

# **Probability of Causation: SYNERGY sensitivity analysis**

**Lung cancer due to occupational asbestos exposure**

Javier Mancilla-Galindo      Susan Peters      Jenny Deng  
Henk van der Molen      Hans Kromhout      Lützen Portengen  
Roel Vermeulen      Dick Heederik

2025-08-14

**Asbestos Exposure accross studies**

Asbestos exposure summary by study and control source

Source of controls	Study name	Median exposure (ff/ml-years)	IQR (ff/ml-years)	Range (ff/ml-years)	Cases	Total
H	CAPUA	0.000	0 - 1.12	0 - 16.913	559	1071
H	INCO_Czech Republic	0.000	0 - 2.596	0 - 35.432	304	756
H	INCO_Hungary	0.000	0 - 3.688	0 - 64.6	391	696
H	INCO_Romania	0.000	0 - 0.62	0 - 15.954	179	404
H	INCO_Russia	0.709	0 - 4.064	0 - 26.571	599	1179
H	INCO_Slovakia	0.000	0 - 2.146	0 - 23.255	345	630
H	LUCA	0.000	0 - 0.875	0 - 6.933	280	562
H	PARIS	0.000	0 - 1.493	0 - 7.547	169	396
H	ROME	0.000	0 - 1.811	0 - 9.367	326	647
P	AUT	0.000	0 - 2.055	0 - 59.509	3180	6429
P	EAGLE	0.000	0 - 0.421	0 - 10.853	1908	3973
P	HdA	0.236	0 - 2.618	0 - 32.909	1004	2006
P	ICARE	0.000	0 - 0.445	0 - 15.236	2739	6188
P	INCO_Poland	0.000	0 - 1.007	0 - 24.183	793	1628
P	INCO_UK	0.076	0 - 0.372	0 - 5.889	441	1357
P	LUCAS	0.000	0 - 0.11	0 - 2.117	1014	3321
P	MONTREAL	0.000	0 - 0.231	0 - 4.94	1176	2681
P	MORGEN	0.000	0 - 0	0 - 6.942	43	158
P	TURIN	0.000	0 - 1.753	0 - 19.084	1086	2575
P	Toronto	0.000	0 - 0	0 - 2.378	365	1209

## Mixed effects model - cum0 with interaction term

A mixed-effects logistic regression model is used to estimate the PoC, with `asbestos_cum0` as the main explanatory variable, and adjusted for:

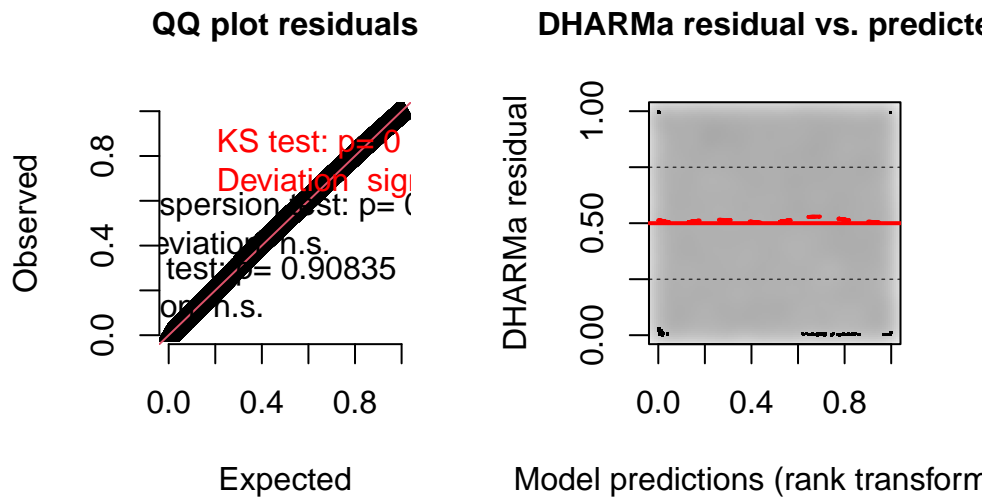
- The study source of participants (`study_name`)
- Age category (`agegroup`). The age groups are: <45, 45–49, 50–54, 55–59, 60–64, 65–69, 70–74, and > 74 years.
- Sex (`sex`), female and male.
- Smoking (`packyrs`), cigarette pack-years.
- Time since smoking cessation (`time_quit`). The categories are: current smokers; stopping smoking 2–7, 8–15, 16–25, and  $\geq 26$  years before interview/diagnosis; and never-smokers.

Adding random effects with a random intercept for each study source (`study_name`) and random slopes for the exposure (`asbestos_cum0`) within each study source.

An interaction term between `asbestos_cum0` and `source_controls` is included to assess whether the exposure-response relationship differs between hospital and population control studies.

```
mod_int <- glmer(status ~ asbestos_cum0 * source_controls + agegroup + sex +  
  packyrs + time_quit + (1 + asbestos_cum0 | study_name),  
  data      = df,  
  family    = binomial,  
  nAGQ      = 0,  
  control    = glmerControl(optimizer = "bobyqa")  
)
```

## DHARMa residual



	GVIF	Df	GVIF <sup>1/(2*Df)</sup>
asbestos_cum0	2.155903	1	1.468299
source_controls	1.316874	1	1.147551
agegroup	1.190896	7	1.012557
sex	1.235039	1	1.111323
packyrs	5.544476	1	2.354671
time_quit	5.959662	5	1.195425
asbestos_cum0:source_controls	2.416766	1	1.554595

[1] "I<sup>2</sup> for mod\_int (intercept): 3.32 %"

[1] "I<sup>2</sup> for asbestos slopes per study: 0.02 %"

# Asbestos Exposure-Response by Control Source in SYNERGY population

Model	Risk Increase per Fibre-year		Min Exposure for 50% PoC (fibre-years) <sup>1</sup>		Cases per 10,000 above 50% PoC in SYNERGY	
	Estimate	95% Prediction Interval	Point Estimate	Presumably Plausible	Point Estimate	Presumably Plausible
Population controls	9.3%	2.7% ; 16.2%	7.81	4.61	165	528
Hospital controls	1.9%	-4.3% ; 8.4%	37.35	8.56	0	97

<sup>1</sup>Probability of Causation (PoC)

Refer to the *save\_plots.R* script in the **R/scripts** directory for the relevant code used to generate the plots from this sensitivity analysis.

Model	Risk Increase per Fibre-year <sup>a</sup>		Min Exposure for 50% PoC (fibre-years) <sup>b</sup>		Cases per 10,000 above 50% PoC in SYNERGY	
	Estimate	95% Prediction Interval	Point Estimate	Presumably Plausible	Point Estimate	Presumably Plausible
<b>Meta-regression of published studies (ECHA)</b>						
Spline	0.9%	0% ; 2.6%	81.71	26.09	0	3
<b>Meta-analysis of individual participant data (SYNERGY)</b>						
Population-average	6.8%	0% ; 17.5%	10.59	4.31	115	845
<b>Sensitivity analysis stratified by source of controls (SYNERGY)</b>						
Population controls	9.3%	2.7% ; 16.2%	7.81	4.61	165	528
Hospital controls	1.9%	0% ; 8.4%	37.35	8.56	0	97

<sup>a</sup>Negative values truncated to 0 due to monotonicity assumption. For spline models: instantaneous slope at 1.5 fibre-years (median exposure in cases).

<sup>b</sup>Probability of Causation (PoC) of 50% corresponds to a doubling in lung cancer risk (i.e.,  $RR = 2$ ) for the point estimate and the upper 95% prediction interval for the presumably plausible threshold.

## Summary of coefficients

Coefficients from both the ECHA and the SYNERGY study are summarized with the following code.

## Package References

For specific information on the operating system, R version, and R package versions used, please refer to the **R/session** folder in the GitHub repository.

- Auguie B (2017). *gridExtra: Miscellaneous Functions for “Grid” Graphics*. doi:10.32614/CRAN.package.gridExtra <https://doi.org/10.32614/CRAN.package.gridExtra>, R package version 2.3, <https://CRAN.R-project.org/package=gridExtra>.
- Bates D, Mächler M, Bolker B, Walker S (2015). “Fitting Linear Mixed-Effects Models Using lme4.” *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01 <https://doi.org/10.18637/jss.v067.i01>.
- Bates D, Maechler M, Jagan M (2025). *Matrix: Sparse and Dense Matrix Classes and Methods*. doi:10.32614/CRAN.package.Matrix <https://doi.org/10.32614/CRAN.package.Matrix>, R package version 1.7-3, <https://CRAN.R-project.org/package=Matrix>.
- Csárdi G, Hester J (2025). *pak: Another Approach to Package Installation*. doi:10.32614/CRAN.package.pak <https://doi.org/10.32614/CRAN.package.pak>, R package version 0.9.0, <https://CRAN.R-project.org/package=pak>.
- Fox J, Weisberg S (2019). *An R Companion to Applied Regression*, Third edition. Sage, Thousand Oaks CA. <https://www.john-fox.ca/Companion/>.
- Fox J, Weisberg S, Price B (2022). *carData: Companion to Applied Regression Data Sets*. doi:10.32614/CRAN.package.carData <https://doi.org/10.32614/CRAN.package.carData>, R package version 3.0-5, <https://CRAN.R-project.org/package=carData>.
- Gohel D, Skintzos P (2025). *flextable: Functions for Tabular Reporting*. doi:10.32614/CRAN.package.flextable <https://doi.org/10.32614/CRAN.package.flextable>, R package version 0.9.9, <https://CRAN.R-project.org/package=flextable>.
- Grolemund G, Wickham H (2011). “Dates and Times Made Easy with lubridate.” *Journal of Statistical Software*, 40(3), 1-25. <https://www.jstatsoft.org/v40/i03/>.
- Hartig F (2024). *DHARMA: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. doi:10.32614/CRAN.package.DHARMA <https://doi.org/10.32614/CRAN.package.DHARMA>, R package version 0.4.7, <https://CRAN.R-project.org/package=DHARMA>.
- Iannone R, Cheng J, Schloerke B, Hughes E, Lauer A, Seo J, Brevoort K, Roy O (2025). *gt: Easily Create Presentation-Ready Display Tables*. doi:10.32614/CRAN.package.gt <https://doi.org/10.32614/CRAN.package.gt>, R package version 1.0.0, <https://CRAN.R-project.org/package=gt>.
- Lüdtke D (2025). *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.9.0, <https://CRAN.R-project.org/package=sjPlot>.
- Makowski D, Lüdtke D, Patil I, Thériault R, Ben-Shachar M, Wiernik B (2023). “Automated Results Reporting as a Practical Tool to Improve Reproducibility and Methodological Best Practices Adoption.” *CRAN*. <https://easystats.github.io/report/>.
- Müller K, Wickham H (2025). *tibble: Simple Data Frames*. doi:10.32614/CRAN.package.tibble <https://doi.org/10.32614/CRAN.package.tibble>, R package version 3.3.0, <https://CRAN.R-project.org/package=tibble>.

- [//CRAN.R-project.org/package=tibble](https://CRAN.R-project.org/package=tibble).
- R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
  - Rich B (2023). *table1: Tables of Descriptive Statistics in HTML*. doi:10.32614/CRAN.package.table1 <https://doi.org/10.32614/CRAN.package.table1>, R package version 1.4.3, <https://CRAN.R-project.org/package=table1>.
  - Rinker TW, Kurkiewicz D (2018). *pacman: Package Management for R*. version 0.5.0, <http://github.com/trinker/pacman>.
  - Sera F, Armstrong B, Blangiardo M, Gasparrini A (2019). “An extended mixed-effects framework for meta-analysis.” *Statistics in Medicine*, DOI: 10.1002/sim.8362. [http://www.ag-myresearch.com/2019\\_sera\\_statmed.html](http://www.ag-myresearch.com/2019_sera_statmed.html).
  - Venables WN, Ripley BD (2002). *Modern Applied Statistics with S*, Fourth edition. Springer, New York. ISBN 0-387-95457-0, <https://www.stats.ox.ac.uk/pub/MASS4/>.
  - Wickham H (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, <https://ggplot2.tidyverse.org>.
  - Wickham H (2023). *conflicted: An Alternative Conflict Resolution Strategy*. doi:10.32614/CRAN.package.conflicted <https://doi.org/10.32614/CRAN.package.conflicted>, R package version 1.2.0, <https://CRAN.R-project.org/package=conflicted>.
  - Wickham H (2023). *forcats: Tools for Working with Categorical Variables (Factors)*. doi:10.32614/CRAN.package.forcats <https://doi.org/10.32614/CRAN.package.forcats>, R package version 1.0.0, <https://CRAN.R-project.org/package=forcats>.
  - Wickham H (2023). *stringr: Simple, Consistent Wrappers for Common String Operations*. doi:10.32614/CRAN.package.stringr <https://doi.org/10.32614/CRAN.package.stringr>, R package version 1.5.1, <https://CRAN.R-project.org/package=stringr>.
  - Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H (2019). “Welcome to the tidyverse.” *Journal of Open Source Software*, 4(43), 1686. doi:10.21105/joss.01686 <https://doi.org/10.21105/joss.01686>.
  - Wickham H, François R, Henry L, Müller K, Vaughan D (2023). *dplyr: A Grammar of Data Manipulation*. doi:10.32614/CRAN.package.dplyr <https://doi.org/10.32614/CRAN.package.dplyr>, R package version 1.1.4, <https://CRAN.R-project.org/package=dplyr>.
  - Wickham H, Henry L (2025). *purrr: Functional Programming Tools*. doi:10.32614/CRAN.package.purrr <https://doi.org/10.32614/CRAN.package.purrr>, R package version 1.1.0, <https://CRAN.R-project.org/package=purrr>.
  - Wickham H, Hester J, Bryan J (2024). *readr: Read Rectangular Text Data*. doi:10.32614/CRAN.package.readr <https://doi.org/10.32614/CRAN.package.readr>, R package version 2.1.5, <https://CRAN.R-project.org/package=readr>.
  - Wickham H, Vaughan D, Girlich M (2024). *tidyr: Tidy Messy Data*. doi:10.32614/CRAN.package.tidyr <https://doi.org/10.32614/CRAN.package.tidyr>, R package version 1.3.1, <https://CRAN.R-project.org/package=tidyr>.