De beoordeling van wetenschappelijke evidentie. To GRADE or not to GRADE.

Webinar Academische Werkplaats Gezonde Leefomgeving

3 april 2023

Gerard Hoek
IRAS, University Utrecht
The Netherlands

Universiteit Utrecht

Achtergrond

- Groot aantal studies naar milieu en gezondheid
- Epidemiologisch en experimenteel
- Beoordeling van evidentie via systematische reviews inclusief meta-analyse
- Vaste procedures ontwikkeld
- Element daarin beoordeling kwaliteit evidence base epidemiologische studies
- GRADE vanuit klinische studies
- Inclusief Risk of Bias tools
- Veel toegepast maar zijn niet zonder problemen
- Goed voor GGD medewerkers om daarvan op de hoogte te zijn

Presentatie

- Illustratie belang inzicht in “GRADE”
- Toepassing in WHO systematic review
- Toepassing in HEI systematic review
- Beschouwing

Confidence assessment using GRADE

- Developed in clinical studies to assess the confidence in the body of evidence that a certain intervention improves health
- Systematic methodology
- Applied more and more to observational studies e.g. on the environment
- Example: the noise guidelines by WHO

Noise guidelines World Health Organization 2018

Table 6. Average exposure levels (L_{den}) for priority health outcomes from road traffic noise

Summary of priority health outcome evidence	Benchmark level	Evidence quality
Incidence of IHD The 5% relevant risk increase occurs at a noise exposure level of 59.3 dB L_{den}. The weighted average of the lowest noise levels measured in the studies was 53 dB L_{den} and the RR increase per 10 dB is 1.08.	5% increase of RR	High quality
Incidence of hypertension One study met the inclusion criteria. There was no significant increase of risk associated with increased noise exposure in this study.	10% increase of RR	Low quality
Prevalence of highly annoyed population There was an absolute risk of 10% at a noise exposure level of 53.3 dB $L_{den.}$	10% absolute risk	Moderate quality
Permanent hearing impairment	No increase	No studies met the inclusion criteria
Reading skills and oral comprehension in children	One-month delay	Very low quality

Evidence quality affects the strength of the WHO recommendation

GRADE

- The Grading of Recommendations Assessment, Development and Evaluation (short GRADE) working group began in the year 2000 as an informal collaboration of people with an interest in addressing the shortcomings of grading systems in health care. The working group has developed a common, sensible and transparent approach to grading quality (or certainty) of evidence and strength of recommendations. Many international organizations have provided input into the development of the GRADE approach which is now considered the standard in guideline development.
- https://www.gradeworkinggroup.org/
- https://nl.gradeworkinggroup.org/

GRADE elementen

- Transparant systeem ipv narratief
- Vaste structuur van beoordeelde issues
- Toegepast door veel organisaties waaronder de WHO
- Niet door bv US EPA in ISA en IARC in carcinogeniteit assessment

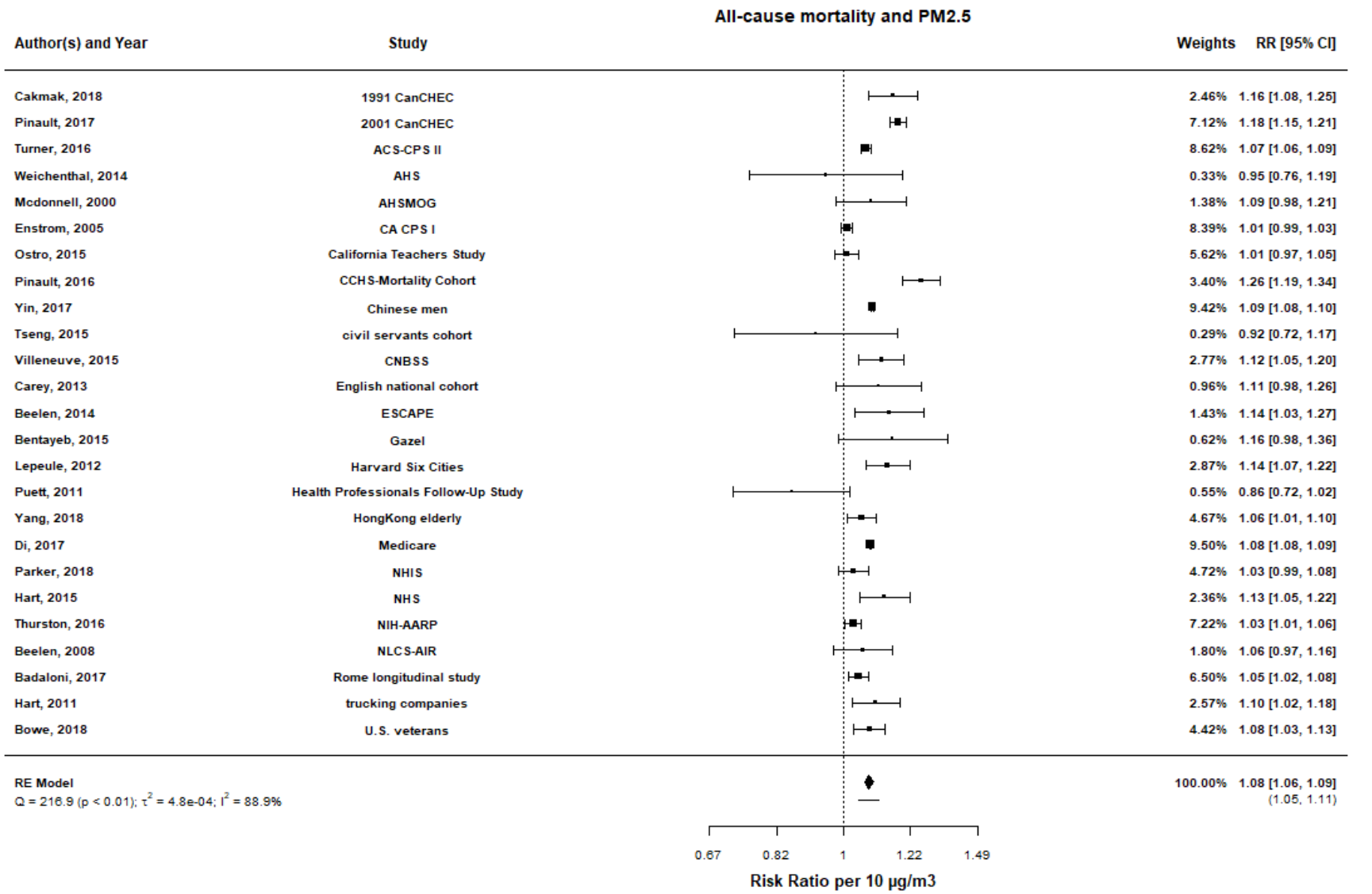
WHO systematic review air pollution

- Application GRADE obligatory
- Long-terme PM2.5, PM10 and cause-specific mortality
- Systematic review including meta-analysis
- GRADE inclusief Risk of Bias

Methods of systematic review

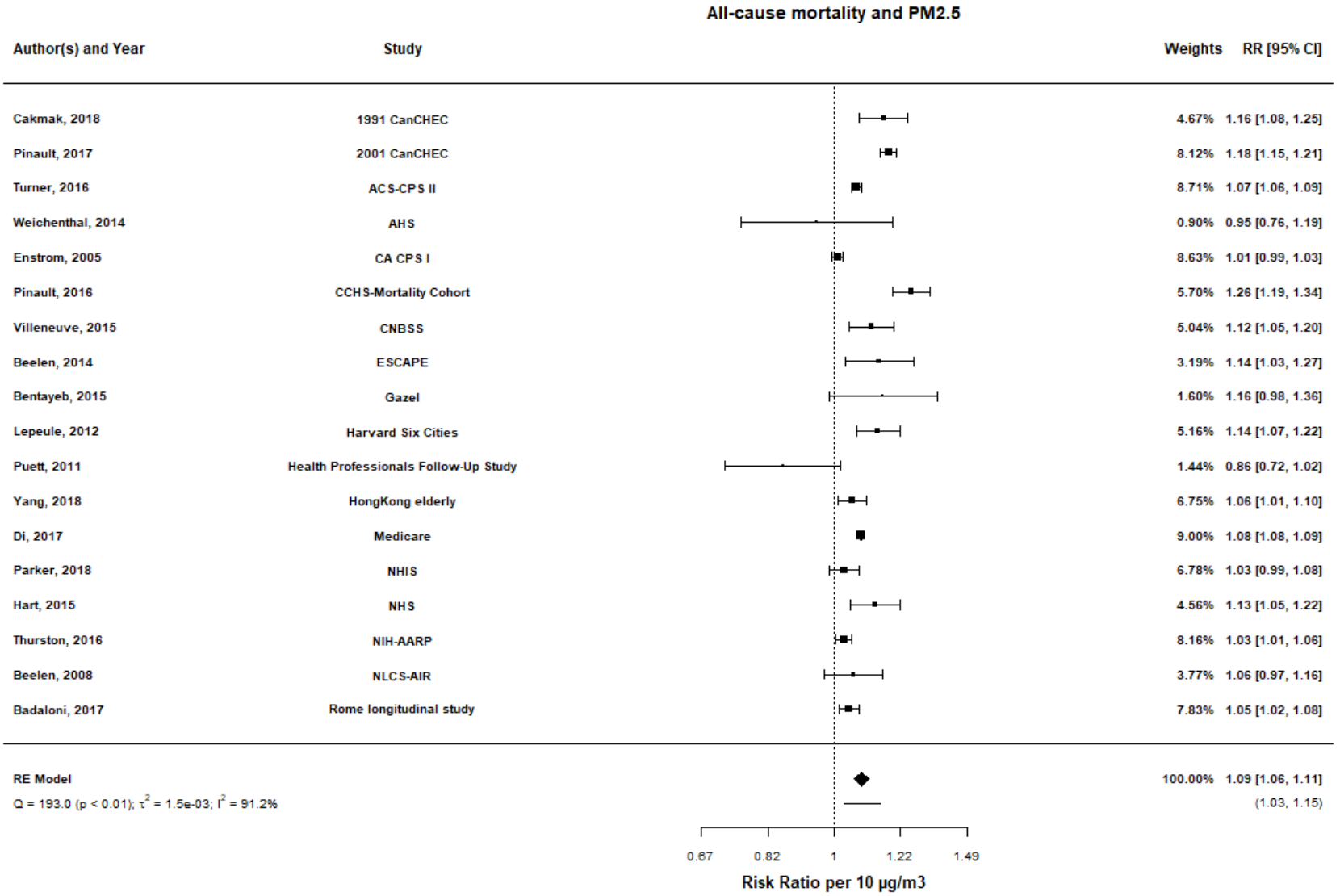
1. Systematic search in Pubmed and Embase using Mesh terms and free terms
2. Evaluation abstracts by two investigators
3. Evaluation of full-text papers by two investigators-> selection of studies
4. Data extraction with standard form two investigators
5. Meta-analysis
6. Risk of bias
7. Grade

Figure 3 PM2.5 and all-cause mortality: meta-analysis



Domain	Subdomain	Low-risk	Moderate-risk	High-risk
1. Confounding	Were all confounders considered adjusted for in the analysis?	3	22	
	Validity of measuring of confounding factors	23	2	
	Control in analysis (Did the authors use an appropriate analysis method or study design that controlled for confounding domains?)	25		
	Overall	3	22	
2. Selection bias	Selection of participants into the study (includes non-response)	25		
	Overall	25		
3. Exposure assessment	Methods used for exposure assessment	22	3	
	Exposure measurement methods comparable across the range of exposure	25		
	Change in exposure status (for long-term studies only)	22	3	
	Exposure contrast	24	1	
	Overall	22	3	
4. Outcome measurement	Blinding of outcome measurement	25		
	Validity of outcome measurements	25		
	Outcome measurement	25		
	Overall	25		
5. Missing data	Missing data of outcome measures	25		
	Missing data of exposures	24	1	
	Overall	24	1	
6. Selective reporting	Authors reported a priori primary and secondary study aims	24		1

Figure 60 Meta-analysis of PM2.5 and all-cause mortality: studies with high/moderate risk of bias excluded

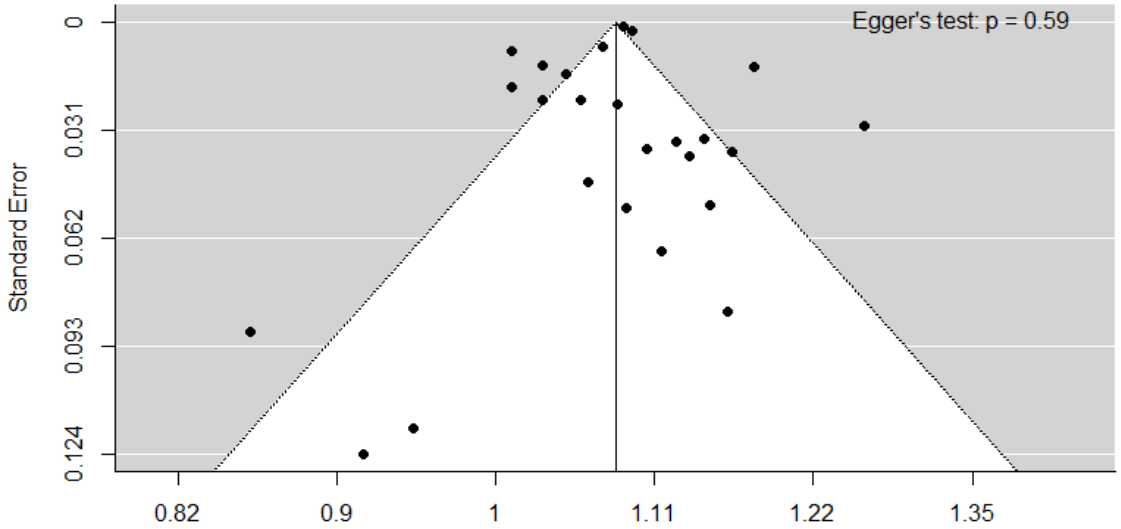


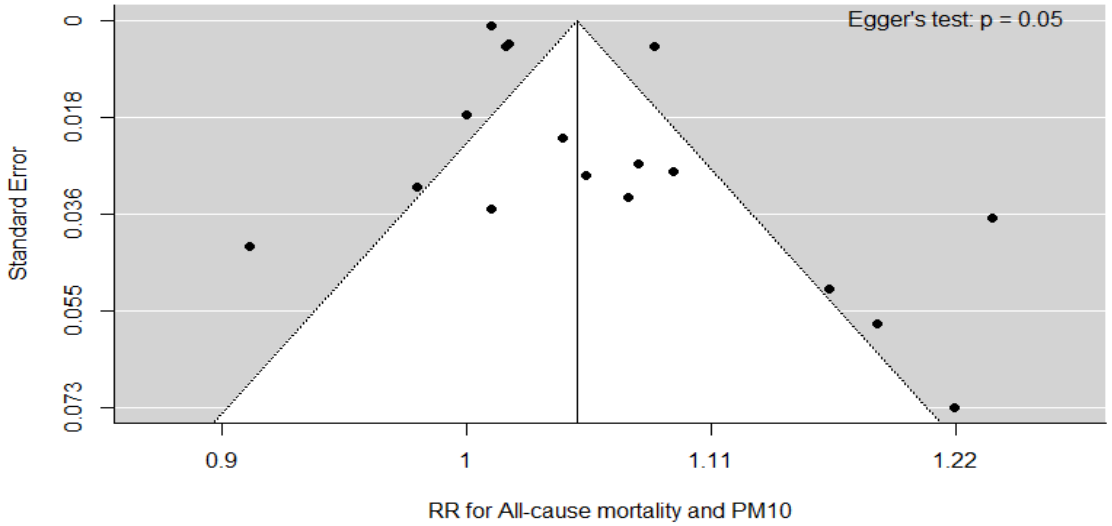
GRADE assessment of quality of evidence

A1 = limitations in studies (risk of bias); A2 = indirectness;
A3 = inconsistency; A4 = imprecision; A5 = publication bias

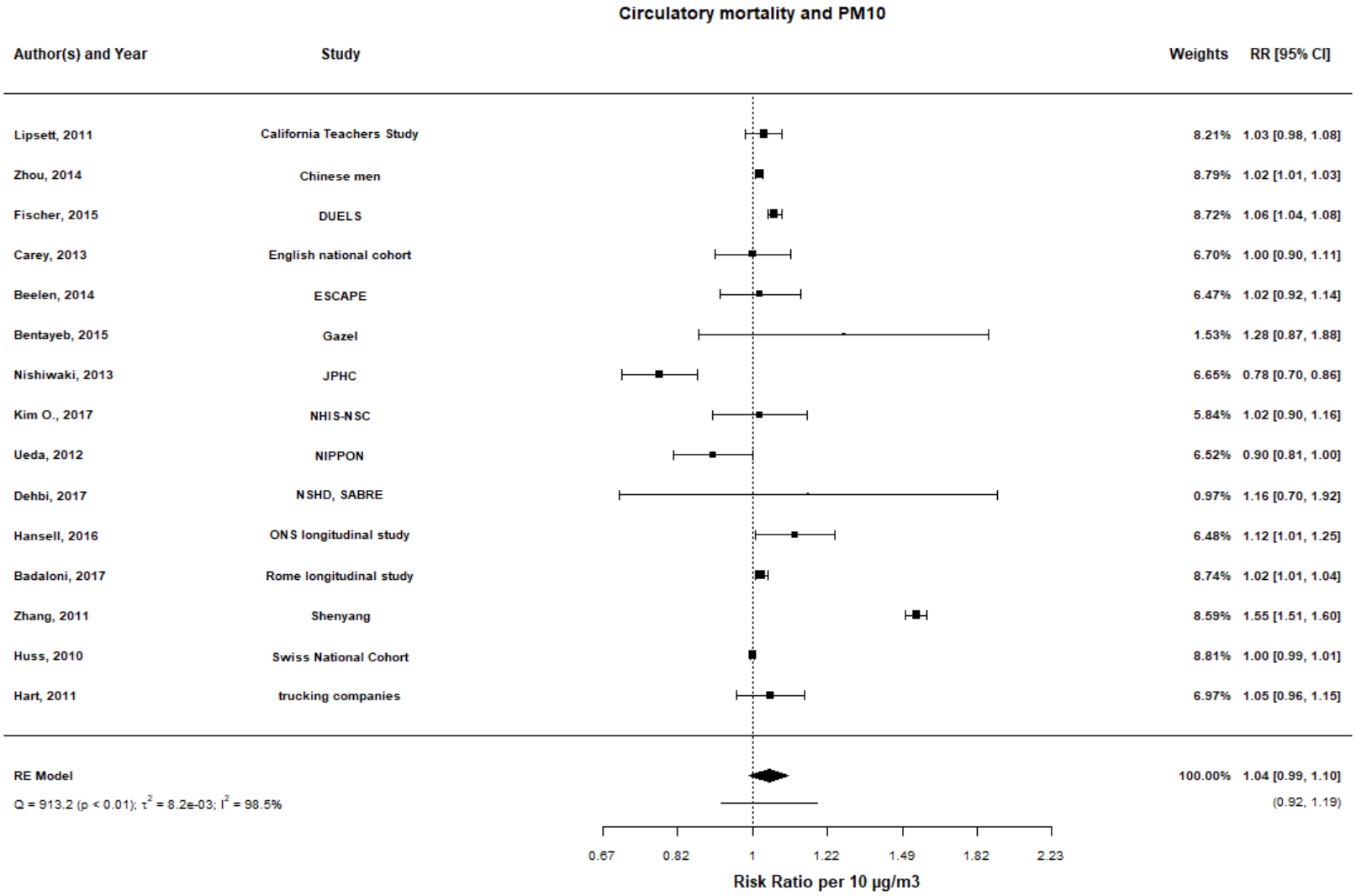
	reasons for downgrading									
	A1	rationale	A2	rationale	A3	rationale	A4	rationale	A5	rationale
PM2.5 and all-cause	0	little influence on the overall effect	0	no evidence of indirectness	0	prediction interval does not include unity	0	sample size large enough to assess RR with sufficient precision	0	no evidence of publication bias
PM10 and all-cause	0	little influence on the overall effect	0	no evidence of indirectness	0	prediction interval does not include unity	0	sample size large enough to assess RR with sufficient precision	0	no evidence of publication bias
PM2.5 and circulatory	0	little influence on the overall effect	0	no evidence of indirectness	0	prediction interval does not include unity	0	sample size large enough to assess RR with sufficient precision	0	no evidence of publication bias
PM10 and circulatory	0	little influence on the overall effect	0	no evidence of indirectness	-1	prediction interval includes unity	0	sample size large enough to assess RR with sufficient precision	0	no evidence of publication bias

Funnel plots to evaluate publication bias





Universiteit



Universiteit Utrecht

Lung Cancer mortality and PM10

Author(s) and Year	Study	Weights	RR [95% CI]
Abbey, 1999	AHSMOG	1.28%	1.37 [0.94, 2.00]
Lipsett, 2011	California Teachers Study	6.06%	0.93 [0.81, 1.07]
Zhou, 2014	Chinese men	14.35%	1.01 [0.99, 1.03]
Fischer, 2015	DUELS	13.46%	1.26 [1.22, 1.31]
Carey, 2013	English national cohort	4.89%	1.03 [0.88, 1.22]
Heinrich, 2013	German	0.59%	2.39 [1.35, 4.23]
Nishiwaki, 2013	JPHC	4.99%	0.91 [0.77, 1.07]
Kim O., 2017	NHIS-NSC	5.09%	0.96 [0.82, 1.13]
Chen X., 2016	Northern China	14.58%	1.05 [1.03, 1.06]
Hansell, 2016	ONS longitudinal study	3.30%	1.60 [1.29, 1.99]
Huss, 2010	Swiss National Cohort	14.53%	1.05 [1.04, 1.07]
Katanoda, 2011	Three-prefecture Cohort Study	10.57%	1.16 [1.08, 1.25]
Hart, 2011	trucking companies	6.32%	1.00 [0.87, 1.14]
RE Model		100.00%	1.08 [1.03, 1.13]
			(0.99, 1.18)

Q = 153.9 (p < 0.01); τ^2 = 3.5e-03; I^2 = 92.2%

GRADE assessment of quality of evidence for each exposure-outcome

B1 = large RR; B2 = all confounding decreases observed RR; B3= dose-response gradient

	reasons for upgrading						Change	Overall assessment
	B1	rationale	B2	rationale	B3	rationale		
PM2.5 and all-cause	0	E-value=2.06 (40 µg/m3 vs. 10 µg/m3)	0	confounders would shift the RR in both directions	+1	evidence of increase in risk with increasing exposure	+1	High
PM10 and all-cause	0	E-value=1.73 (40 µg/m3 vs. 10 µg/m3)	0	confounders would shift the RR in both directions	+1	evidence of increase in risk with increasing exposure	+1	High
PM2.5 and circulatory	0	E-value=2.40 (40 µg/m3 vs. 10 µg/m3)	0	confounders would shift the RR in both directions	+1	evidence of increase in risk with increasing exposure	+1	High
PM10 and circulatory	0	no upgrading because of downgrading	0	no upgrading because of downgrading	0	no upgrading because of downgrading	-1	Low

Systematic review of mortality effects of long-term exposure to traffic-related air pollution

Gerard Hoek on behalf of the Panel

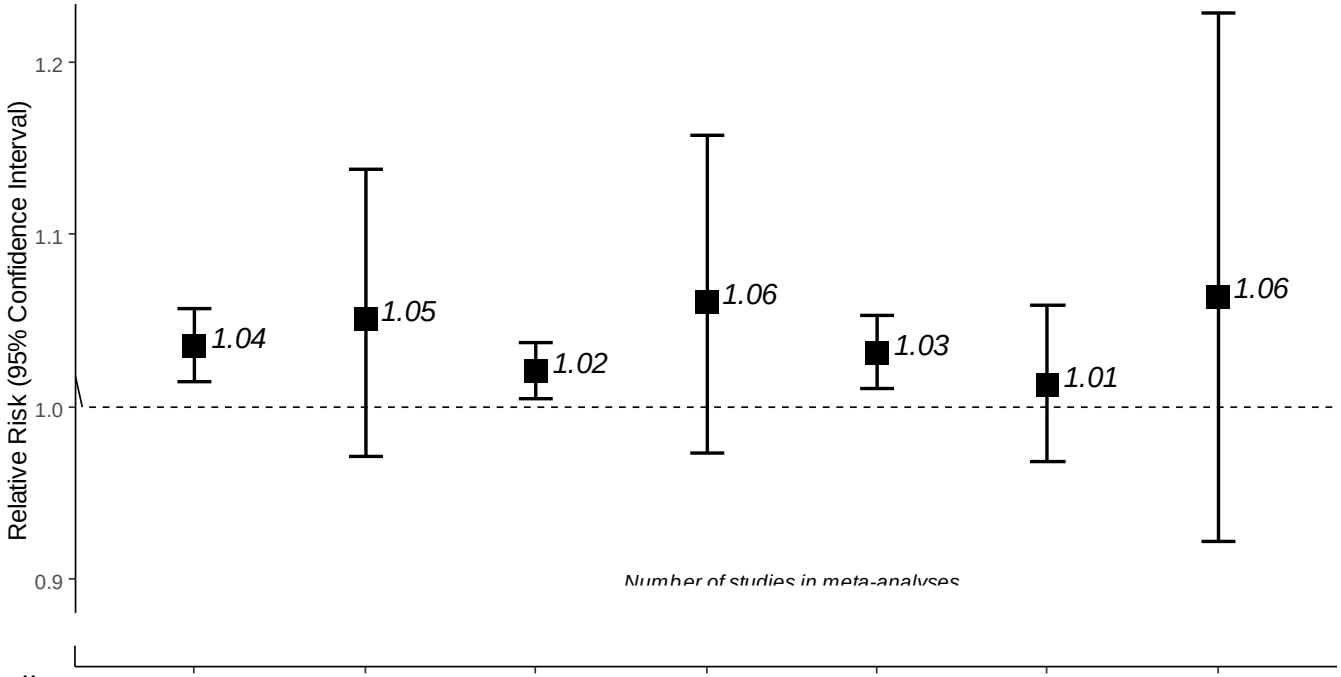
Institute for risk assessment sciences (IRAS)

Utrecht University

The Netherlands

RespiraMi IV meeting, June 17, 2022

Meta-analysis of associations between traffic-related air pollutants and all-cause mortality



Pollutant:	NO_2	NO_x	EC	PM_{10}	$PM_{2.5}$	$Cu-PM_{2.5}$	$Fe-PM_{2.5}$
Number of studies in meta-analysis:	11	5	11	6	12	3	3
Increments:	10 µg/m³	20 µg/m³	1 µg/m³	10 µg/m³	5 µg/m³	5 ng/m³	500 ng/m³

Effect estimates cannot be directly compared across the different traffic-related pollutants because the selected increments do not necessarily represent the same contrast in exposure. The individual pollutants are considered as indicators of the TRAP mixture.

Universiteit Utrecht

Confidence rating NO2 as example

Table 11.5. Confidence Rating in the Quality of the Body of Evidence for Traffic-Related Air Pollutants and All-Cause Mortality[a]

	High ++++ Moderate +++ Low ++ Very low +		Factors Decreasing Confidence (0 if no concern; – if serious concern to downgrade confidence)				Factors Increasing Confidence (0 if not present; + if sufficient to upgrade confidence)			
Pollutant	Study Design	Initial Confidence Rating (# studies)	Risk of Bias	Unexplained Inconsistency	Imprecision	Publication Bias	Monotonic Exposure–Response	Consideration of Residual Confounding	Consistency Across Populations	Final Confidence Rating
NO_2	Cohort	+++ ($N = 11$)	0	0	0	0	+	0	+	++++ (High)
	Rationale	Cohort design initially rated as moderate.	Few studies high RoB and robust effect estimates in low and moderate RoB studies.	High heterogeneity ($I^2 = 83\%$) due to magnitude not direction.	Sample size met, and confidence interval does not include unity.	No evidence found in plot and test.	Clear evidence of plausible shape of ERF (Cesaroni 2013; Crouse 2015; Dirgawati 2019; Hvidtfeldt 2019; Raaschou-Nielsen 2012).	Confounding in both directions possible.	Across geographic regions robust effect.	

Universiteit Utrecht

Risk of bias assessment by study

Domain	Low	Moderate	High
Confounding	11	4	5
Selection bias	17	2	1
Exposure	9	10	1
Outcome	20	0	0
Missing data	18	0	2
Selective reporting	20	0	0

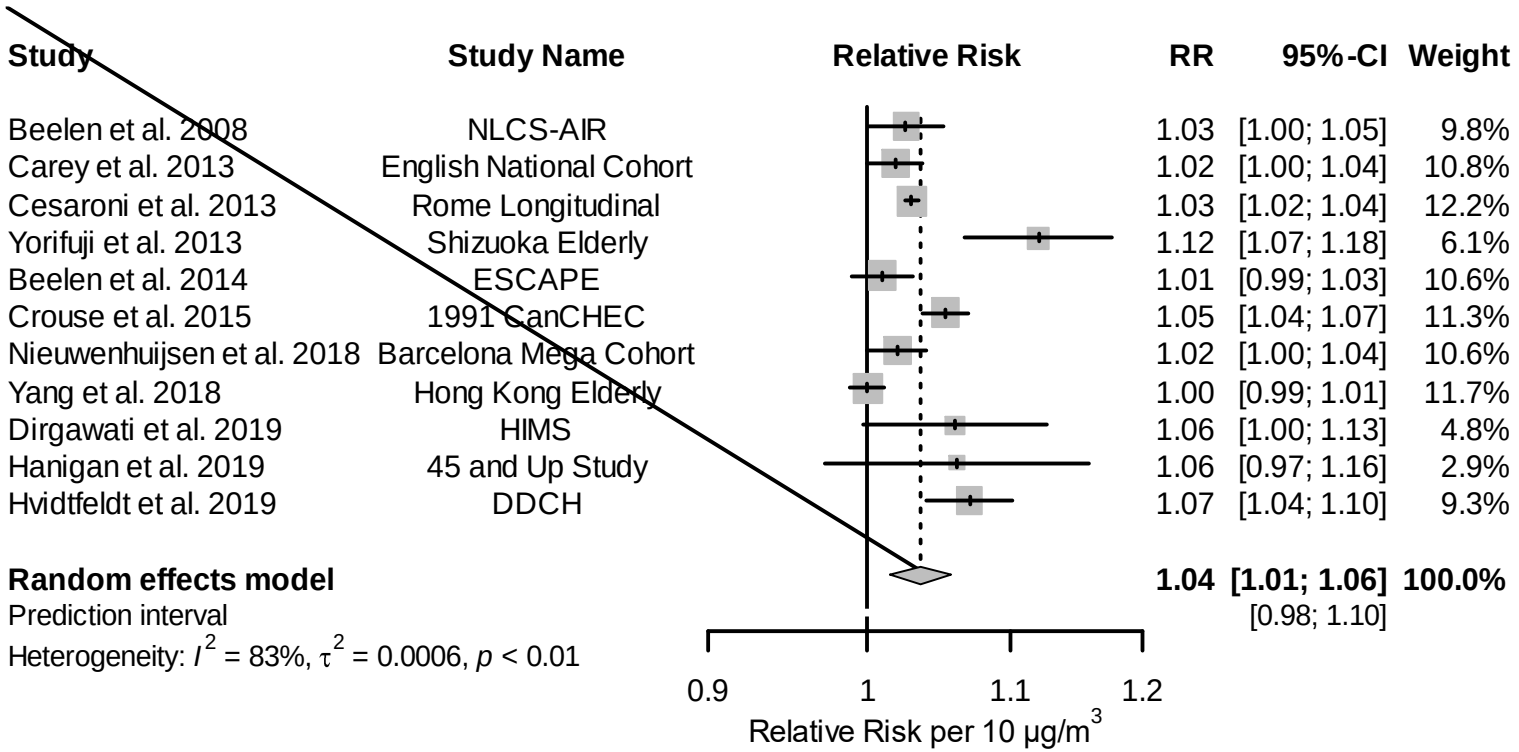
Effect estimates by Risk of bias confounding

NO_2 - total mortality by Risk of bias assessment on confounding

Study	Study Name	RR	95%-CI
Low/Moderate			
Beelen et al. 2008	NLCS-AIR	1.03	[1.00; 1.05]
Carey et al. 2013	English National Cohort	1.02	[1.00; 1.04]
Yorifuji et al. 2013	Shizuoka Elderly	1.12	[1.07; 1.18]
Beelen et al. 2014	ESCAPE	1.01	[0.99; 1.03]
Crouse et al. 2015	1991 CanCHEC	1.05	[1.04; 1.07]
Yang et al. 2018	Hong Kong Elderly	1.00	[0.99; 1.01]
Dirgawati et al. 2019	HIMS	1.06	[1.00; 1.13]
Hanigan et al. 2019	45 and Up Study	1.06	[0.97; 1.16]
Hvidtfeldt et al. 2019	DDCH	1.07	[1.04; 1.10]
Random effects model		**1.04**	**[1.01; 1.07]**
Heterogeneity: $I^2 = 86\%$, $\tau^2 = 0.0009$, $p < 0.01$			
High			
Cesaroni et al. 2013	Rome Longitudinal	1.03	[1.02; 1.04]
Nieuwenhuijsen et al. 2018	Barcelona Mega Cohort	1.02	[1.00; 1.04]

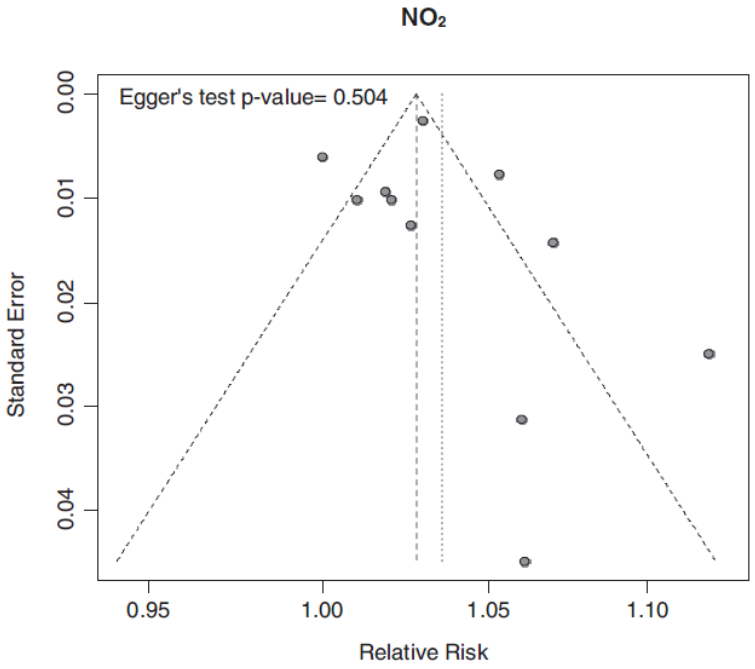
Universiteit Utrecht

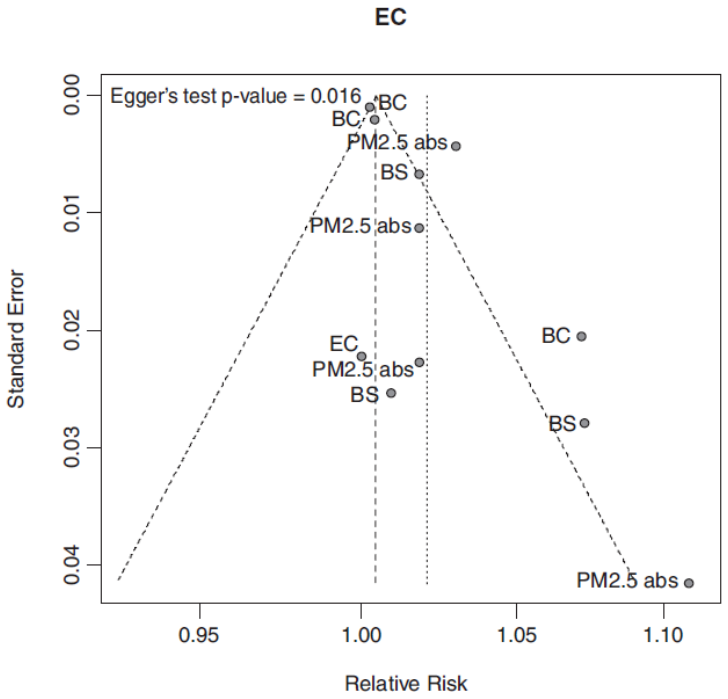
Meta-analysis NO_2 – All cause mortality



Universiteit Utrecht

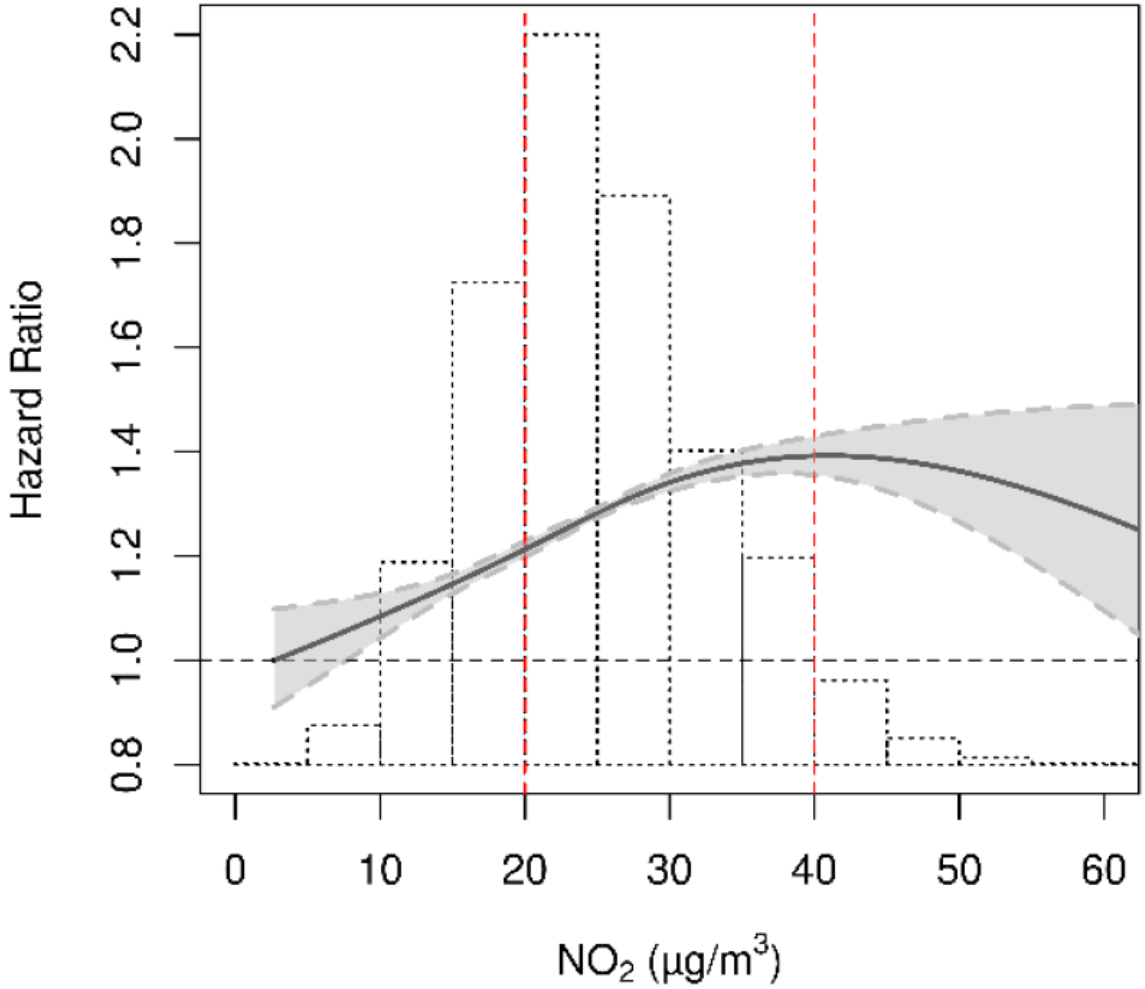
Publication bias or heterogeneity?





Universiteit Utrecht

Long-term NO2 exposure and mortality



Universiteit Utre

Final modified OHAT assessment for TRAP

- Upgrades for monotonic exposure response function (NO2, PM2.5, NOX and PM10) and consistency across regions (NO2)

• PM2.5, NO2, EC	high confidence
• NO2 and PM10	moderate confidence
• Cu, Fe	low confidence
• **TRAP combined**	**high confidence**

Problemen in GRADE type assessments

- Formule voor uiteindelijk oordeel, downgrade voor alle items zelfde (bv confounding en publikatie bias)
- Implementatie vergt veel keuzes die controversieel zijn (bv welke confounders, imprecision, hoe omgaan met RoB)
- Keuze voor initieele confidence
- Aanname dat clinical trial beter is, dubieus bij milieufactoren (ethisch, populatie)
- Voor miliefactoren: Alleen epidemiologie, geen toxicologie, mechanistische evidentie

Alternatieven

- IARC carcinogenicity determination (https://monographs.iarc.who.int/)
- US EPA causality determination: Integrated science assessment

Voor GGD medewerkers

- Goed om bewust te zijn van de gemaakte keuzes
- Niet blind kwaliteitsoordelen overnemen
- In systematic review nadenken over systematiek bv HEI narrative naast GRADE type