



HAL
open science

Lung cancer mortality in the European cohort of titanium dioxide workers: a reanalysis of the exposure–response relationship

Irina Guseva Canu, Alan Gaillen-Guedy, Ahti Antilla, Sandrine Charles, Sandrine Fraize-Frontier, Danièle Luce, Damien Martin Mcelvenny, Franco Merletti, Cecile Michel, Eero Pukkala, et al.

► To cite this version:

Irina Guseva Canu, Alan Gaillen-Guedy, Ahti Antilla, Sandrine Charles, Sandrine Fraize-Frontier, et al.. Lung cancer mortality in the European cohort of titanium dioxide workers: a reanalysis of the exposure–response relationship. *Occupational and Environmental Medicine*, 2022, 79 (9), pp.637-640. 10.1136/oemed-2021-108030 . anses-03663925

HAL Id: anses-03663925

<https://anses.hal.science/anses-03663925>

Submitted on 28 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Lung cancer mortality in the European cohort of titanium dioxide workers: a re-analysis of**
2 **the exposure-response relationship**

3 Irina Guseva Canu¹, Alan Gaillen-Guedy¹, Ahti Anttila², Sandrine Charles³, Sandrine Fraize-
4 Frontier³, Danièle Luce⁴, Damien McElvenny^{5,6}, Franco Merletti⁷, Cecile Michel³, Eero
5 Pukkala^{2,8}, Mary Schubauer-Berigan⁹, Kurt Straif^{10,11}, Pascal Wild¹, David B. Richardson¹²

6 Author affiliations:

7 1-Center for Primary Care and Public Health (Unisanté), University of Lausanne, Lausanne,
8 Switzerland

9 2-Finnish Cancer Registry - Institute for Statistical and Epidemiological Cancer Research,
10 Helsinki, Finland

11 3- French Agency for Food, Environmental and Occupational Health & Safety (ANSES), Maisons-
12 Alfort, France

13 4- Inserm, Univ Rennes, Univ Angers, EHESP, Irset (Institut de recherche en santé, environnement
14 et travail) – UMR_S 1085, Pointe-à-Pitre, France.

15 5-Institute of Occupational Medicine, Edinburgh, United Kingdom

16 6-University of Manchester, Manchester, United Kingdom

17 7-Cancer Epidemiology Unit, Department of Medical Sciences, University of Turin and CPO-
18 Piemonte, Turin, Italy

19 8-Faculty of Social Sciences, Tampere University, Tampere, Finland

20 9-International Agency for Research on Cancer, Lyon, France

21 10- ISGlobal, Barcelona, Spain

1 11- Boston College, Chestnut Hill, MA, USA

2 12-Department of Epidemiology, School of Public Health, University of North Carolina
3 Chapel Hill, NC, USA

4

5 Running Head: Titanium dioxide exposure and lung cancer

6 Correspondence to:

7 Prof. Irina Guseva Canu

8 Unisanté, Departement of Occupational and Environmental Health

9 Route de la Cornishe, 2

10 1066 Epalinges-Lausanne,

11 Switzerland

12 irina.guseva-canu@unisante.ch

13 Disclaimer : Where authors are identified as personnel of the International Agency for Research
14 on Cancer / World Health Organization, the authors alone are responsible for the views expressed
15 in this article and they do not necessarily represent the decisions, policy or views of the
16 International Agency for Research on Cancer / World Health Organization.

17

18

19

20

21

22

1 ABSTRACT

2 Objectives: Animal bioassays have demonstrated convincing evidence of the potential
3 carcinogenicity to humans of titanium dioxide (TiO₂), but limitations in cohort studies have been
4 identified, among which is the healthy worker survivor effect (HWSE). We aimed to address this
5 bias in a pooled study of four cohorts of TiO₂ workers.

6 Methods: We re-analysed data on respirable TiO₂ dust exposure and lung cancer mortality among
7 7341 male workers employed in TiO₂ production in Finland, France, United Kingdom and Italy
8 using the parametric g-formula, considering three hypothetical interventions: setting annual
9 exposures at 2.4 (U.S. occupational exposure limit), 0.3 (German limit), and 0 mg/m³ for 25 and
10 35 years.

11 Results: The HWSE was evidenced. Taking this into account, we observed a positive association
12 between lagged cumulative exposure to TiO₂ and lung cancer mortality. The estimated number of
13 lung cancer deaths at each age group decreased across increasingly stringent intervention levels.
14 At age 70 years, the estimated number of lung cancer deaths expected in the cohort after 35-year
15 exposure was 293 for exposure set at 2.4 mg/m³, 235 for exposure set at 0.3 mg/m³, and 211 for
16 exposure set at 0 mg/m³.

17 Conclusion: This analysis shows that HWSE can hide an exposure-response relationship. It also
18 shows that TiO₂ epidemiological data could demonstrate an exposure-effects relationship if
19 analysed appropriately. More epidemiological studies and similar re-analyses of existing cohort
20 studies are warranted to corroborate the human carcinogenicity of TiO₂. This human evidence,
21 when combined with the animal evidence, strengthens the overall evidence of carcinogenicity of
22 TiO₂.

23 KEY WORDS

1 Occupational exposure; g-formula; causal inference; retrospective cohort

2 INTRODUCTION

3 There is substantial interest in human evidence regarding the carcinogenicity of titanium dioxide
4 (TiO_2), an odorless white pigment and opacifying agent widely used since the 1920s. TiO_2 is
5 classified as possibly carcinogenic to humans by the International Agency for Research on Cancer
6 based on sufficient evidence from cancer bioassay studies and inadequate evidence from human
7 cancer studies. Since 2020, TiO_2 is also classified as a suspected human carcinogen by inhalation
8 in the European Union according to Regulation n°1272/2008.

9 Prior large cohorts of TiO_2 workers reported increased mortality from lung cancer, but failed to
10 observe an exposure–response relationship with cumulative exposure to TiO_2 ,[1] except in a
11 French cohort of TiO_2 workers.[2] One of the key limitations noted in the occupational cohorts
12 was the potential healthy worker survivor effect (HWSE) which can mask association between
13 cumulative exposure and lung cancer mortality.[3] For instance, in the pooled European study of
14 TiO_2 workers, a key study on this topic, no evidence of an association between respirable TiO_2
15 exposure and lung cancer mortality was observed despite the excess of lung cancer mortality
16 among male TiO_2 workers as compared to the general population (standardised mortality
17 ratio=1.23; 95% confidence interval (95% CI)=1.10-1.38).[4]

18 In this study, we re-analysed a subset of the pooled European cohort of TiO_2 workers,[4] to
19 examine the evidence of HWSE and the exposure-response relationship between cumulative
20 exposure to TiO_2 and lung cancer mortality. We implemented the g-computation algorithm
21 formula (g-formula) recommended for statistical analysis of cohort data in the presence of time-
22 varying confounders affected by prior exposure, typical of HWSE.[5] The g-formula has been
23 proven to be an essential method for estimating human health effects of exposures and

1 interventions on exposures in such scenarios.[5] Therefore, we applied it to assess the effect of
2 three hypothetical interventions of TiO₂ exposure limitation.

3 METHODS

4 *Study population*

5 The original pooled European cohort included workers who had been employed at least 1 month
6 in one of 11 TiO₂ production factories in six European countries (Finland, France, Italy, Norway,
7 Germany, and the UK).[4] All female workers and male workers with missing death certificates
8 and/or lacking quantitative exposure estimates were excluded from the exposure-response analysis
9 of this cohort.[4] We used the same criteria of worker inclusion/exclusion as in original cohort,
10 but restricted the study to four countries (Finland, France, Italy, and the UK), for which data were
11 still available and ethical approvals obtained.

12 *Exposure assessment*

13 The factories produced mainly pigment-grade TiO₂, although TiO₂ form (e.g., particle size and
14 crystalline phase) was ignored.[6] Information on demographic and employment characteristics
15 was collected from factories' records describing date of birth, sex, race, and dates of hire, job or
16 department change, and termination. Estimated cumulative occupational exposure to respirable
17 TiO₂ dust was derived from job title and work history.[6] Exposure assessments were carried out
18 at the level of occupational titles for each plant for discrete time periods throughout the history of
19 plant operations. Lists of occupational titles were compiled and coded for each factory. Exposure
20 measurement data were obtained from company files along with information such as the area(s) of
21 the plant where measurements were made, the presence of any local ventilation, the type of the
22 materials being handled and the purpose of sampling. These were then linked to the work history
23 of each individual in the cohort to provide exposure estimates.

1 *Statistical analysis*

2 The lengths of follow-up varied between countries and ranged from 1950–1972 until 1997–2001.

3 The primary outcome of interest was death for which the underlying cause was attributed to
4 cancers of the trachea, bronchus and lung (ICD-9 code 162).

5 A data tabulation of person-periods and events was constructed with one record for each person-
6 year of observation from date of entry into the analysis until end of follow-up or administrative
7 censoring of workers alive at age 90 years. Using the observed data, we fitted logistic regression
8 models for the probability of the outcome of interest, for the probability of remaining at work, and
9 for the probability of dying from a competing cause, as a function of covariates and estimated
10 exposure. The cumulative TiO₂ exposure was 10-year lagged (Supplementary material Figure S1
11 and Technical Appendix 1).

12 The g-formula was implemented by a Monte Carlo simulation based on the regression model
13 estimates of the probability of termination of employment and death.[5, 7] Ten Monte Carlo
14 samples per exposure scenario were drawn randomly from the observed cohort and the estimated
15 parameters from the parametric models to recreate the study data for each person in the sample
16 under specified exposure intervention. Three hypothetical interventions were considered: setting
17 workers' annual exposures to 2.4 mg/m³, 0.3 mg/m³ (the currently recommended TiO₂
18 occupational exposure limits in the US [8] and Germany [9], respectively), and zero exposure. For
19 each intervention, we assumed two possible exposure durations: 25 and 35 years and estimated the
20 expected lung cancer mortality at 60, 70, 80, and 90 years of age. The associated 95%-CIs were
21 calculated using bootstrap samples.

22 RESULTS

1 The cohort included 7341 workers (Table S2). At the end of the follow-up, 139 lung cancer deaths
2 were observed. The presence of the HWSE was evident (Figure S1). Being in employment reduced
3 the risk of lung cancer mortality (OR=0.14, 95%CI=0.08-0.22) and the probability of leaving the
4 employment increased as a function of TiO₂ exposure (OR=3.55 95%CI=2.82-4.46). The OR of
5 lung cancer death associated with lagged cumulative exposure to TiO₂ was estimated at 1.03 per
6 1 mg/m³-year (95%CI=0.99-1.07), after adjustment for the employment status in previous and
7 current years and employment duration.

8 G-estimates of lung cancer mortality, derived under the three hypothetical interventions, are shown
9 in Table 1. The estimated number of lung cancer deaths at 60, 70, 80, and 90 years of age all
10 decrease across the three interventions considered and for both exposure durations.

11 DISCUSSION

12 This re-analysis provides the first evidence of an exposure-response relationship between TiO₂
13 cumulative exposure and lung cancer mortality using the parametric g-formula. Adjustment of a
14 standard regression model for employment status orexposure duration is not sufficient for
15 complete HWSE correction.[3] However, rank ordering of lung cancer deaths across levels of the
16 intervention estimated by g-formula is consistent with a positive exposure-response association
17 between TiO₂ and lung cancer (3% per 1 mg/m³-year of respiratory TiO₂).

18 A limitation of the g-formula is the g-null paradox. The g-formula may be guaranteed some degree
19 of model misspecification if there is treatment-confounder feedback and the sharp causal null
20 hypothesis (i.e., the intervention has no effect on any individual's outcome at any time) holds.[10]

21 To address concern that model misspecification may lead to bias, we followed recent guidance
22 [11] to avoid overly parsimonious models for the components of the g-formula in this analysis.

1 A causal interpretation of the findings requires certain statistical assumptions, including
2 consistency, positivity, exchangeability.[12] The consistency assumption may be challenging
3 given the complexity of historical exposure conditions and the various TiO₂ forms, the latter being
4 insufficiently documented. We nevertheless believe that exposure contrasts are defined here well
5 enough to support meaningful inference regarding TiO₂'s effect. The positivity hypothesis (i.e.
6 observations on exposed and unexposed workers through covariate levels) was difficult to confirm
7 as 82% of workers were exposed. However, the positive exposure-response relationship based on
8 continuous cumulative lagged exposure to TiO₂ supports these assumptions.

9 The exchangeability assumption (i.e., no unmeasured confounding) is challenging given the
10 limitations of available data on smoking. The most complete smoking data were available for the
11 French cohort and showed no effect of adjustment for smoking on estimates of TiO₂-lung cancer
12 mortality associations.[2] Moreover, we were able to assess some other occupational co-exposures
13 that are suspected or known lung carcinogens, including asbestos, welding fumes, and other
14 mineral dusts; adjustment for these exposures had no effect on the association with TiO₂.

15 A confounder typically of concern in occupational cohort mortality studies is the HWSE.[3] In the
16 current study, this source of confounding was addressed by the g-formula. In prior analyses of
17 these data using standard multivariable regression method, the association was likely masked by
18 the HWSE and potentially the exposure misclassification in the early years of follow-up. Prior to
19 undertaking the g-formula analysis, we assessed the most important HWSE components in these
20 data and confirmed their presence. Moreover, prior reports of SMR analyses indicated patterns
21 consistent with the HWSE[3] and in the analysis of the French cohort, duration of employment
22 was negatively associated with lung cancer mortality.[2] These underscore the relevance of g-
23 methods in these cohorts and future investigations of TiO₂'s effects on human health.

1 The estimates with overlapping CI are likely due to the limited statistical power in this study,
2 which might be addressed by additional follow-up of these cohorts. More epidemiological studies
3 and similar re-analyses of updated existing cohort studies are warranted to corroborate the human
4 carcinogenicity of TiO₂. This human evidence, when combined with the animal evidence,
5 strengthens the overall evidence of carcinogenicity of TiO₂.

6 ACKNOWLEDGEMENTS

7 None.

8 CONTRIBUTORS

9 IGC designed and conducted this study and drafted the manuscript, AG-G, DBR, and PW
10 conducted statistical analyses. KS and MSB centralised the data, facilitated data access, and
11 obtained IARC ethical approval. SC, SFF, and CM contributed to study coordination and French
12 ethical approval. All authors discussed the study methodology, read the manuscript, critically
13 reviewed it and agreed on the final version.

14 FUNDING

15 This work was supported by the ANSES/IST Number 2017-CRD-18.

16 COMPETING INTERESTS

17 None declared.

18 PATIENT CONSENT FOR PUBLICATION

19 Not required.

20 ETHICS APPROVAL

21 The study was authorised by the IARC ethics Committee (IEC project number 18-32). The use of
22 the data for the French cohort was approved by the French Data Protection Authority (CNIL),
23 Authorisation No 999250. The use of the data for the UK cohort was approved by the EPS ethics

1 committee of Heriot-Watt University (approval number 19/EA/JC/1). The use of the data for the
2 Italian cohort was approved by the Comitato Etico Interaziendale A.S.L Citta di Torino (approval
3 number CS2/1250). The use of the data for the Finnish cohort was approved by Statistics Finland
4 (approval number TK-53-988-19).

5

6 DATA AVAILABILITY STATEMENT

7 Data may be obtained from a third party and are not publicly available.

8

9 KEY MESSAGES

10 **What is already known about this subject?**

11 Titanium dioxide (TiO₂) is classified by the International Agency for Research on Cancer as
12 possibly carcinogenic to humans. In 2020, the European Chemicals Agency classified TiO₂ under
13 all forms as suspected human carcinogen by inhalation. Prior large cohorts of TiO₂ workers
14 reported increased mortality from lung cancer but failed to observe an exposure–response
15 relationship with cumulative exposure to TiO₂, except in a French cohort of TiO₂ workers. A
16 concern of potential healthy worker survivor effect (HWSE) has been raised.

17 **What are the new findings?**

18 We re-analysed data from the European cohort of TiO₂ workers and found strong evidence of the
19 HWSE. Taking this into account, a positive exposure–response relationship with 10-years lagged
20 cumulative exposure to TiO₂ was observed. The estimated number of lung cancer deaths at each
21 age group decreased across increasingly stringent exposure limits.

22 **How might this impact on policy or clinical practice in the foreseeable future?**

23 This analysis shows that HWSE can hide exposure–response relationship. It also shows that TiO₂
24 epidemiological data could demonstrate an exposure–effects relationship if analysed appropriately.

1 This human evidence, when combined with the animal evidence, strengthens the overall evidence
2 of carcinogenicity of TiO₂.

3

4 REFERENCES

- 5 1. Guseva Canu I, Fraize-Frontier S, Michel C, Charles S. Weight of epidemiological evidence
6 for titanium dioxide risk assessment: current state and further needs. *J Expo Sci Environ*
7 *Epidemiol.* 2020;30:430–5.
- 8 2. Guseva Canu I, Gaillen-Guedy A, Wild P, Straif K, Luce D. Lung cancer mortality in the
9 French cohort of titanium dioxide workers: some aetiological insights. *Occup Environ Med.*
10 2020;77(11):795-7.
- 11 3. Buckley JP, Keil AP, McGrath LJ, Edwards JK. Evolving methods for inference in the
12 presence of healthy worker survivor bias. *Epidemiology.* 2015;26(2):204-12.
- 13 4. Boffetta P, Soutar A, Cherrie JW, Granath F, Andersen A, Anttila A, et al. Mortality among
14 workers employed in the titanium dioxide production industry in Europe. *Cancer Causes*
15 *Control.* 2004;15(7):697-706.
- 16 5. Hernan MA, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol*
17 *Community Health.* 2006;60(7):578-86.
- 18 6. Boffetta P, Soutar A, Weiderpass E, Cherrie J, W. G, Andersen A, et al. Historical cohort
19 study of workers employed in the titanium dioxide production industry in Europe. Results of
20 mortality follow-up. Final report. Stockholm, Sweden: Department of Medical Epidemiology,
21 Karolinska Institute 2003.
- 22 7. Naimi AI, Cole SR, Kennedy EH. An introduction to g methods. *Int J Epidemiol.*
23 2017;46(2):756-62.
- 24 8. US-NIOSH. Current Intelligence Bulletin 63. Occupational Exposure to Titanium
25 Dioxide 2011. 140 p.
- 26 9. MAK. Titanium dioxide (respirable fraction). In: The MAK-collection for occupational health
27 and safety. 2009:1–66.
- 28 10. Robins JM, Wasserman L, editors. Estimation of effects of sequential treatments by
29 reparameterizing directed acyclic graphs. The Thirteenth Conference on Uncertainty in
30 Artificial Intelligence; 1997 San Francisco.
- 31 11. McGrath S, Young JG, Hernan MA. Revisiting the g-null paradox. *arXiv.* 2021:1-16.
- 32 12. Cole SR, Hernan MA. Fallibility in estimating direct effects. *Int J Epidemiol.* 2002;31:163-
33 5.

34

35

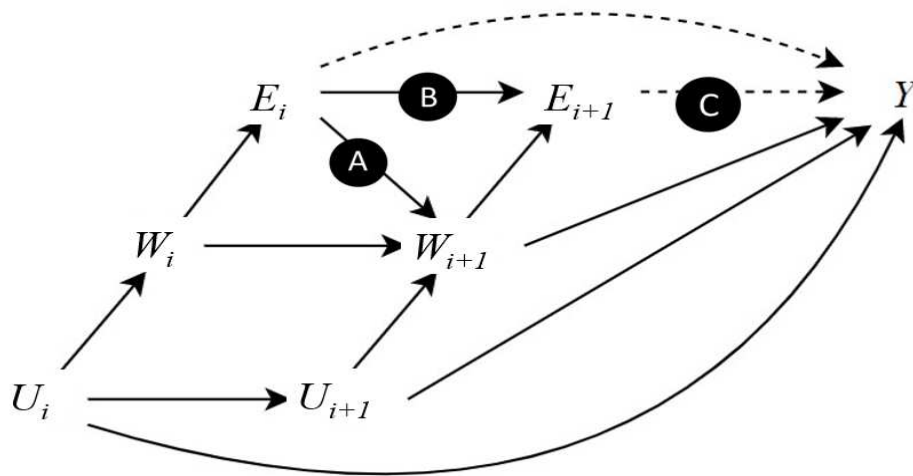
1 Table 1. Estimated cumulative lung cancer mortality at age 60, 70, 80 and 90 years under three
 2 hypothetical interventions on the TiO₂ exposure and assuming two durations of this exposure.
 3 European TiO₂ worker cohort, 1955-1990.

4

Attained age (in years)	Cumulative number of lung cancer deaths (95%CI) assuming 25-year exposure			Cumulative number of lung cancer deaths (95%CI) assuming 35-year exposure		
	Set		Set	Set		Set
	exposure to 2.4 mg/m ³	exposure to 0.3 mg/m ³	exposure to 0.0 mg/m ³	exposure to 2.4 mg/m ³	exposure to 0.3 mg/m ³	exposure to 0.0 mg/m ³
60	50 (46-55)	38 (33-44)	39 (34-45)	98 (94-102)	79 (72-86)	74 (68-80)
70	134 (130-139)	105 (97-112)	95 (86-104)	293 (288-299)	235 (226-243)	211 (200-223)
80	168 (162-174)	128 (121-136)	116 (107-125)	417 (407-427)	329 (320-337)	299 (281-316)
90	169 (163-174)	129 (121-136)	116 (108-125)	423 (412-433)	333 (325-342)	303 (286-321)

5
6

1 ONLINE SUPPLEMENTARY MATERIAL



2

3 Figure S1. Directed acyclic graph representing structural relationships underlying the healthy
4 worker survivor effect.

5

6 This graph represents the evolution over a worker's lifetime of his or her employment status (1 if
7 employed, 0 if terminated), exposure and mortality from lung cancer (1 if the cause of death is
8 lung cancer, 0 for any other situation) as well as the presumed causal links between these factors.

9 Dashed lines show causal effect of the exposure. A, B, and C represent relationships necessary to

10 model for implementation of the g-formula to control for potential confounding bias due to the

11 healthy worker survival effect. E_i corresponds to exposure to TiO_2 in year i ; W_i corresponds to

12 employment status and U_i to health status in year i . Finally, Y denotes mortality from lung cancer.

13 The unmeasured health status U_i in year i influences employment status in the same year and health

14 status U_{i+1} in the next year. Similarly, W_i influences E_i and W_{i+1} . Finally, E_i affects E_{i+1} .

15 Note that this is a simplified directed acyclic graph showing only two time intervals and omitting

16 the baseline covariates.

1 **Technical Appendix 1.**

2

3 Given the objectives of the project, the choice of the g-formula seemed the most judicious, as it
4 is the only method allowing the evaluation of the effect of hypothetical interventions. To apply
5 this method we began by modeling the different presumed causal associations presented in Figure
6 S1: the association between employment status and survival, the association between previous
7 exposure to TiO₂ and current employment status. The third key component association (i.e., an
8 association between employment status and subsequent exposure) is considered deterministic,
9 because the end of employment leads to the end of exposure. For this, the following models were
10 applied to the observed data in the study population (pooled cohort of four countries, n=7341):

- 11 1. A logistic regression model to predict the probability of job termination for a given person-
12 year (W_{i+1} as a function of W_i and E_i);
- 13 2. A logistic regression model to predict exposure in that year (E_{i+1} as a function of W_i and E_i
14 when $W_{i+1} = 1$) and a linear regression model to predict exposure level (as a function of
15 active employment and exposure);
- 16 3. A logistic regression model to predict lung cancer death as a function of employment status
17 and exposure (Y as a function of E_i);
- 18 4. A logistic regression model for other causes of death.

19 The beta coefficients of all variables included in each of these models are provided in the
20 technical report of the study¹, available upon request to the corresponding author.

¹ Irina Guseva Canu, Alan Gaillen-Guedy, Ahti Anttila, Danièle Luce, Damien McElvenny, Franco Merletti, Eero Pukkala, Mary Schubauer-Berigan, Kurt Straif, Pascal Wild, David B. Richardson. Rapport des résultats du projet scientifique « RealyTi » Reanalysis of human TiO₂ data. Troisième version. 22 décembre 2020. 66p.

1 In a second step, we used the results of these models to estimate the potential effect of different
2 exposure scenarios. The results of this step were expressed in terms of estimated number of deaths
3 per attained age for each hypothetical intervention on the exposure.

4 We considered three hypothetical interventions:

5 - A first intervention sets the annual respiratory TiO_2 exposure at 2.4 mg/m^3

6 - A second intervention sets the annual respiratory TiO_2 exposure at 0.3 mg/m^3

7 - A third intervention sets the exposures of all subjects during their entire career to zero

8 These first two values correspond to the currently recommended OELs for TiO_2 fine dust
9 (respirable fraction) in the USA and Germany, respectively.

10 The occupational careers, dates of death and dates of lung cancer death of the cohort participants
11 were resampled (Monte Carlo) 10 times for each scenario, resulting in 10 pseudo-cohorts for each
12 scenario.

13 In practice, we retained information on the baseline variables for each participant (identification
14 number, year of hire, age at hire, exposure at the start of follow-up, center identification code in
15 each country) and drew, year by year, exposure status, cumulative exposure level fixed by
16 exposure scenario lung cancer death (yes/no), and, if applicable, death by another cause of death
17 (concurrent causes).

18 Specifically, from year $m=1$, employment status is assigned using the conditional probability,
19 estimated from the parametric employment status model. If the person is employed, the exposure
20 of this person is defined at the specified level by a new simulation. Then, the probability of lung
21 cancer and competing causes of death is estimated based on the joint distribution of exposure and
22 covariates. A binary indicator for each outcome is drawn from a Bernoulli distribution with the

1 associated probability. If the individual is still alive at the end of year m , the process is repeated
2 for $m + 1$, until death or the end of follow-up date. Within each person-year, we assume the
3 temporal order of the component variables assigned to each person-year as follows: fixed
4 covariates, employment status, conditional TiO_2 exposure, lung cancer death, and death from
5 another cause.

6 The final step was to calculate for each simulated joint pseudo-cohort the number of deaths in
7 each age group assuming 25 and 35-year exposure.

8 In order to select the most relevant explanatory variables for each of the four models explained
9 above, we based ourselves on *Akaike Information Criterion* (AIC). This criterion is based on a
10 compromise between the quality of the fit and the complexity of the model. This penalizes models
11 with a large number of variables to limit the effects of over-fitting. The model with the lowest
12 AIC value was therefore considered the best for predicting employment status, exposure status,
13 cumulative exposure, and lung cancer death and death from other causes in year i respectively.

14
15 The final risk model for lung cancer death for each subject in each year is as follows:

$$\begin{aligned} 17 \quad & \textit{logit}(\textit{cancer}) \\ 18 \quad & = \beta_1 \textit{age} + \beta_2 \textit{age}^2 + \beta_3 \textit{age}^2 \mathbb{1}_{\textit{age} > 65} + \beta_4 \textit{age}^2 \mathbb{1}_{\textit{age} > 70} \\ 19 \quad & + \beta_5 (\textit{year of hire} - 1960)^2 + \beta_6 \textit{cumulative exposure}_{\textit{lag} 10} \\ 20 \quad & + \beta_7 (\textit{cumulative exposure}_{\textit{lag} 10})^2 + \beta_8 \mathbb{1}_{\textit{active employment}} \\ 21 \quad & + \beta_9 \mathbb{1}_{\textit{active employment}_{\textit{lag} 1}} + \beta_{10} \textit{employment duration} \\ 22 \quad & + \textit{intercept} \end{aligned}$$

1 with the β coefficients of each variable calculated for the entire cohort of four countries.
2 Table S1 confirms that the selected model simulates well the natural course in the joint cohort.
3 The cumulative exposure variable Lag10 is the lagged cumulative exposure of 10 years. The
4 application of lag on the cumulative exposure variable consists of introducing a lag time between
5 exposure and cancer occurrence when modeling the relationship between the two. In practice, this
6 involves deducting from the cumulative exposure variable the annual exposures over the last 10
7 years preceding the occurrence of cancer. The 10-year value is used by convention for solid
8 cancers, although it is possible to derive other, more specific values for a given cancer by
9 modelling.

10
11

1 Table S1. Observed and simulated cumulative deaths due to lung cancer by attained age

2

Age (years)	Number of deaths	
	Number of observed death	(associated 95- % confidence interval) estimated by baseline model
50	9	10 (8-12)
60	43	44 (40-49)
70	99	113 (105-120)
80	138	139 (133-146)
90	139	140 (133-146)

3

4

1 Table S2. Characteristics of the cohort of male titanium dioxide workers

2

Characteristics	<i>N (%)</i>
Number of workers	7341 (100)
Number of workers exposed to TiO ₂	6019 (82.0)
Vital status at the end of follow-up	
Alive	5945 (81.0)
Deceased	1223 (16.7)
From lung cancer	139 (11.4)
From unknown cause	13 (1.1)
Lost to follow-up	173 (2.4)
Number of person-years	179040 (100)
Characteristics	<i>Mean (Standard deviation)</i>
Attained age at hire (years)	28.5 (9.2)
Attained age at start of follow-up (years)	30.2 (9.3)
Attained age at end of follow-up (years)	52.1 (10.9)
Employment duration (years)	13.6 (10.1)
Exposure to TiO ₂ (respiratory fraction)*	<i>Mean (Standard deviation) [Min-Max]</i>
Cumulative (mg/m ³ -years)	6.17 (10.47) [0.0003-142.8]
Annual average (mg/m ³)	0.47 (0.79) [0.00001-8.0]

3 *Exposure concentration among exposed workers

4

5

6