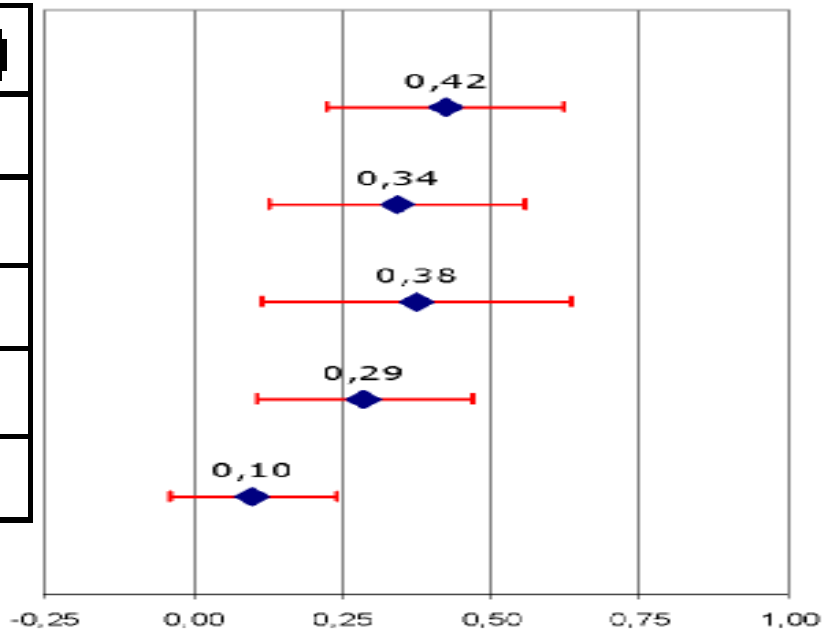
- No(t) (more) popular in Psychology:
  Glass meta-analysis, p-value aggregation according to Rosenthal, Bayesian approaches

HO-Meta-Analysis: Synthesis & Analysis

Step 1:
Z[r] Transformation

$$ES_i = Zr_i = \frac{1}{2}\ln\left[\frac{1+r_i}{1-r_i}\right]$$

	N	r	Z[r]	SE(Z[r])
S 1	100	0,40	0,42	0,10
S 2	86	0,33	0,34	0,11
S 3	60	0,36	0,38	0,13
S 4	120	0,28	0,29	0,09
S 5	200	0,10	0,10	0,07



HO-Meta-Analysis: Synthesis & Analysis

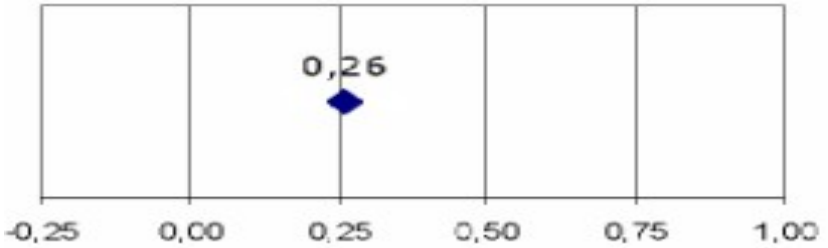
$$\overline{ES} = \frac{\sum_{i=1}^{k} (w_i \times ES_i)}{\sum_{i=1}^{k} w_i}$$

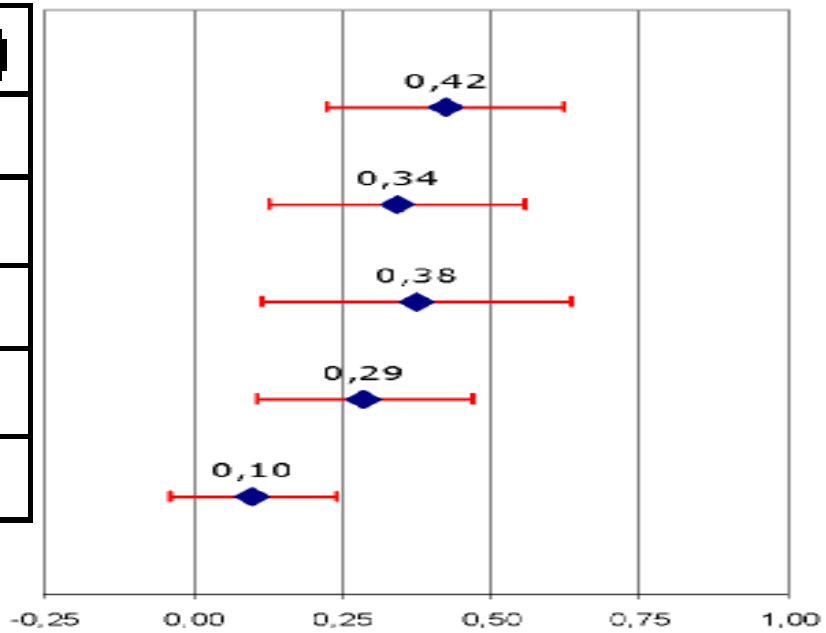
$$w_i = \frac{1}{SE_i^2}$$

$$SE_i = \sqrt{\frac{1}{n-3}}$$

Step 2:
Weighted Synthesis

	N	r	Z[r]	SE(Z[r])
S 1	100	0,40	0,42	0,10
S 2	86	0,33	0,34	0,11
S 3	60	0,36	0,38	0,13
S 4	120	0,28	0,29	0,09
S 5	200	0,10	0,10	0,07



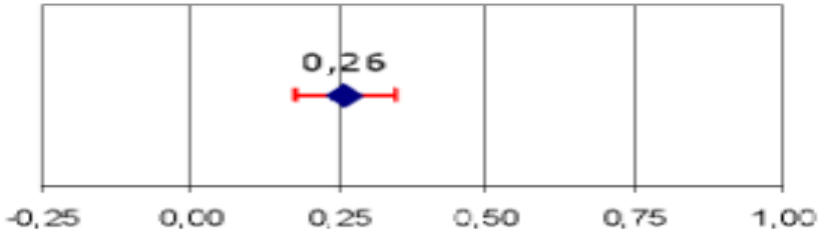


HO-Meta-Analysis: Synthesis & Analysis

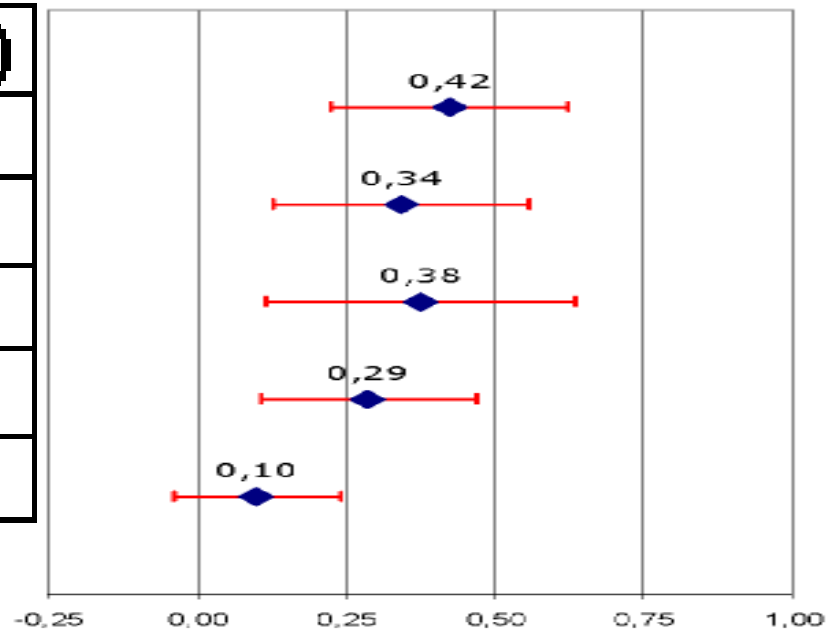
$$\overline{ES} \pm 1.96(SE_{\overline{ES}})$$

$$SE_{\overline{ES}} = \sqrt{\frac{1}{\sum_{i=1}^{k} w_i}}$$

Step 3:
Significance Testing



	N	r	Z[r]	SE(Z[r])
S 1	100	0,40	0,42	0,10
S 2	86	0,33	0,34	0,11
S 3	60	0,36	0,38	0,13
S 4	120	0,28	0,29	0,09
S 5	200	0,10	0,10	0,07

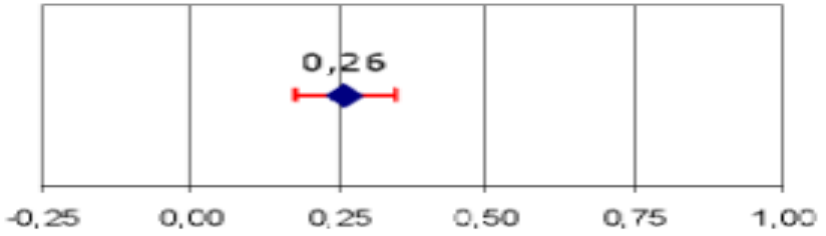


HO-Meta-Analysis: Synthesis & Analysis

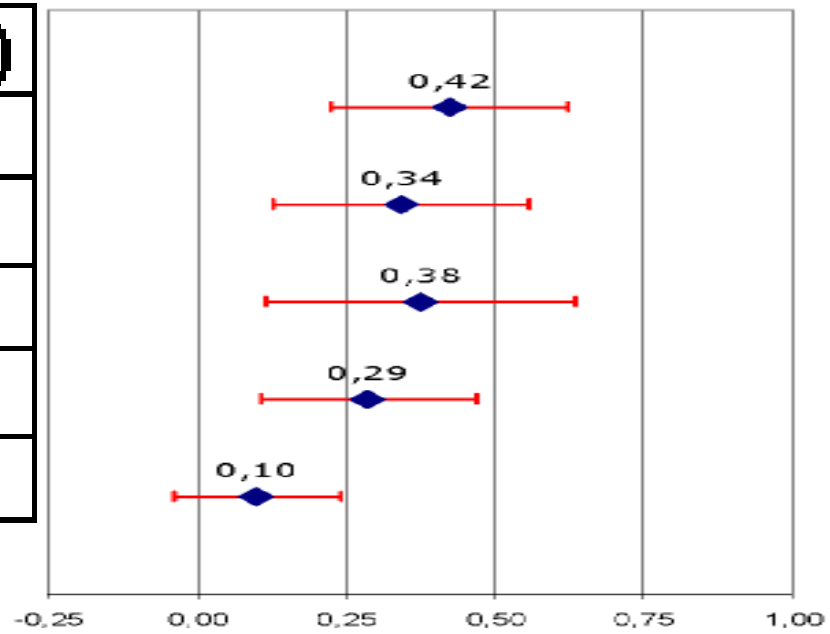
$$Q_T = \sum_{i=1}^{k} \frac{(ES_i - \overline{ES})^2}{SE_i^2} = \sum_{i=1}^{k} w_i (ES_i - \overline{ES})^2$$

Q_T is approx. chi² distributed with *df*= k-1

Step 4:
Homogeneity Analysis



	N	r	Z[r]	SE(Z[r])
S 1	100	0,40	0,42	0,10
S 2	86	0,33	0,34	0,11
S 3	60	0,36	0,38	0,13
S 4	120	0,28	0,29	0,09
S 5	200	0,10	0,10	0,07



Psychometric MA: Data Generation Model

Attenuation model

$$\rho_o = A * \rho$$

„True“ non-attenuated correlation

$$\rho_o = r_o + e$$

$$E(e) = 0$$

$$A = \prod_{j=1}^{m} a_j$$

Examples for a(j):
a(1,2): Measurement errors (att. reliability),
a(3,4): Artificial dichotomization.

Psychometric MA : Procedure

- Disattenuation model:

$$\rho = \frac{\rho_o}{A} = \frac{r_o}{A} + \frac{e}{A} = r_c + e_c$$

- Integration model:

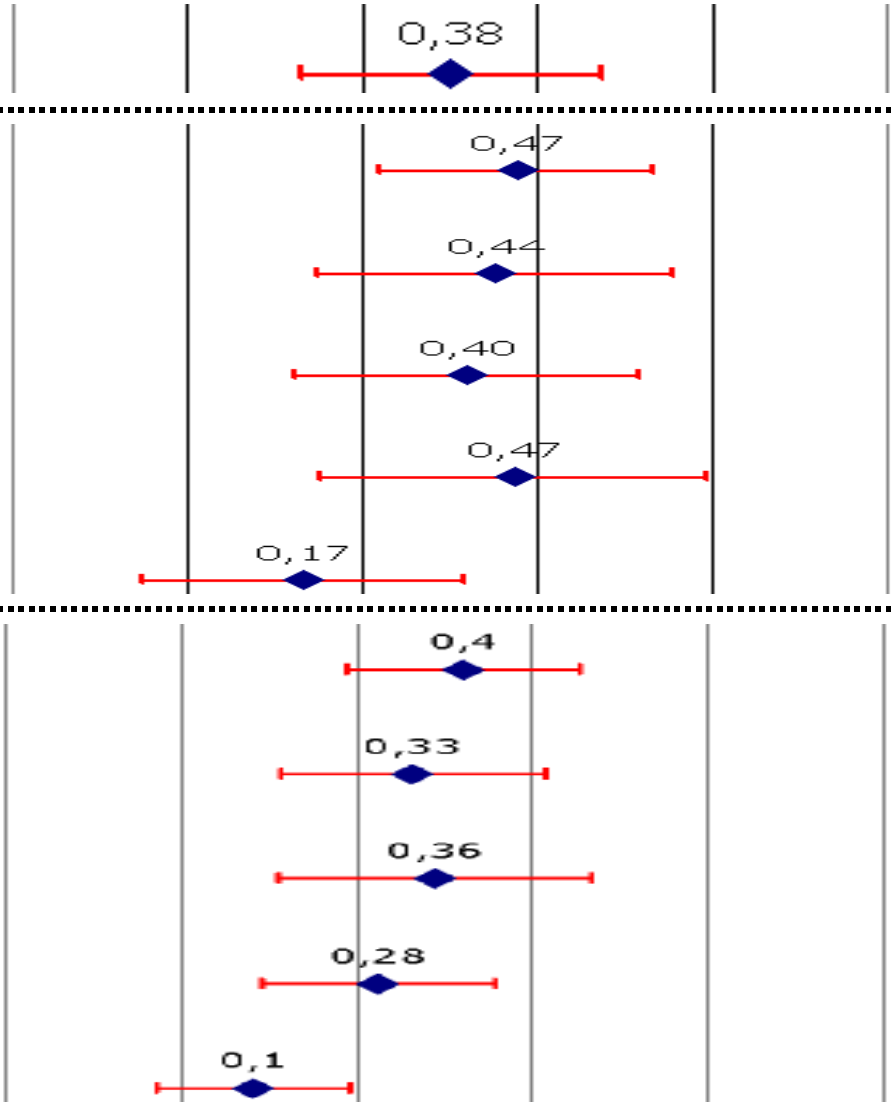
$$\overline{r_c} = \frac{\sum_{i=1}^{k} \left(w_i * r_{c(i)} \right)}{\sum_{i=1}^{k} w_i} \qquad \overline{SE_c^2} = \frac{\sum_{i=1}^{k} \left(w_i * SE_{c(i)}^2 \right)}{\sum_{i=1}^{k} w_i} \qquad w_i = N_i * A_i^2$$

- 75% rule for homogeneity ´testing´ (rule of thumb)

Psychometric MA : Synthesis

	A	r[c]	SE(r[c])
S 1	0,85	0,47	0,10
S 2	0,75	0,44	0,13
S 3	0,90	0,40	0,13
S 4	0,60	0,47	0,14
S 5	0,60	0,17	0,12

	N	r[o]	SE(r[o])
S 1	100	0,40	0,08
S 2	86	0,33	0,10
S 3	60	0,36	0,11
S 4	120	0,28	0,08
S 5	200	0,10	0,07



Discussion : „Evergreens“

- HO- versus psychometric approach?
- Incommensurability:
  - Multiple (general) vs single (narrow) criteria
  - Incommensurability of aggregated constructs (apple-and-oranges problem).
  - Incommensurability of corrected versus uncorrected estimates (´statistical fruit salad´)
- Selective publication?
  file-drawer problem and publication bias
- Quality of primary studies?
- Dependent effect sizes?

Some Recent Trends

- Cumulative and meta-meta-analyses

- Multivariate meta-analyses (MA-SEM):
  - How to synthesize correlation matrices?
  - How to test homogeneity of correlation matrices?
  - Uneven No of available effect sizes*?*
  - Which N to use when doing SEM on synthesized matrices?
  - Correlation versus covariance as input for SEM?

Agenda

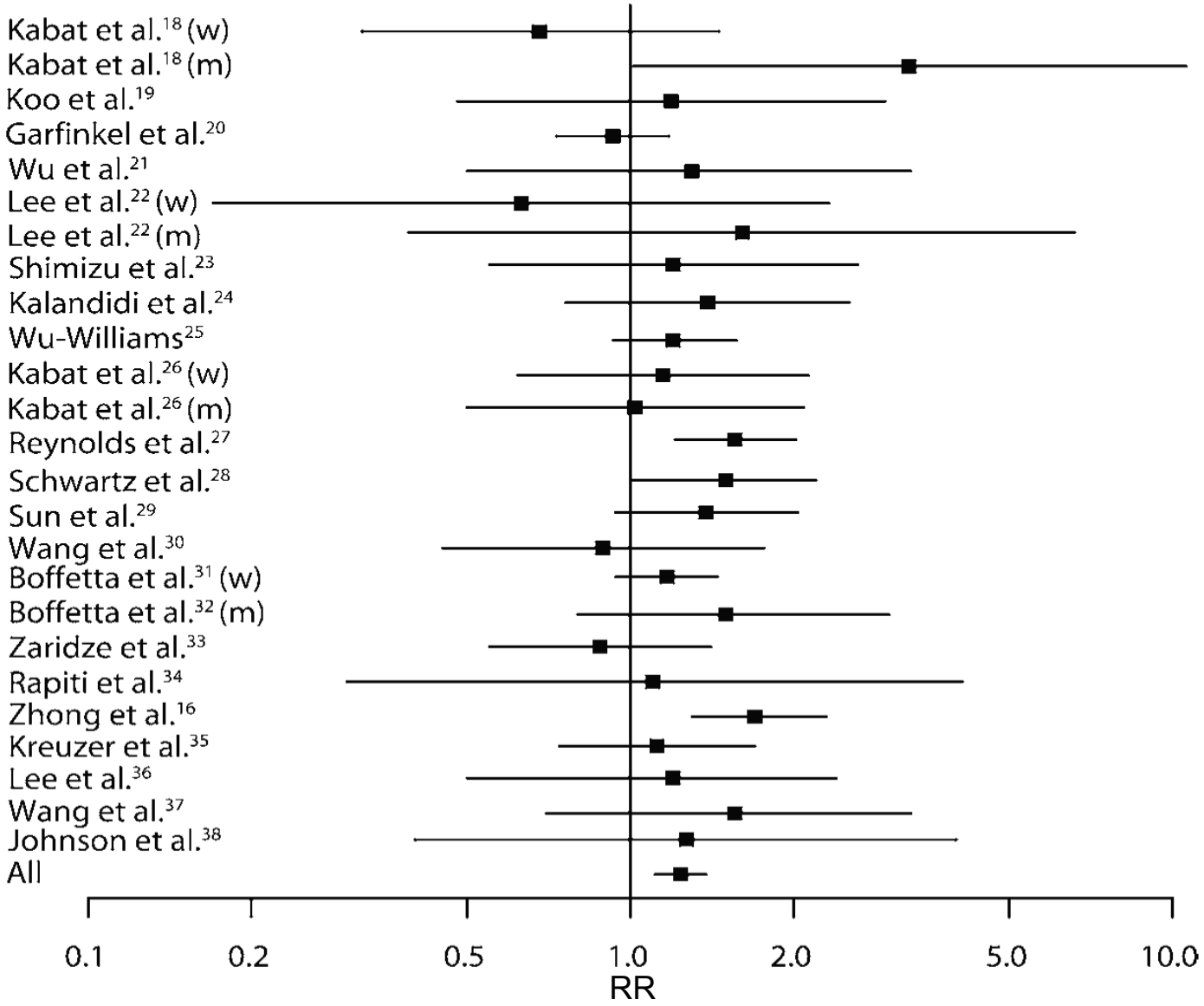
1. Meta-analytic principles and sketch of two common methods (in Psychology)
2. The bone of contention: Meta-analysis on the effects of ETS at workplace on lung cancer (Stayner et al., 2007)
3. Are meta-analyses rock solid? Understanding subjectivity in meta-analytic research processes: The ´crossroads model´ applied to the Stayner data.

Meta-Analysis Stayner et al. (2007)

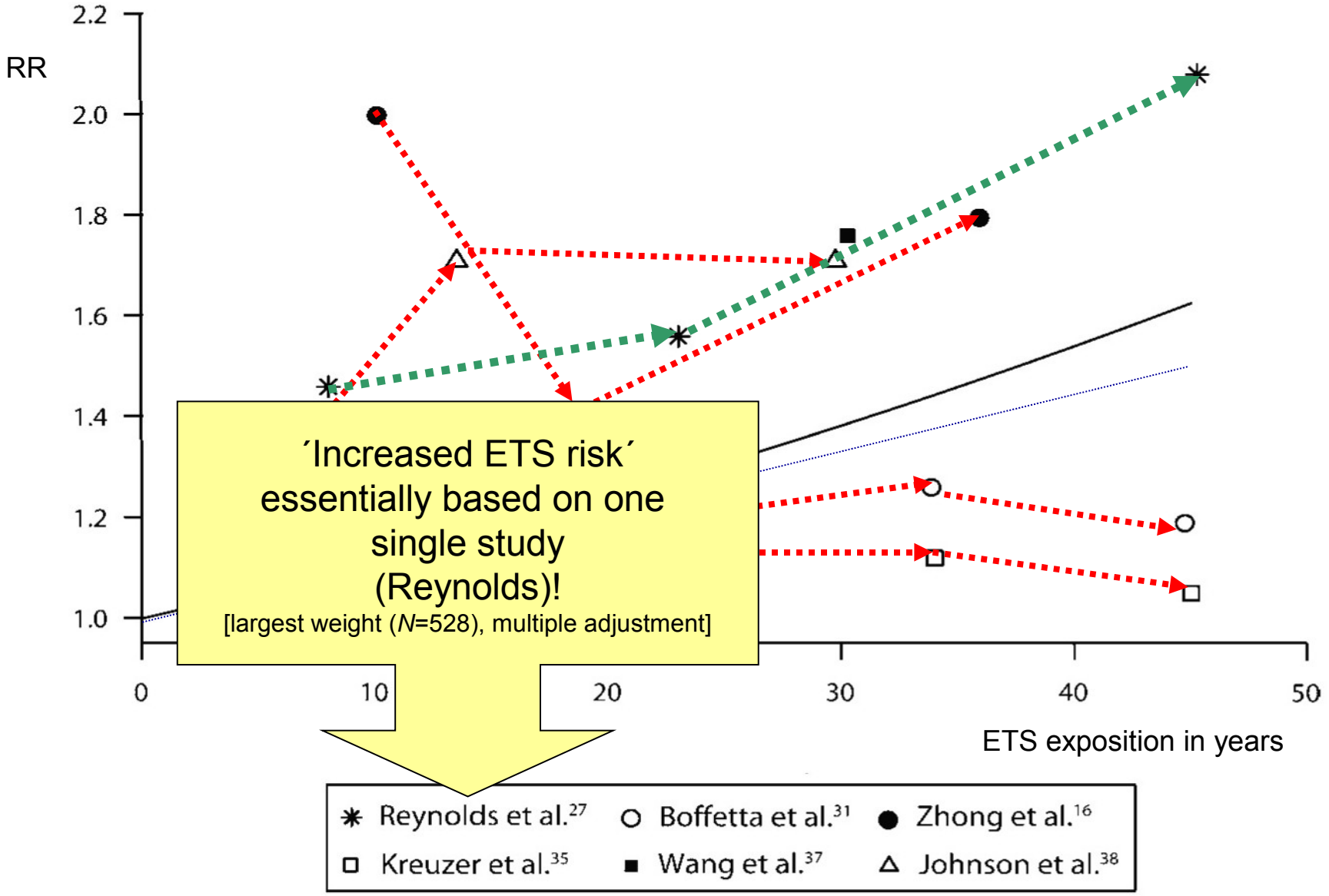
- Increased lung cancer risk in nonsmokers due to ETS exposition at workplace?

- 25 effect sizes (RR) based on 22 papers
- 18 effect sizes were (multiply) adjusted (statistically corrected for the influence of potential confounding variables)
- Mean RR was estimated, homogeneity analysis
- Meta regression to detect moderators
- Dosage-response analysis for time based on 6 primary studies and 17 RR estimates
- Technical procedure: Normand (1999), SAS Proc Mixed, HO meta-analysis

Stayner et al. (2007) : Results I



Stayner et al. (2007) : Results II



Critical Issues : Stayner et al. (2007) I

Quality of documentation (selection):

- Criteria for inclusion/exclusion? Search strategy?
- Information of coding quality?
- Problem of dependent effect sizes?
- Procedure followed for adjusting ES?
- „Statistical fruit salad“: Adjusted and unadjusted estimates mixed.
- Sensitivity and outlier analysis only sketched.

> *Replicabililty low, logical flow of analyses partly unclear*

Documentation in (Medical) Meta-Analyses

Cochrane Collaboration**:

- QUORUM Statement
  (Standards of Reporting of Meta-Analyses)
- STARLITE Statement
  (Standards for Reporting Literature searches)

Coding quality**:

- CONSORT Statement (for parallel-group randomized trials)
- TREND Statement (for non-randomized trials)
- STROBE Statement (Strengthening the Reporting of Observational studies in Epidemiology)

**http://www.cochrane.de/de/statements.htm

Stayner et al. (2007): Conclusions

Robustness of results in this specific case?

General principle(s) influencing robustness of meta-analyses in general? (*Crossroads model*)

Agenda

1. Meta-analytic principles and sketch of two common methods (in Psychology)
2. The bone of contention:
   Meta-analysis on the effects of ETS at workplace on lung cancer
   (Stayner et al., 2007)
3. Are meta-analyses rock solid? Understanding subjectivity in meta-analytic research processes: The ´crossroads model´ applied to the Stayner data.

Subjectivity & Meta-Analyses: Past Studies

Mengersen et al. (1995):

- Method choice influences meta-analytic results in passive smoking research syntheses.
- Choices investigated:
  (a) approximate versus exact statistical techniques, (b) fixed versus random effects models, (c) publication bias / choice of studies included (> *data generation phase*).

Barnes & Bero (1998):

Affiliation of ETS research synthesis authors (tobacco industry / non tobacco industry) single best predictor of results and final conclusion drawn.

Overall Research Question(s)

Replicability of the Stayner et. al. (2007) meta-analysis based on …

1. … the data reported in the paper?
2. … newly coded data?
3. … data taking subjective decisions within a meta-analytic process into account?

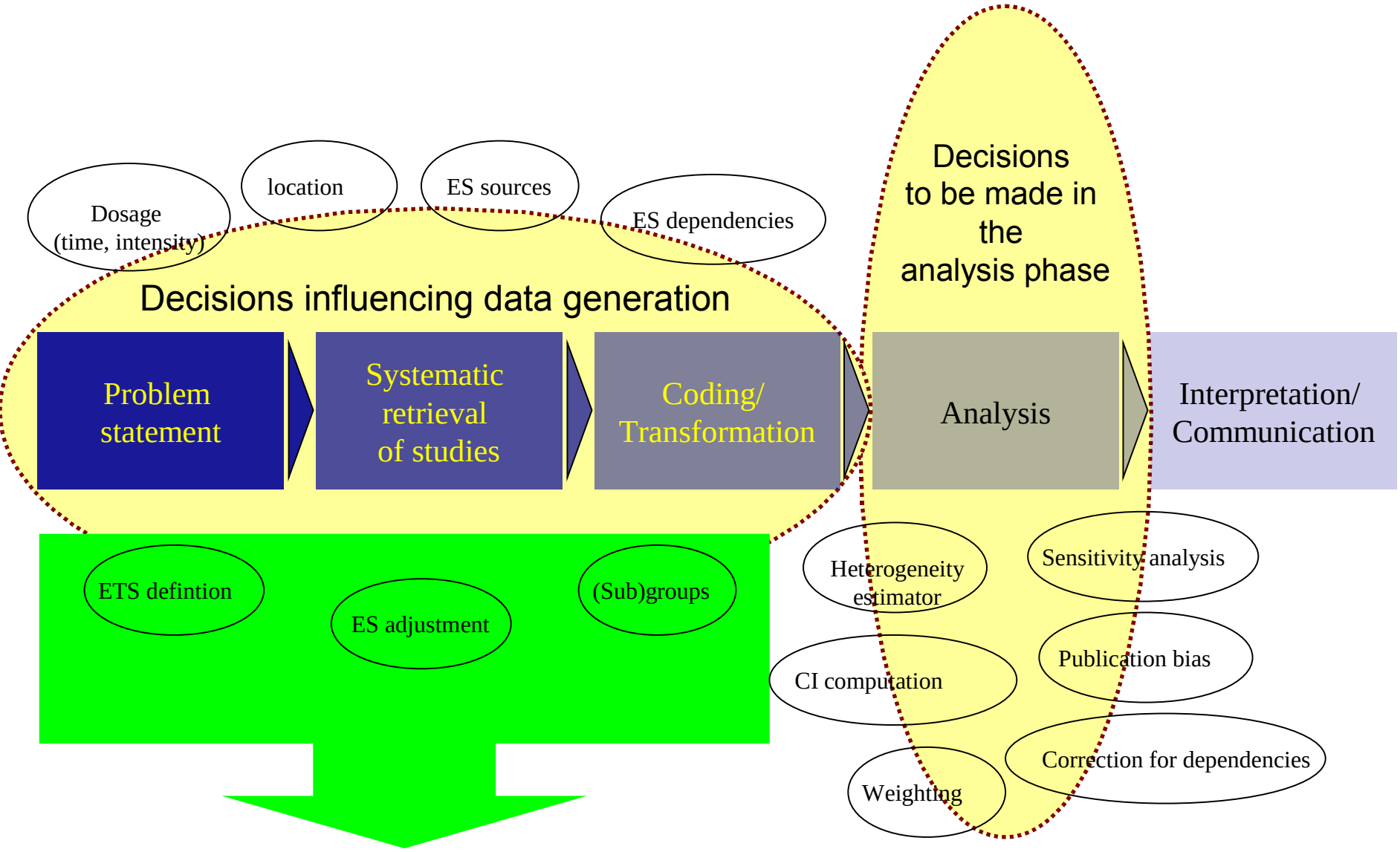
A ´Crossroads Model´ : Basic Idea

- Modeling all possible and reasonable (but still subjective) decisions within a meta-analytic process on the level of (a) data generation and (b) data analysis.

- Current focus of project: Data generation phase (excluding literature retrieval)

- Next step(s), not reported here, simulating the effect of different data analysis approaches on the results (or ranges of "reasonable" results).

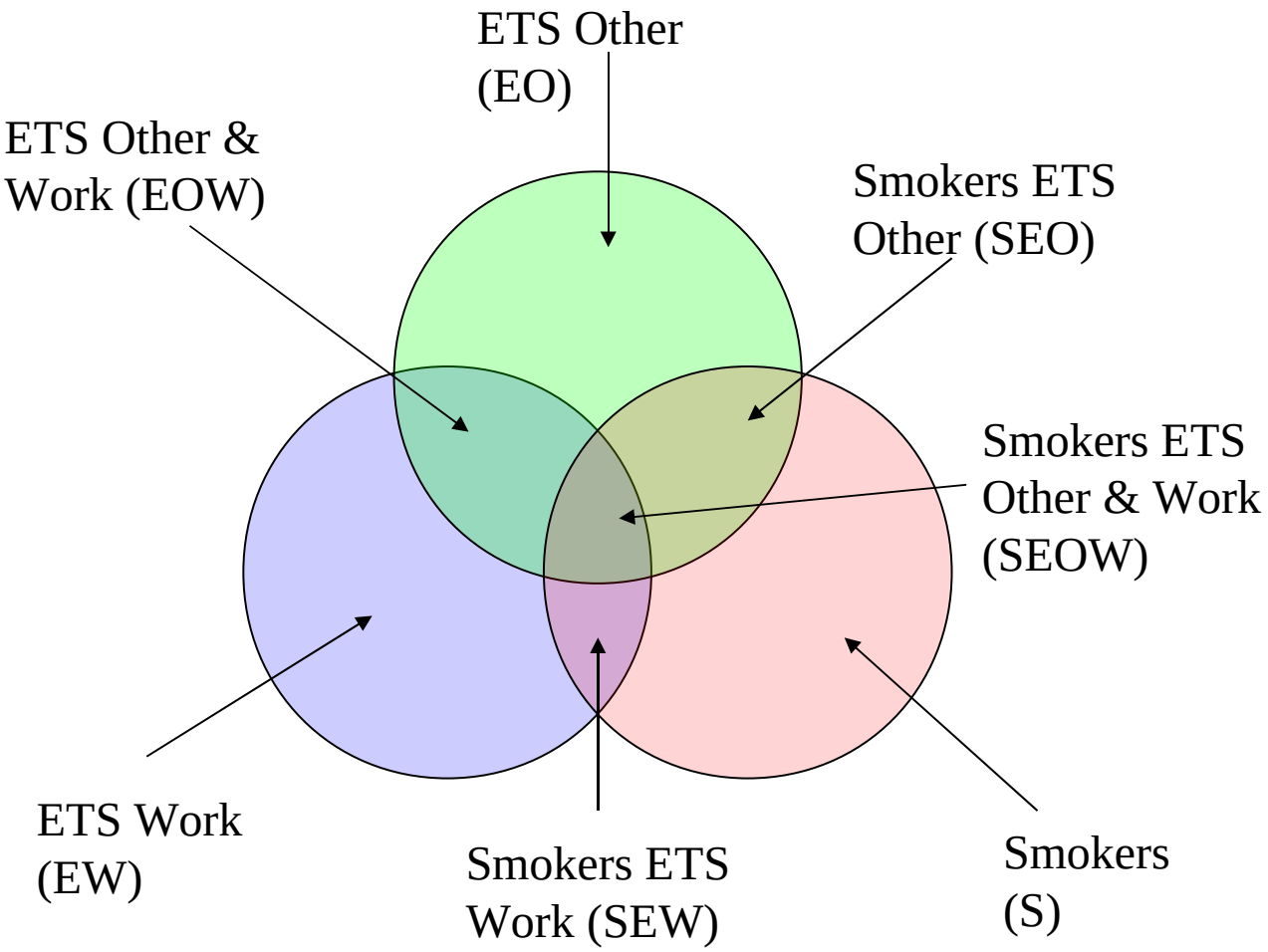
A ´Crossroads Model´ : Examples



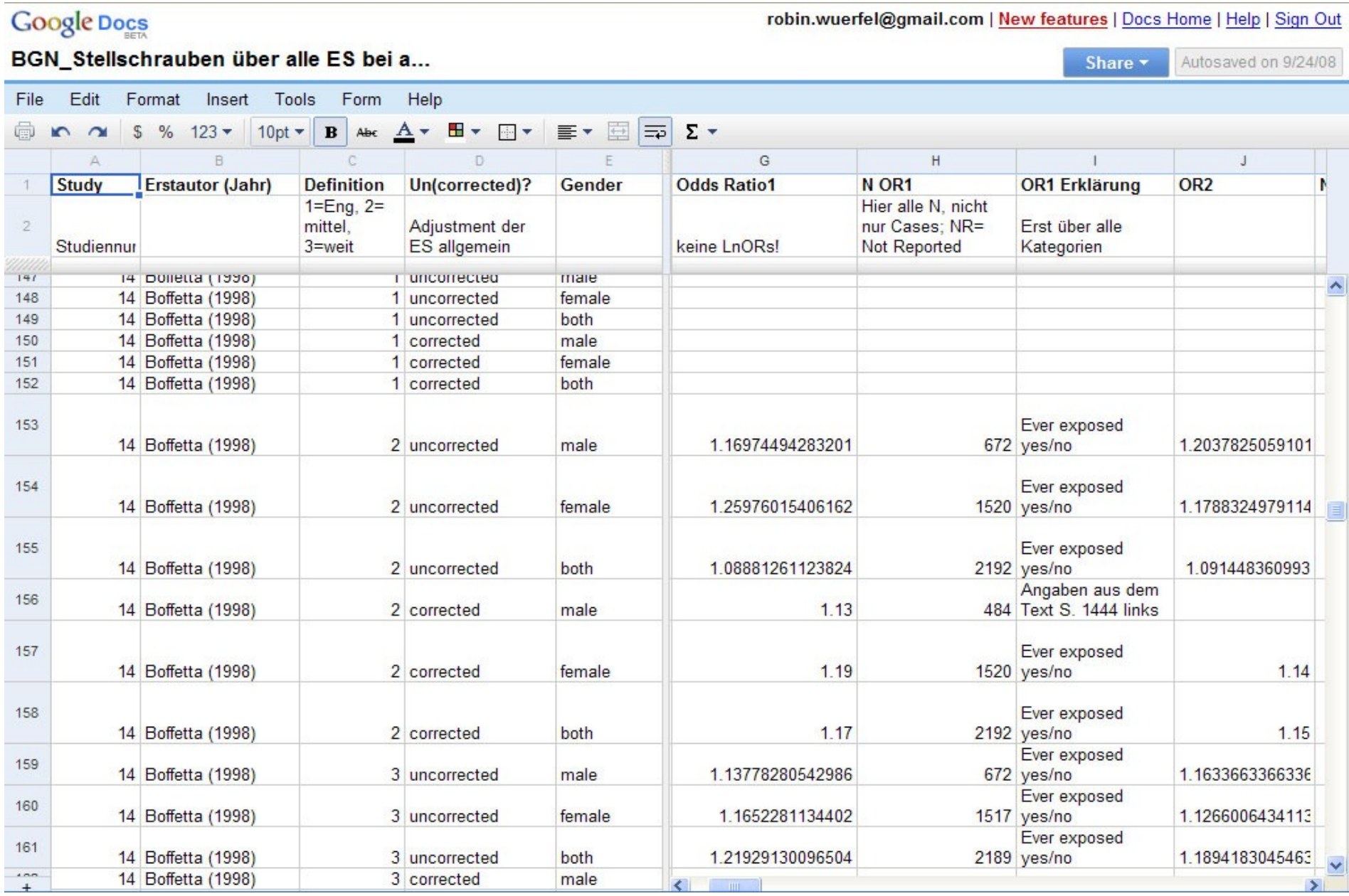
Overview of Methods

- Retrieval of all primary studies used in the Stayner et al. (2007) meta-analysis
- Comprehensive re-coding of data by employing:
  - Different ETS definitions implicitly used in the Stayner paper
  - ES computed for different groups (male/female/both)
  - Adjusted and unadjusted ES computed
  - Quality criteria, intercoder reliability, etc.
- Simulating the effect of all subjective decisions under different restrictions (work in progress)
- Today: Tails of the simulation results (max./min. RR) if only three decisions on the level of data generation are addressed.

ETS (at Work) Definitions



Definitions (implicitly) used in the Stayner et al (2007) paper:
(a) EW, (b) EW & EOW (c) EW & EOW & EO



	A	B	C	D	E	G	H	I	J
1	Study	Erstautor (Jahr)	Definition	Un(corrected)?	Gender	Odds Ratio1	N OR1	OR1 Erklärung	OR2
2	Studiennur		1=Eng, 2= mittel, 3=weit	Adjustment der ES allgemein		keine LnORs!	Hier alle N, nicht nur Cases; NR= Not Reported	Erst über alle Kategorien	
147	14	Boffetta (1998)	1	uncorrected	male				
148	14	Boffetta (1998)	1	uncorrected	female				
149	14	Boffetta (1998)	1	uncorrected	both				
150	14	Boffetta (1998)	1	corrected	male				
151	14	Boffetta (1998)	1	corrected	female				
152	14	Boffetta (1998)	1	corrected	both				
153	14	Boffetta (1998)	2	uncorrected	male	1.16974494283201	672	Ever exposed yes/no	1.2037825059101
154	14	Boffetta (1998)	2	uncorrected	female	1.25976015406162	1520	Ever exposed yes/no	1.1788324979114
155	14	Boffetta (1998)	2	uncorrected	both	1.08881261123824	2192	Ever exposed yes/no	1.091448360993
156	14	Boffetta (1998)	2	corrected	male	1.13	484	Angaben aus dem Text S. 1444 links	
157	14	Boffetta (1998)	2	corrected	female	1.19	1520	Ever exposed yes/no	1.14
158	14	Boffetta (1998)	2	corrected	both	1.17	2192	Ever exposed yes/no	1.15
159	14	Boffetta (1998)	3	uncorrected	male	1.13778280542986	672	Ever exposed yes/no	1.1633663366336
160	14	Boffetta (1998)	3	uncorrected	female	1.1652281134402	1517	Ever exposed yes/no	1.1266006434113
161	14	Boffetta (1998)	3	uncorrected	both	1.21929130096504	2189	Ever exposed yes/no	1.1894183045463
	14	Boffetta (1998)	3	corrected	male				

Overview of Analyses

Block 0: Recap.: Published Results (fixed/random)

Block 1: Replication based on the data reported in the published paper (fixed, random)

Block 2: Replication based on newly coded data (fixed, random)

Block 3: Exploring the tails of the first simulation

- „good guy" analysis (unfavorable towards ETS; fixed, random)
- „bad guy" analysis (favorable towards ETS; fixed random)

Results : Block 0

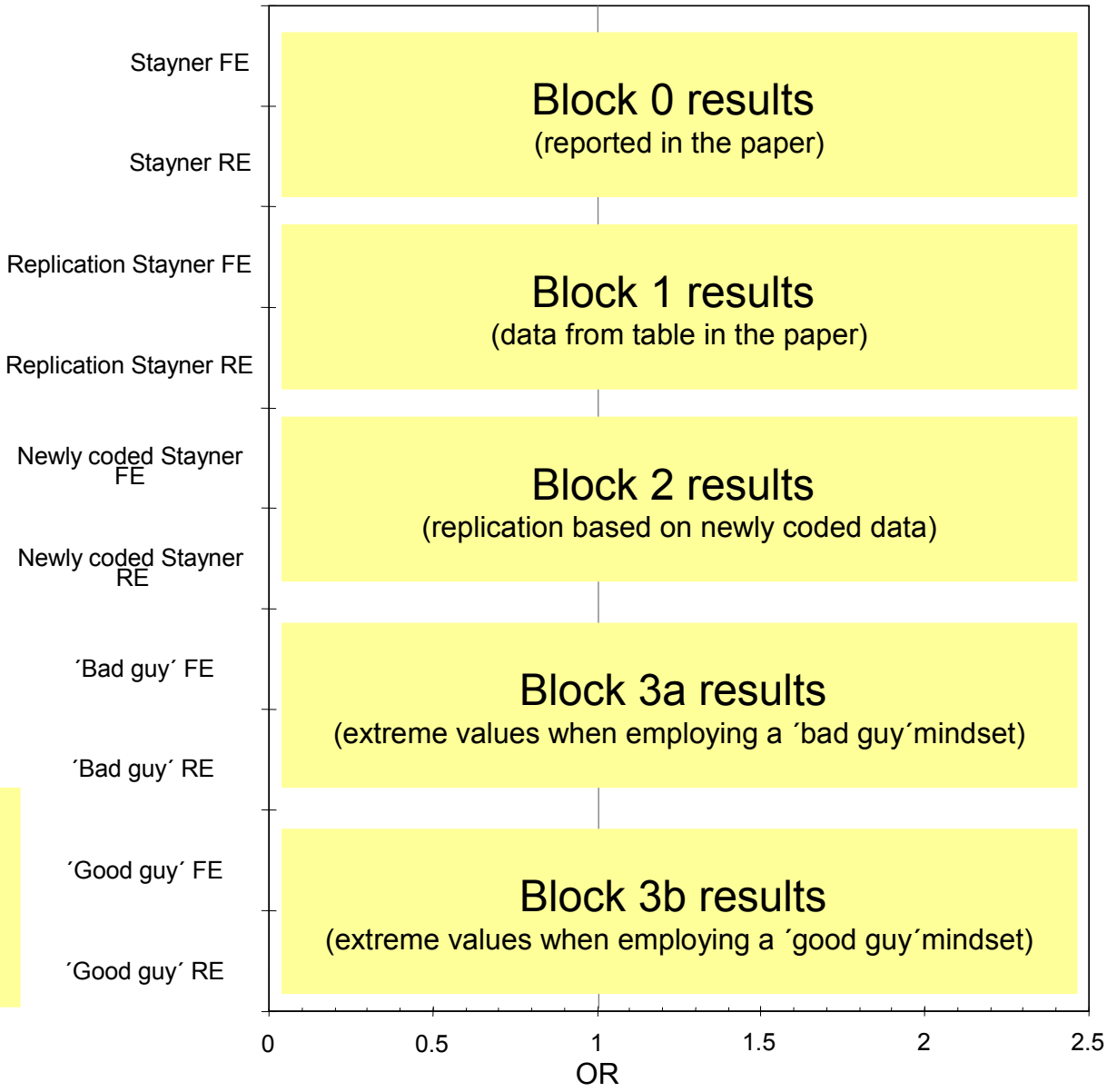
Block 0 results:

FE model:
OR= 1.24
95% CI= 1.18, 1.29

RE model:
OR= 1.24
95% CI= 1.17, 1.31

… based on 25 ES.

Please note:
We are using OR here, not RR (as done in the paper).



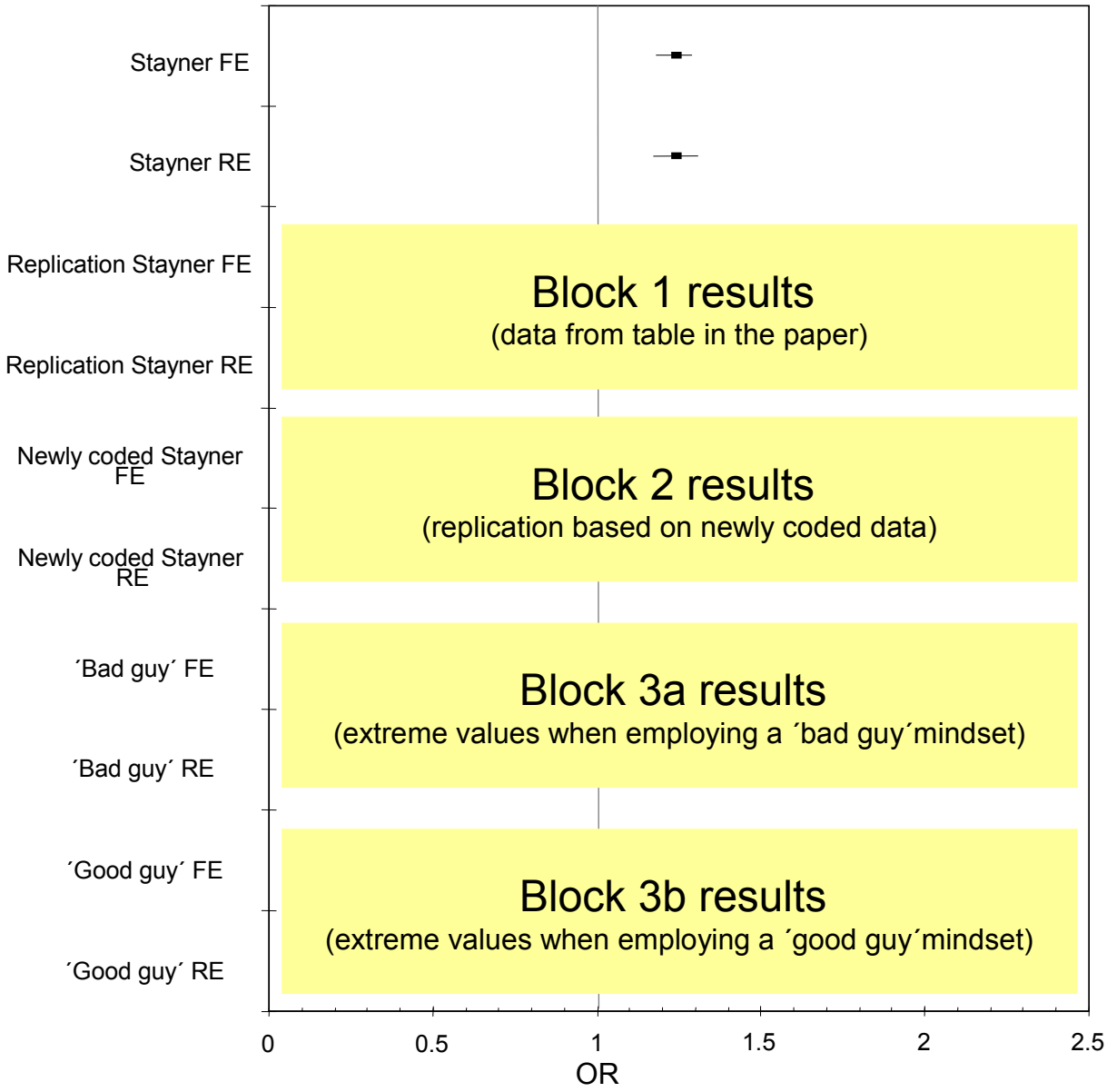
Results : Block 1

Block 1 results:

FE model:
OR= 1.27
95% CI= 1.17, 1.40

RE model:
OR= 1.28
95% CI= 1.16, 1.41

… based on 25 ES.



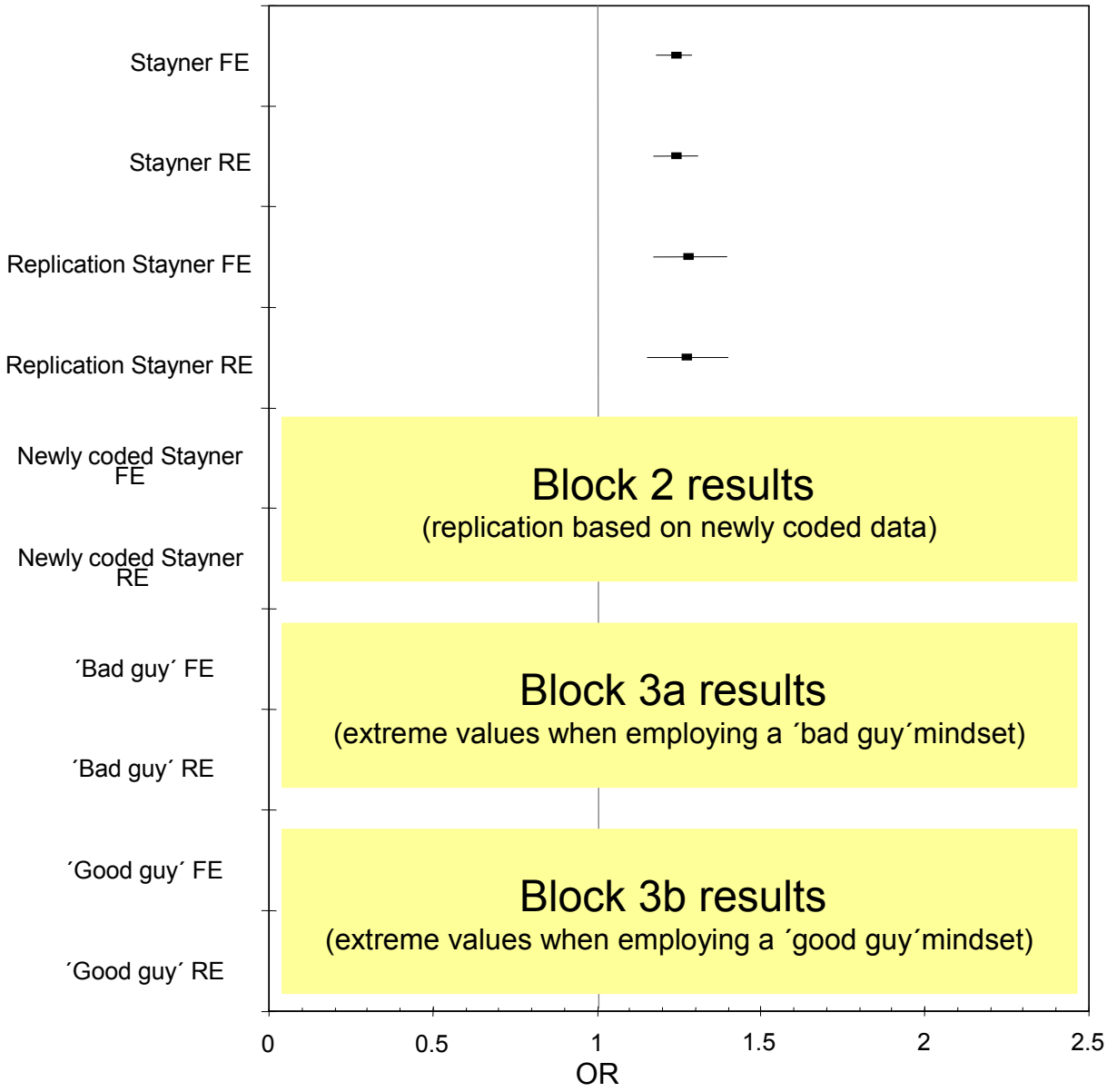
Results : Block 2

Block 2 results:

FE model:
OR= 1.22
95% CI= 1.12, 1.35

RE model:
OR= 1.23
95% CI= 1.10, 1.37

… based on 23 ES.



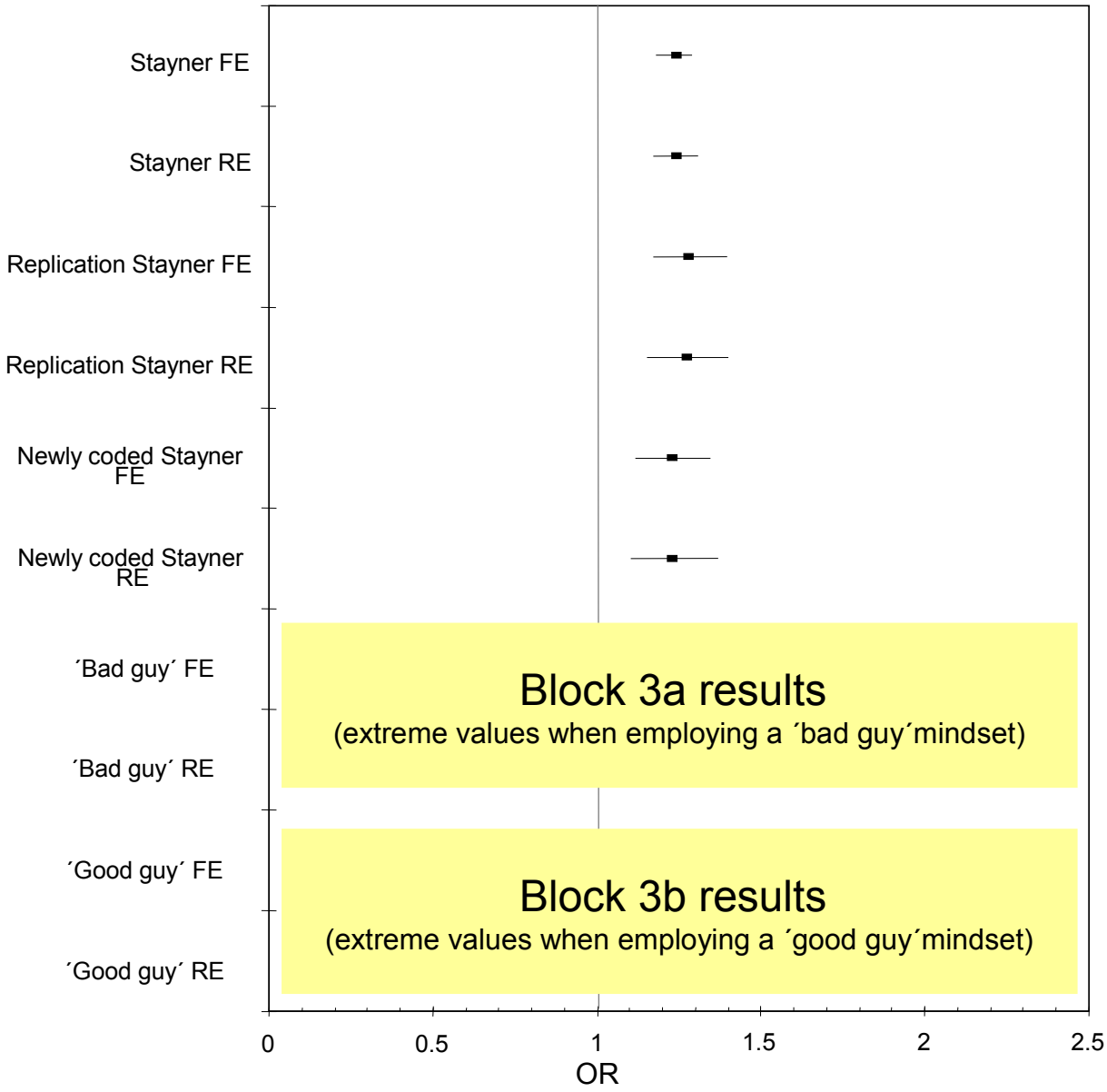
Results : Block 3a

Block 3a results:

FE model:
OR= 0.90
95% CI= 0.84, 0.93

RE model:
OR= 1.00
95% CI= 0.87, 1.14

… based on 22 ES.



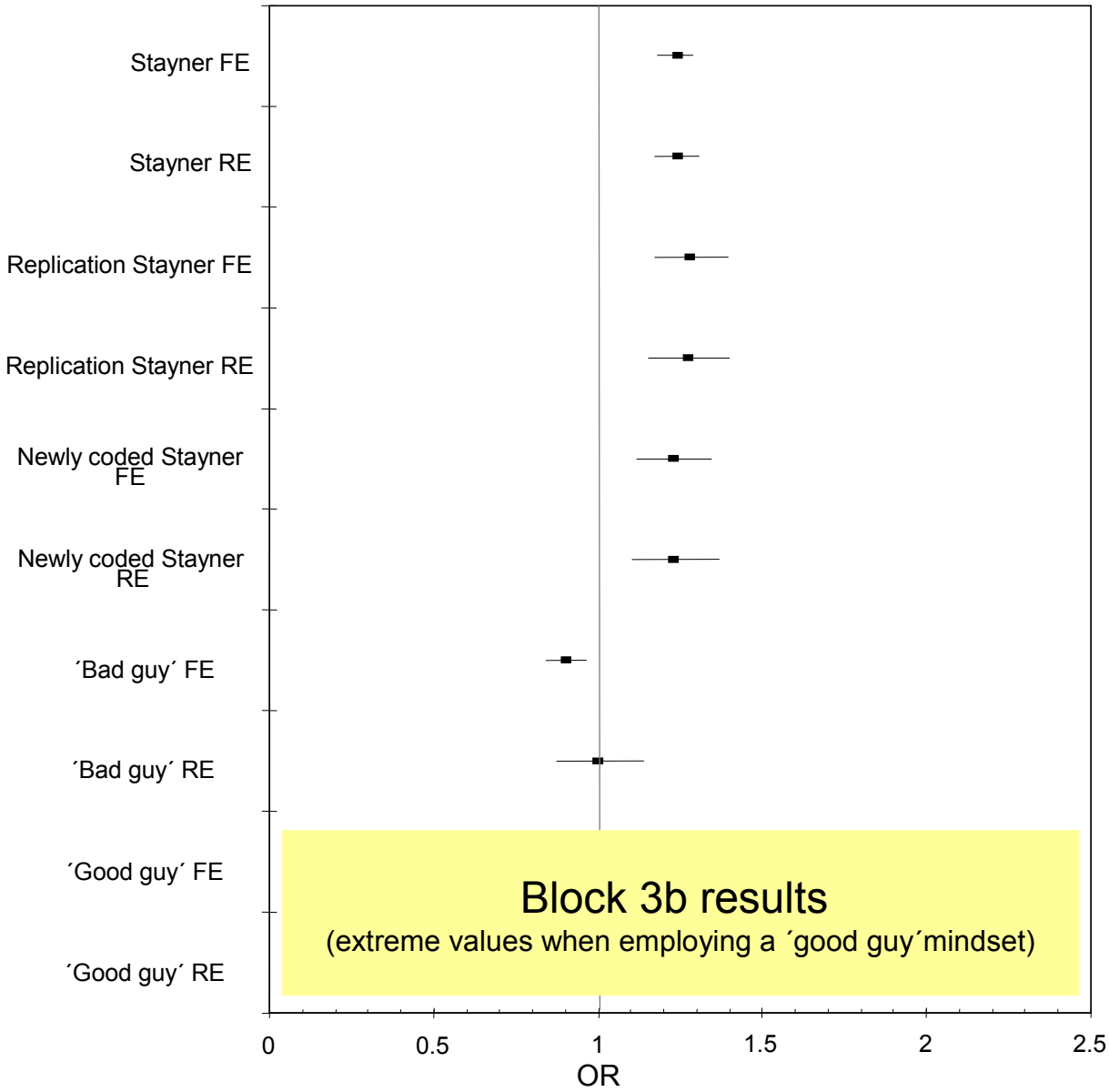
Results : Block 3b

Block 3b results:

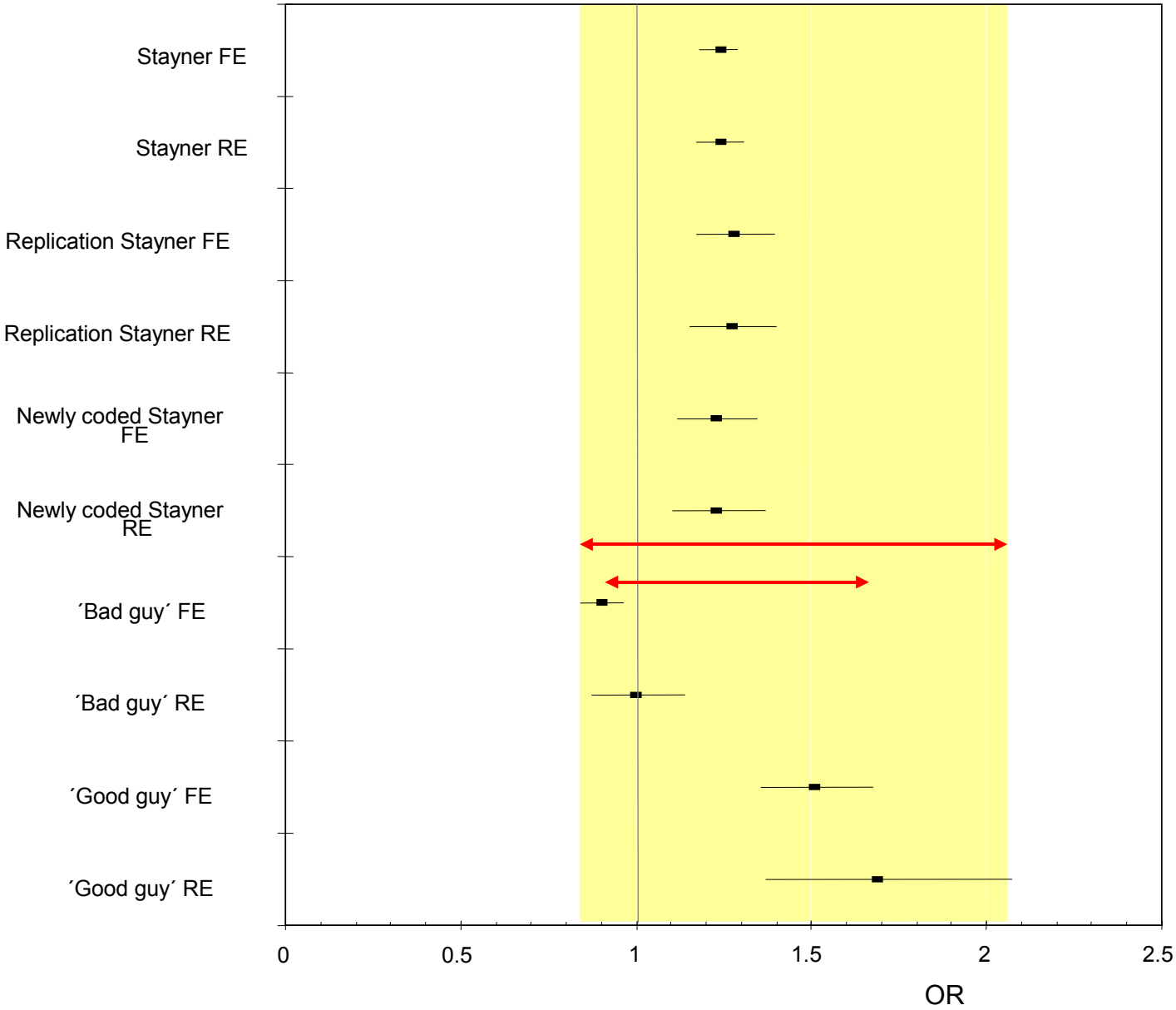
FE model:
OR= 1.51
95% CI= 1.35, 1.68

RE model:
OR= 1.69
95% CI= 1.36, 2.07

… based on 22 ES.



Overall Range *(with only 3 CR variables)*



Preliminary Conclusions and Next Steps

- Modeling only three subjective decisions on the level of data generation may already nullify the reported results of the Stayner et al. (2007) meta-analysis!
- Extending and refining the crossroads model idea may be a valuable contribution to reconstruct subjective decisions in meta-analyses in general …
- … and to stringently test the robustness of results (and the conclusions based on them).
- Next step(s):
  - further developing, extending, and refining the crossroads idea
  - detecting dependencies
  - simulating different results under realistic restrictions.

Thank you!

- Sources on meta-analysis:
  http://www.meta-analysis.eu

- Related paper (in press, unfortunately in German):

  Bosnjak, M. & Viechtbauer, W. (in press). Die Methode der Meta-Analyse zur Evidenz-basierung von Gesundheitsrisiken: Beiträge der Sozial-, Verhaltens- und Wirtschaftswissenschaften. *Zentralblatt für Arbeitsmedizin, Arbeitsschutz und Ergonomie.*

- e-mail:
  michael.bosnjak@unibz.it

- Slides:
  http://invited-presentations.bosnjak.eu

Literature

- Barnes, D.E., & Bero. L.A. (1998). Why review articles on the health effects of passive smoking reach different conclusions. *Journal of the American Medical Association, 279*(19), 1566-1570.
- Mengersen, K.L., Tweedie, R.L., & Biggerstaff, B. (1995). The impact of method choice on meta-analysis. *Australian Journal of Statistics, 37*(1), 19-44.
- Normand, S.-L.T. (1999). Meta-analysis: Formulating, evaluating, combining, and reporting. *Statistics in Medicine, 18*, 321-359.
- Stayner, L., Bena, J., Sasco, A.J., Smith, R., Steenland, K., Kreuzer, M., & Straif, K. (2007). Lung cancer risk and workplace exposure to environmental tabacco smoke. *American Journal of Public Health, 97*(3), 545-551.

Supplements

Relative Risk (RR)

$$RR = \frac{P(Erkrankung|mitRisikofaktor)}{P(Erkrankung|ohneRisikofaktor)}$$

Berechnung des relativen Risikos aus einer 2*2 Häufigkeitstabelle:

	mit Risikofaktor	**ohne Risikofaktor**
erkrankt	a	b
nicht erkrankt	c	d

$$RR = \frac{a/(a+c)}{b/(b+d)}$$

Odds Ratio (OR)

Berechnung des Odds Ratio aus einer 2*2 Häufigkeitstabelle:

	mit Risikofaktor	**ohne Risikofaktor**
erkrankt	a	b
nicht erkrankt	c	d

$$\text{Odds Ratio} = \frac{a/c}{b/d} = \frac{a \cdot d}{b \cdot c}$$

RR versus OR

Vergleich unterschiedlicher ES anhand einer **fiktiven** 2*2 Häufigkeitstabelle:

	mit Risikofaktor	**ohne Risikofaktor**
erkrankt	130	70
nicht erkrankt	1870	7930

$$RR = \frac{130/(130+1870)}{70/(70+7930)} \approx 7,4$$

$$\text{Odds Ratio} = \frac{130 \cdot 7930}{70 \cdot 1870} \approx 7,88$$