

1 **Variability in the association between long-term exposure to ambient air pollution and**
2 **mortality by exposure assessment method and covariate adjustment: a census-based**
3 **country-wide cohort study**

4 *Mariska Bauwelinck^a, Jie Chen^b, Kees de Hoogh^c, Klea Katsouyanni^{d,e}, Sophia Rodopoulou^d,*
5 *Evangelia Samoli^d, Zorana J. Andersen^f, Richard Atkinson^g, Lidia Casas^{h,i}, Patrick*
6 *Deboosere^a, Claire Demoury^j, Nicole Janssen^k, Jochem O. Klompmaker^{k,l}, Wouter Lefebvre^m,*
7 *Amar Jayant Mehta^f, Tim S. Nawrot^{h,n}, Bente Oftedal^o, Matteo Renzi^p, Massimo Stafoggia^{p,q},*
8 *Maciek Strak^{b,r}, Hadewijch Vandenheede^a, Charlotte Vanpoucke^s, An Van Nieuwenhuyse^{h,j,l},*
9 *Danielle Vienneau^c, Bert Brunekreef^b, Gerard Hoek^b*

10

11 ^aInterface Demography – Department of Sociology, Vrije Universiteit Brussel, Brussels,
12 Belgium

13 mariska.bauwelinck@vub.be, patrick.deboosere@vub.be, hadewijch.vandenheede@vub.be

14 ^bInstitute for Risk Assessment Sciences, Utrecht University, Utrecht, the Netherlands

15 j.chen1@uu.nl, maciek.strak@rivm.nl, B.Brunekreef@uu.nl, g.hoek@uu.nl

16 ^cSwiss Tropical and Public Health Institute, Basel, Switzerland

17 University of Basel, Basel, Switzerland

18 c.dehoogh@swisstph.ch, danielle.vienneau@swisstph.ch

19 ^dDepartment of Hygiene, Epidemiology and Medical Statistics, Medical School, National and
20 Kapodistrian University of Athens, Athens, Greece

21 kkatsouy@med.uoa.gr, srodopoyl@med.uoa.gr, esamoli@med.uoa.gr

22 ^eEnvironmental Research Group Imperial College, London, London UK

23 kkatsouy@med.uoa.gr

24 ^fDepartment of Public Health, University of Copenhagen, Copenhagen, Denmark

25 vlq961@sund.ku.dk, amar.mehta@sund.ku.dk

26 ^gPopulation Health Research, Institute St George's, University of London, London, UK

27 atkinson@sgul.ac.uk

28 ^hCentre for Environment and Health, Department of Public Health and Primary Care, KU

29 Leuven, Leuven, Belgium
30 Lidia.CasasRuiz@uantwerpen.be, tim.nawrot@uhasselt.be, An.vanNieuwenhuyse@Ins.etat.lu

31 ⁱMedical Sociology and Health Policy, Department of Epidemiology and Social Medicine,
32 University of Antwerp, Wilrijk, Belgium
33 Lidia.CasasRuiz@uantwerpen.be

34 ^jRisk and Health Impact Assessment Unit, Sciensano, Brussels, Belgium
35 Claire.Demoury@sciensano.be, An.vanNieuwenhuyse@Ins.etat.lu

36 ^kNational Institute for Public Health and the Environment, Bilthoven, the Netherlands
37 nicole.janssen@rivm.nl, jklompmaker@hsph.harvard.edu

38 ^lNational Institute for Public Health and the Environment, Bilthoven, the Netherlands
39 Harvard T.H. Chan School of Public Health, Boston, MA, USA
40 jklompmaker@hsph.harvard.edu

41 ^mVlaamse Instelling voor Technologisch Onderzoek (VITO), Mol, Belgium
42 wouter.lefebvre@vito.be

43 ⁿCentre for Environmental Sciences, University of Hasselt, Diepenbeek, Belgium
44 tim.nawrot@uhasselt.be

45 ^oDepartment of Environmental Health, Norwegian Institute of Public Health, Oslo, Norway
46 BenteMargaret.Oftedal@fhi.no

47 ^pDepartment of Epidemiology, Lazio Region Health Service / ASL Roma 1, Rome, Italy
48 m.renzi@deplazio.it, m.stafoggia@deplazio.it

49 ^qInstitute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden
50 m.stafoggia@deplazio.it

51 ^rNational Institute for Public Health and the Environment, Bilthoven, the Netherlands
52 maciek.strak@rivm.nl

53 ^sBelgian Interregional Environment Agency (IRCEL-CELINE), Brussel, Belgium
54 vanpoucke@irceline.be

55 ¹*Present address: Department of Health Protection, Laboratoire National de Santé (LNS),*
56 *Dudelange, Luxembourg*
57 An.vanNieuwenhuysen@lms.etat.lu

58

59 **Corresponding Author**

60 Mariska Bauwelinck, Interface Demography, Vrije Universiteit Brussel, Pleinlaan 5, 1050
61 Brussels, Belgium, Tel: +3226148132, Fax: +3226148135, e-mail:
62 mariska.bauwelinck@vub.be

63

Abstract

64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87

Background: Ambient air pollution exposure has been associated with higher mortality risk in numerous studies. We assessed potential variability in the magnitude of this association for non-accidental, cardiovascular disease, respiratory disease, and lung cancer mortality in a country-wide administrative cohort by exposure assessment method and by adjustment for geographic subdivisions.

Methods: We used the Belgian 2001 census linked to population and mortality register including nearly 5.5 million adults aged ≥ 30 (mean follow-up: 9.97 years). Annual mean concentrations for fine particulate matter (PM_{2.5}), nitrogen dioxide (NO₂), black carbon (BC) and ozone (O₃) were assessed at baseline residential address using two exposure methods; Europe-wide hybrid land use regression (LUR) models [100x100m], and Belgium-wide interpolation-dispersion (RIO-IFDM) models [25x25m]. We used Cox proportional hazards models with age as the underlying time scale and adjusted for various individual and area-level covariates. We further adjusted main models for two different area-levels following the European Nomenclature of Territorial Units for Statistics (NUTS); NUTS-1 (n=3), or NUTS-3 (n=43).

Results: We found no consistent differences between both exposure methods. We observed most robust associations with lung cancer mortality. Hazard Ratios (HRs) per 10 $\mu\text{g}/\text{m}^3$ increase for NO₂ were 1.060 (95%CI 1.042-1.078) [hybrid LUR] and 1.040 (95%CI 1.022-1.058) [RIO-IFDM]. Associations with non-accidental, respiratory disease and cardiovascular disease mortality were generally null in main models but were enhanced after further adjustment for NUTS-1 or NUTS-3. HRs for non-accidental mortality per 5 $\mu\text{g}/\text{m}^3$ increase for PM_{2.5} for the main model using hybrid LUR exposure were 1.023 (95%CI 1.011-1.035). After including random effects HRs were 1.044 (95%CI 1.033-1.057) [NUTS-1] and 1.076 (95%CI 1.060-

88 1.092) [NUTS-3].

89 **Conclusion:** Long-term air pollution exposure was associated with higher lung cancer
90 mortality risk but not consistently with the other studied causes. Magnitude of associations
91 varied by adjustment for geographic subdivisions, area-level socio-economic covariates and
92 less by exposure assessment method.

93

94 **Keywords:**
95 **population-based**
96 **environmental hazard**
97 **exposure assessment**
98 **survival analysis**
99 **cause-specific mortality**
100 **health effects**

101

102 **Highlights:**

103 **Large prospective country-wide cohort study including nearly 5.5 million adults**
104 **Non-accidental and cause-specific mortality over long-term ten years follow-up**
105 **Several ambient air pollutants evaluated using two exposure assessment models**
106 **Most robust associations observed between both NO₂ or BC and lung cancer mortality**
107 **Associations varied mildly between hybrid LUR and interpolation-dispersion model**
108 **Magnitude associations differed by differential adjustment for area-level indicators**

109

110

111 **1 Introduction**

112 Over the past few years, a relatively large number of studies on the association between long-
113 term exposure to ambient air pollution and mortality has been published (Huangfu and
114 Atkinson, 2020; Chen and Hoek, 2020; Huang et al., 2021). The majority of studies reported
115 increased mortality risks, although large variation has been observed in magnitude of the effect
116 estimates both between and within countries (Huangfu and Atkinson, 2020; Chen and Hoek,
117 2020; Huang et al., 2021). Part of this heterogeneity in air pollution epidemiological studies
118 might be explained by methodological differences in exposure assessment method, study design
119 or statistical data analysis approach, or by study-specific contextual differences. So far there is
120 little evidence on how air pollution exposure assessment method affects mortality risk estimates
121 (Yap et al., 2012; Jerrett et al., 2016; Klompmaker et al. 2020; Samoli et al., 2020; Butland et
122 al., 2020; Gariazzo et al., 2021). Multicenter studies provide a great opportunity to investigate
123 some of this heterogeneity. This study forms part of the Effects of Low-level Air Pollution: A
124 Study in Europe (ELAPSE) project (www.elapseproject.eu) (Klompmaker et al., 2020;
125 Hvidtfeldt et al., 2020), where Belgium is one of the seven participating European countries
126 contributing to the project with large administrative cohort data. The project's central approach
127 was to harmonize to the greatest extent possible exposure assessment, outcome and confounder
128 definitions as well as statistical methods between different administrative cohorts. Study-
129 specific contextual heterogeneity is likely to remain notwithstanding large harmonization
130 efforts and may potentially affect health effect estimates in relation to long-term exposure to
131 air pollution. Study-specific between-area variability in mortality patterns has been widely
132 observed in several country-wide studies, including in Belgium (Deboosere and Gadeyne, 2002;
133 Van Hemelrijck et al., 2016). Air pollution health effect estimates may be affected if broad
134 scale air pollution patterns are correlated to regional mortality patterns. In recent North

135 American cohort studies, investigators have adjusted for geographic subdivisions of the country
136 to account for potential variability in spatial patterns (Crouse et al. 2012, 2015; Di et al. 2017).
137 The current study presents results for the Belgian administrative cohort on the association
138 between long-term exposure to several ambient air pollutants (fine particulate matter (PM_{2.5}),
139 nitrogen dioxide (NO₂), black carbon (BC) and ozone (O₃)) and non-accidental,
140 cardiovascular disease, respiratory disease, and lung cancer mortality during a ten-year follow-
141 up period for about 5.5 million Belgian adults. The aim of this study was to explore and assess
142 potential variability in mortality effect estimates by different air pollution exposure assessment
143 methods and by additional adjustment for geographic subdivisions of the country.

144

145 **2 Methods**

146 **2.1 Data design and study population**

147 Administrative cohort data was based on the Belgian 2001 census which was linked to
148 population, emigration and mortality follow-up data for the study period October 1, 2001-
149 December 31, 2011 (10.25 years). Data were made available by the Belgian statistical office
150 (Statbel) and contained individual information for the entire population officially residing in
151 Belgium at the time of the census. Individuals were geolocated based on the XY-coordinate of
152 their residential address at baseline, near-complete with 98.7% of individuals included. All
153 adults aged 30 and older with complete covariate information were included in the present
154 study. We excluded about 15% of individuals with missing data on main covariates.

155 Individual sociodemographic covariates were collected through a census questionnaire at
156 baseline, and included: age, sex, marital status (single, cohabiting/married, separated/divorced
157 and widowed), country of origin (local vs foreign), education level (no/primary, secondary and
158 tertiary), and occupational status (employed/self-employed, unemployed, homemaker and
159 retired). Available area-level socio-economic position (SEP) covariates consisted of mean

160 income (i.e. mean household net taxable income), unemployment (i.e. percentage of working
161 age population unemployed), low education (i.e. percentage of population with no/primary
162 education), and ethnicity (i.e. percentage of non-Western migrants). All area-level SEP
163 indicators were retrieved from the Belgian 2011 census, except for ethnicity which was only
164 obtainable for the year 2001. Area-level SEP variables were available at two different area-
165 levels: 1) neighbourhood (n=6,344), i.e. geographical units having a size in between those of
166 census tracts (n=19,781) and local administrative units (LAU) (n=589); and 2) NUTS-3 (n=43),
167 i.e. as defined by the European Nomenclature of Territorial Units for Statistics (NUTS)
168 (Eurostat, 2018). Both aforementioned area-level SEP definitions and selected spatial levels
169 were based on the statistical protocol of ELAPSE (Klompaker et al., 2020).

170

171 **2.2 Air pollution exposure assessment**

172 Air pollution exposure assessment was done using two approaches: Europe-wide hybrid land use
173 regression (LUR) and Belgian interpolation-dispersion (RIO-IFDM) exposure models. Annual
174 mean concentrations for different ambient air pollutants (PM_{2.5}, NO₂, BC and O₃) for the year
175 2010 were assigned to the residential geocode at baseline (01/10/2001). The measurements for
176 O₃ were obtained by averaging warm season months from April through September. A brief
177 description of the methodologies of both models is given below and an overview of the
178 differences can be found in Supplementary table S1.

179

180 **2.2.1 European hybrid LUR model**

181 In the framework of ELAPSE, Europe-wide air pollution exposure assessment was developed
182 and validated following a harmonised protocol, described in detail by de Hoogh et al. (2018). In
183 brief, hybrid LUR models were developed by combining air pollution monitoring data with
184 predictor variables obtained from satellite derived air pollution data, chemical transport model

185 data, and land cover and road traffic data. Monitoring data for PM_{2.5}, NO₂ and O₃ warm season
186 were derived from Airbase version 8 routine data (EEA, 2020; de Hoogh et al., 2016). As Airbase
187 data were not available for BC, European Study of Cohorts for Air Pollution Effects (ESCAPE)
188 monitoring data were used instead (Eeftens et al., 2012a; 2012b). Models were developed at a
189 spatial resolution of 100 x 100 m for the year 2010 (annual mean). Estimates for PM_{2.5}, NO₂
190 and O₃ were expressed in micrograms per cubic meter ($\mu\text{g}/\text{m}^3$) and for BC in 10^{-5}m^{-1} (i.e. similar
191 magnitude compared to BC in $\mu\text{g}/\text{m}^3$).

192

193 **2.2.2 Belgian RIO-IFDM model**

194 Air quality model exposure predictions for the same pollutants and year were provided by the
195 Belgian Interregional Environment Agency (IRCEL-CELINE). The estimates were obtained
196 through the coupling of a spatial interpolation model (RIO) and a dispersion model (IFDM).
197 The interpolation model uses air quality measurements from fixed measuring stations and
198 CORINE Land Cover data (EEA, 2019; Hooyberghs et al., 2006). These background results
199 were combined with a dispersion receptor model using emissions from industrial point and
200 traffic line sources and meteorological data (Lefebvre and Vranckx, 2013). The results are
201 modelled on high-resolution grids of 25 x 25 m. Further details regarding the applied model
202 chain can be consulted in the following technical report by Lefebvre and Vranckx (2013). All
203 annual mean concentrations were expressed in micrograms per cubic meter ($\mu\text{g}/\text{m}^3$).

204

205 **2.3 Mortality outcomes**

206 The studied mortality outcomes were identified through the WHO International Classification
207 of Diseases, Tenth Revision codes (ICD-10) (W.H.O., 2004), based on the selection of the
208 underlying cause of death on the death certificates. We considered non-accidental (ICD-10:
209 A00-R99), cardiovascular disease (ICD-10: I10-I70), respiratory disease (ICD-10: J00-J99),

210 and lung cancer mortality (ICD-10: C34.0-C34.9).

211

212 **2.4 Statistical analyses**

213 We assessed the association between the different air pollutants and mortality outcomes using
214 Cox proportional hazard models with age as the underlying time scale. Individuals were right
215 censored when information about their survival time was incomplete, i.e. death to another cause
216 not under study for cause-specific outcomes, loss to follow-up due to emigration or end of
217 follow-up (31/12/2011).

218 Three models with increasing degree of adjustment were defined a priori within the ELAPSE
219 project (Klompaker et al., 2020; Hvidtfeldt et al., 2020): model 1 (M1) stratified by sex and
220 accounted for within-area correlations of the individuals by including a cluster term for
221 neighbourhood (Therneau, 2015); model 2 (M2) adding to M1 with additional adjustment for
222 individual sociodemographic covariates (marital status, country of origin, education level and
223 occupational status), and model 3 (M3) adding to M2 with additional control for area-level SEP
224 indicators (mean income, unemployment, low education, and ethnicity). In the analysis, area-
225 level SEP was operationalized as the NUTS-3 area-level SEP variable and the deviation
226 between NUTS-3 and neighbourhood area-level SEP variable. In ELAPSE we a priori decided
227 to adjust for multiple dimensions of SEP at both a neighbourhood and regional scale to adjust
228 for potential confounding by socio-economic indicators.

229 We evaluated the shape of the concentration-response curves for the relationship between the
230 different air pollutants and mortality outcomes. We specified natural spline plots for three
231 degrees of freedom (df) (Eisen et al., 2004) and compared the goodness of fit of these models
232 with the models specified with a linear term (M3) using the Bayesian Information Criterion
233 (BIC). No clear deviation from linearity was found based on the model fit nor the splines (i.e.

234 large uncertainty observed about the shape at low and high end of the distribution as indicated
235 by the 95% CIs), thus exposure hazard ratios (HR) were reported as a continuous linear term
236 (Supplementary Figure S1). For linear models, results are presented as HRs with 95% CIs using
237 pollutant-specific increments based on the ESCAPE project: 5 $\mu\text{g}/\text{m}^3$ for $\text{PM}_{2.5}$, 10 $\mu\text{g}/\text{m}^3$ for
238 NO_2 , $0.5 \cdot 10^{-5}\text{m}^{-1}$ (hybrid LUR) or 0.5 $\mu\text{g}/\text{m}^3$ (RIO-IFDM) for BC, and 10 $\mu\text{g}/\text{m}^3$ for O_3 .

239 Based on the single pollutant main model (M3), we specified two-pollutant models where
240 pollutants within the same exposure model (i.e. hybrid LUR and RIO-IFDM) were
241 simultaneously entered in the model to assess potential co-pollutant confounding.

242 In additional analyses, we specified two alternative mixed-effect Cox models with random
243 intercept. Both included additional levels of spatial correlation to account for potential
244 differences in mortality rate between geographical areas not accounted for in the main model.

245 The first model adjusted for both neighbourhood and large geographical NUTS-1 area-level
246 ($n=3$), whereas the second model adjusted for both neighbourhood and NUTS-3 area-level
247 ($n=43$). To explore potential effect modification, we included multiplicative interaction terms
248 into our main model between each of the pollutants and age (<65 years or ≥ 65 years), and
249 education level (no/primary education, secondary education or tertiary education). We
250 evaluated the goodness of fit of models with and without interaction term using the Wald test.

251 As sensitivity analyses, we repeated M1 with the full population sample (i.e. complete cases
252 analysis using only M1 covariates) and compared these with the reduced sample of the main
253 model (i.e. complete cases after including M3 covariates). We further evaluated the consistency
254 of our effect estimates to area-level SEP adjustment in our main model (M3) by specifying
255 models where each of the four available area-level SEP indicator was adjusted for separately
256 instead of combined. Additionally, we indirectly adjusted main model HRs to account for
257 important missing health-related behavioral indicators in the census in relation to mortality risk.

258 We used the method proposed by [Shin et al. \(2014\)](#) to apply indirect adjustment for both

259 smoking status (current, former or never) and body mass index (BMI) (underweight <18.5,
260 normal 18.5-24.9, overweight 25-29.9 or obese >30). In brief, the indirect adjustment method
261 extracts ancillary information on aforementioned health-related behavioral indicators from a
262 dataset representative of the study population. We obtained the Belgian 2001 Health Interview
263 Survey (HIS) (<http://www.healthsurvey.be>) matching with the baseline year of the
264 administrative cohort. Besides information on smoking status and BMI, the HIS included the
265 same individual and area-level covariates as in our main model, with the exception of marital
266 status which was not available. We assigned identical exposure models to the HIS participants,
267 following the same procedure as previously described in section 2.2. We then ran multivariate
268 linear regression models with the harmonized HIS data to retrieve the estimates based on the
269 association between the air pollutants and the available health-related behavioral indicators. In
270 addition to aforementioned estimates, the indirect adjustment method also uses estimates based
271 on the association between the health-related behavioral indicators and the different mortality
272 outcomes under study, which have been retrieved from literature. More methodological details
273 are described in the work of [Shin et al. \(2014\)](#).

274 Statistical analyses and exposure data linkages were performed in R version 3.4.0 (R Core Team
275 2019) and RStudio (RStudio Team, 2019) using the following packages: survival (Therneau,
276 2015), coxme (Therneau, 2018), ggplot2 (Wickham, 2009), data.table (Dowle and Srinivasan,
277 2017), gdalUtils (Greenberg and Mattiuzzi, 2015), raster (Hijmans, 2016), rgdal (Bivand et al.,
278 2017), and base and dependency packages.

279

280 **3 Results**

281 **3.1 Study population and air pollution exposure**

282 The included study population consisted of 5,474,470 adults, with a total of 54,574,471 person-
283 years and mean follow-up period of 9.97 years (Table 1). The number of men and women was

284 nearly equal with a mean age at baseline of 52.6 years. The majority of subjects were born in
285 Belgium (96.6%), were cohabiting/married (68.3%), had obtained secondary education level or
286 higher (76.3%), and were employed (53.3%) at the time of the census. We observed 707,138
287 individuals who died from non-accidental causes of which 33.2% from cardiovascular disease,
288 11.6% from respiratory disease, and 7.4% from lung cancer mortality.

289 The exposure distribution and pairwise correlations for the different pollutants are summarised
290 in Table 2, Supplementary Table S1 and Supplementary Figures S2-S3. For all four pollutants,
291 median values were higher in hybrid LUR compared to RIO-IFDM exposure models, whereas
292 the interquartile range (IQR) was moderately lower in hybrid LUR models (Table 2). Lower
293 variability of the hybrid LUR model is particularly reflected in the lowest and highest
294 percentiles of the distributions, whereas the range of observed concentrations was wider for all
295 different pollutants in the RIO-IFDM model (Supplementary Figure S2). The broad spatial
296 patterns of exposure distributions agreed quite well between both exposure models for all
297 pollutants (Supplementary Figure S3).

298 Pearson correlations between hybrid LUR and RIO-IFDM models were 0.64, 0.86, 0.82 and
299 0.76 for PM_{2.5}, NO₂, BC and O₃, respectively (Supplementary Table S1). Generally, correlations
300 between pollutants were stronger in the RIO-IFDM compared to hybrid LUR exposure model
301 (e.g. 0.83 vs 0.62 between PM_{2.5} and NO₂, respectively). Correlations between different
302 pollutants were moderate to high, especially between NO₂ and BC. Also, expectedly, O₃ was
303 negatively correlated with all other pollutants.

304 **3.2 Association between air pollution and mortality**

305 **3.2.1 Main analyses**

306 Hazard ratios (HRs) from single-pollutant models with increasing confounder adjustment for
307 different mortality outcomes under study are presented in Figure 1 and Supplementary Table

308 S2. HRs were sensitive to incremental adjustment for potential confounders. Overall, hazard
309 ratios increased after individual level covariate adjustment (M2) for PM_{2.5}, NO₂ and BC. After
310 area-level SEP covariate adjustment (M3), HRs mostly attenuated, except for associations with
311 PM_{2.5} where HRs generally increased. In single pollutant main models (M3), we found small
312 HRs both above and below unity with differing patterns depending on the studied outcome.
313 Main model HRs ranged between 0.975 and 1.060 (Figure 1 and Supplementary Table S2). For
314 non-accidental mortality we only found a significant association for PM_{2.5} with the hybrid LUR
315 model (HR: 1.023, 95%CI 1.011-1.035). Observed HRs for cardiovascular mortality were
316 mostly below unity, except for O₃ where HRs were above unity. For both respiratory and lung
317 cancer mortality, HRs were mainly larger than unity, with strongest HRs observed with NO₂
318 and BC. HRs between hybrid LUR versus RIO-IFDM exposure models generally agreed for
319 the different outcomes, although stronger estimates were mainly found in hybrid LUR models
320 (Supplementary Table S3 with M3 HRs per IQR increase). The difference in HRs between the
321 hybrid LUR and RIO-IFDM model exposures was larger in the fully adjusted model (M3) than
322 in the age and sex only model (M1).

323

324 Our main results were relatively robust after further adjustment in two-pollutant models (Table
325 3). However, interpretation of these estimates must be with caution due to potential
326 multicollinearity, especially between NO₂ and BC. The association between non-accidental
327 mortality and PM_{2.5} remained and became slightly stronger after adjustment for NO₂, BC or
328 O₃. Associations with NO₂ became stronger after adjustment for O₃. Associations with O₃
329 became larger than unity and significant after adjustment for the other pollutants with the
330 hybrid LUR exposure model. For cardiovascular mortality, negative associations with O₃
331 remained significant only after adjustment for PM_{2.5} in hybrid LUR and BC in RIO-IFDM
332 exposure models. The significant inverse associations in single pollutant models approached

333 unity after adjustment for O₃. Associations with lung cancer mortality remained in both hybrid
334 LUR and RIO-IFDM exposure models for NO₂ and BC after adjustment for other pollutants,
335 except for BC after NO₂ adjustment. Associations in two-pollutant models were most notable
336 in both respiratory and lung cancer mortality where HRs generally were stronger after
337 adjustment for O₃, in addition to higher estimates for O₃.

338

339 **3.2.2 Additional analyses**

340 In additional analysis, we further accounted for between-area variability by including a random
341 intercept in our main models for neighbourhood and NUTS-1 (n=3) or neighbourhood and
342 NUTS-3 area-level (n=43) (Figure 1 and Supplementary Table S4). Specification of random
343 effects with NUTS-1 area-level only mildly affected HRs, with the exception of non-accidental
344 mortality where associations between PM_{2.5}, NO₂ and BC became larger than unity and
345 statistically significant, albeit with small HRs. Estimates were influenced more when allowing
346 for random effects with the spatially more detailed level of NUTS-3, and generally resulted in
347 substantially larger HRs, mainly for associations with PM_{2.5}. Overall, most HRs that were above
348 unity in our main model (M3) became stronger for PM_{2.5}, NO₂ and BC. HRs in models with
349 aforementioned pollutants that were lower than unity lost statistical significance or became larger
350 than unity with increasing degree of area-level adjustment. HRs for associations with O₃ became
351 inversely statistically significant with increasing area-control for non-accidental, respiratory and
352 lung cancer mortality. Associations with O₃ and cardiovascular mortality did not retain statistical
353 significance. Also, differences in effect estimates between the two exposure assessment
354 methods became smaller and more stable when introducing random effects with NUTS-1 or
355 more pronouncedly including the spatially more refined NUTS-3 area-level.

356 Effect modification analyses by age indicated stronger associations for all mortality outcomes
357 under study with PM_{2.5}, NO₂ and BC in younger age (≤ 65 years), and with O₃ in older age (≥ 65

358 years) (Supplementary Table S5). Observed effect modification patterns by education level
359 were overall suggestive of stronger associations for PM_{2.5}, NO₂ and BC among individuals with
360 tertiary education (Supplementary Table S5).

361

362 **3.2.3 Sensitivity analyses**

363 Effect estimates for M1 including the full population sample (i.e. individuals without any
364 missing value for air pollution exposure, age and sex) were almost identical for non-accidental
365 and cardiovascular mortality and slightly stronger for respiratory and lung cancer mortality,
366 although very similar compared to the reduced sample (i.e. with no missing additional
367 covariates) used in the main models (Supplementary Table S6).

368 HRs were sensitive to the inclusion of different area-level SEP covariates (Supplementary Table
369 S7). When adjusting separately for each area-level SEP variable, HRs differed in both directions
370 from M2 and the main model (M3; i.e. all available area-level SEP indicators combined). For
371 example, for non-accidental and respiratory mortality in model SEP3, effects were downward for
372 PM_{2.5} and upward for NO₂ compared to the main model. The observed sensitivity was less for lung
373 cancer mortality where HRs were larger. No substantial differences were observed between the
374 different exposure models.

375 Study population characteristics between cohort and survey data were fairly similar (Supplementary
376 Table S8), suggesting the use of the survey for the retrieval of ancillary information to be adequate.
377 Indirect adjusted HRs for smoking status and BMI were generally higher in all mortality outcomes
378 and for both exposure models. Strongest effect estimates were consistently observed in mortality
379 associations with PM_{2.5} (Supplementary Table S9).

380

381 **4 Discussion**

382 We observed associations between long-term exposure to ambient air pollution and mortality
383 risk for natural and cause-specific mortality outcomes. Effect estimates were sensitive to

384 exposure assessment method, additional adjustment for geographical subdivisions (NUTS-1 or
385 NUTS-3) of the country and differential adjustment for area-level socio-economic covariates.
386 Mortality risk in relation to ambient air pollution suggested to be highest among individuals
387 younger than 65 years at baseline or with tertiary education. Overall, we observed most robust
388 associations with lung cancer and both NO₂ or BC for both exposure methods, independently
389 of alternative model specifications. Observed consistency of aforementioned results among
390 exposure methods is an important finding, as each method may incorporate different degrees of
391 measurement error. These potentially introduce bias to health effect estimates of which
392 magnitude and direction is hard to quantify.

393 To our knowledge, only four other studies systematically compared potential heterogeneity in
394 effect estimates using different exposure assessment methods when evaluating the association
395 between long-term exposure to ambient air pollution and various mortality outcomes using
396 cohort data (Yap et al. 2012; Jerrett et al., 2016; Klompmaker et al. 2020; Gariazzo et al., 2021).

397 All four aforementioned studies also detected variation in the effect estimates in terms of
398 magnitude, direction or statistical significance depending on the applied exposure assessment
399 method. The study of Klompmaker et al. (2020), using Dutch administrative cohort data, was
400 also part of the ELAPSE project. Similar to our study, they also observed moderate correlations
401 for PM_{2.5} and relatively strong correlations for NO₂ and BC between different exposure
402 methods (Klompmaker et al. 2020). Comparably, differences in HRs for both NO₂ and BC
403 between exposure models were smaller in minimally adjusted models (M1; i.e. including age
404 and sex) versus fully adjusted models (M3), reflecting differential correlation patterns between
405 pollutants and area-level SEP. Further, comparison of effect estimates based on the same hybrid
406 LUR exposure model and non-accidental mortality were almost identical for associations
407 between non-accidental mortality and PM_{2.5} [HR 1.023 (95%CI 1.011-1.035) for the current
408 (Belgian) and HR 1.030 (95%CI 1.019-1.041) for the Dutch administrative cohort

409 (Klompaker et al. 2020)]. Overall observed patterns with hybrid LUR exposure methods were
410 similar in both the Belgian and Dutch administrative cohort, where strongest associations were
411 observed for lung cancer and weakest for cardiovascular mortality (Klompaker et al. 2020).
412 When study-specific between-area variability was additionally accounted for, associations in
413 our study between PM_{2.5}, NO₂ and BC and mortality became stronger; hence, indicating that
414 potential residual confounding does not necessarily lead to effect estimates biased upwards.
415 This finding is consistent with a review reporting that more complete adjustment for area-level
416 indicators tended to increase air pollution effect estimates rather than decrease (Vodanos et al.,
417 2018). In Canadian cohort studies (Crouse et al., 2012 and 2015), HRs also increased after
418 adjustment for large geographical area of the country. Additional adjustment for geographical
419 subdivisions (neighbourhood in addition to NUTS-1 or NUTS-3), reflected broad-scale spatial
420 variation in health due to factors other than air pollution or included socio-economic covariates
421 at individual and area-level. Previous research on spatial variability in mortality patterns in
422 Belgium identified a clear north-south gradient across the country, where mortality rates
423 generally are highest in the south and in former industrial areas (Deboosere and Gadeyne, 2002;
424 Van Hemelrijck et al., 2016). Other possible explanations for this geographic variation in health
425 status have been proposed, such as differences in diagnostic and therapeutic practices, cultural
426 and health-related behaviours and historical context (Deboosere and Gadeyne, 2002; Van
427 Hemelrijck et al., 2016). Although we aimed to maximise the number of available relevant
428 covariates in our study, no data on these specific factors was available for linkage to the Belgian
429 administrative cohort. Therefore, we recognise that some important unobserved residual
430 confounding may remain. With regard to country-wide spatial trends of air pollution, the
431 aforementioned north-south gradient is inverse: observed pollutant levels are highest in the
432 north and decrease towards the south of the country (Supplementary Figure S3). In
433 consequence, additional adjustment for between-area variability as random effects in our main

434 model might have accentuated the generally small exposure contrasts between different area-
435 levels (neighbourhood in addition to NUTS-1 or NUTS-3).

436 Consistent with the majority of prior research evaluating effect modification by age in the
437 association of long-term exposure to air pollution (Huangfu and Atkinson, 2020; Chen and
438 Hoek, 2020), our study confirmed earlier findings showing higher mortality risk in younger
439 individuals (<65 years) with PM_{2.5}, NO₂ and BC. Current evidence on potential effect
440 modification by education level with aforementioned pollutants is still limited and inconclusive.

441 Two other participating administrative cohorts in the ELAPSE project evaluated effect
442 modification by education level. In accordance with our study findings, the Swiss cohort also
443 detected strongest associations among higher educated compared to lower educated with PM_{2.5},
444 NO₂ and BC. Contrarily, the observed pattern was opposite in the Norwegian cohort ("results
445 not shown" or REF ELAPSE report if available). Exposure distributions of studied pollutants
446 were nearly identical between different population subgroups by age or education level. Health
447 and mortality risks are known to be generally higher among individuals with lower versus
448 higher education levels, which is often referred to as the social gradient in health (Wilkinson
449 and Marmot, 2003). This is also true for our study, where we found relative mortality risks to
450 increase two- to three- fold between each category of education level. The aforementioned
451 social gradient among population subgroups has been attributed to several underlying health
452 determinants, such as differences in health-related behaviors (e.g. tobacco and alcohol use,
453 dietary habits or physical activity) or differential access to important resources (e.g. access to
454 health care or basic housing conditions). While in our study we only observed higher mortality
455 risks among younger or higher educated individuals, presumed mortality risks among older or
456 lower educated individuals in relation to long-term exposure to air pollution may also be
457 detected if other, potentially more influential health determinants could be mitigated. We
458 speculate that the absence of such determinants in our data might partially explain observed

459 null-trends for cardiovascular mortality in our main model.

460 When disentangling sensitivity of various area-level SEP indicators into separate models, we
461 observed heterogeneity of patterns in effect estimates for different pollutants and mortality
462 outcomes. This finding points to the multiplicity of the construct of (area-level) SEP, as well as
463 its complex interplay with different air pollutants. Consequently, comprehensive explanation is
464 not straightforward and deserves to be addressed further in future studies focussing on health
465 and environmental inequalities.

466 Previous studies on the health effects of air pollution emphasised the importance of adjustment
467 for SEP indicators at both individual and area-level since associations with health outcomes
468 seemed to be independent (Roux, 2007; Temam et al., 2017; Vodonos et al., 2018).
469 Additionally, it has been argued that adjustment for area-level SEP complementary to
470 individual SEP might be of particular interest in studies where individuals' geographic location
471 is important (Galobardes et al., 2007). Also, the inclusion of various SEP indicators to represent
472 its different dimensions was suggested to be important (Galobardes et al., 2007; Pinault et al.,
473 2016). Given the complexity of SEP and in order to reduce confounding as much as possible, our
474 main model (M3), as has been defined a priori within the ELAPSE project, adjusted for as many
475 individual and area-level SEP indicators as available. Although concerns for potential over-
476 adjustment might be valid, a recent meta-analytic review on associations between PM_{2.5} and
477 several mortality outcomes observed that additional adjustment for area-level SEP unlikely
478 results in upward bias (Vodonos et al., 2018). These findings are in line with our study, where
479 effect estimates for PM_{2.5} increased after area-level SEP adjustment with non-accidental
480 (hybrid LUR), respiratory disease (hybrid LUR and RIO-IFDM) and lung cancer mortality
481 (RIO-IFDM). However, we did not observe a similar pattern for the other pollutants under study.

482 Our study includes a number of limitations. First, and potentially most important, our study lacked
483 individual information on health-related behaviors, such as tobacco and alcohol use, dietary habits

484 or physical activity, as these have been identified as important determinants of mortality risk.
485 However, we tried to address this limitation by indirectly adjusting our main models with
486 information on smoking status and BMI using a survey representative of the study population. Such
487 adjustment resulted mainly in stronger mortality associations with PM_{2.5} for studied outcomes.
488 Indirect adjustment could not further explain observed weaker findings for cardiovascular mortality,
489 nor could it explain apparent stronger findings for lung cancer mortality. Therefore, speculate that
490 exposure to ambient air pollution may be a long-underestimated risk factor in relation to lung
491 cancer. Another limitation of our study is that only time-fixed exposure for the year 2010 could be
492 obtained for both exposure models. Although a decreasing trend in air pollution levels has been
493 observed across Europe over the last years, we assumed its spatial distribution remained relatively
494 stable over the follow-up period. Therefore, we may have underestimated mortality risk related to
495 ambient air pollution, although this could not be evaluated. Additionally, individual and area-level
496 covariates for different time points over the follow-up period were not available, which is a common
497 limitation in most administrative cohorts. Furthermore, updates on residential history were not
498 obtainable either.

499

500 **5 Conclusion**

501 Long-term term exposure to ambient air pollution was associated with higher mortality risk
502 among nearly 5.5 million Belgian adults. We observed variability in the strength of our effect
503 estimates by additional adjustment for geographic subdivisions of the country, area-level SEP
504 covariates and to a limited extent exposure assessment method. Most robust and consistent
505 associations were found between both NO₂ or BC and lung cancer mortality. Future studies
506 should apply caution and carefully evaluate analytic strategies as exposure assessment method,
507 different model specifications and covariate availability might influence both magnitude and
508 direction of health effect estimates related to long-term air pollution exposure.

509

510 **6 Acknowledgements**

511 The authors would like to thank IRCEL-CELINE (Belgian Interregional Environment Agency)
512 for providing the air pollution estimates and Statbel (Directorate-general Statistics - Statistics
513 Belgium) for geocoding the census data and facilitating several waves of data linkages during
514 the exposure assessment process. Furthermore, we acknowledge the help of Johan Surkyn from
515 Interface Demography VUB with data management as well as the support of Katrien De Troeyer
516 with the preparation of the data demand. We are grateful for the input of Benoit Nemery from
517 Centre for Environment and Health, Department of Public Health and Primary Care, KU
518 Leuven for his comments/suggestions on early draft versions of the article. Computational
519 resources and services were provided by the Shared ICT Services Centre funded by the Vrije
520 Universiteit Brussel, the Flemish Supercomputer Center (VSC) and the Research Foundation -
521 Flanders (FWO). The icons used for the graphical abstract were made available by the Noun
522 Project (<https://thenounproject.com>) under a Free Creative Commons (CC) license
523 [attributions: belgium by Matthieu Rodrigues, Fien Robbe and Paul Aimé; pollution by
524 callorine and jokokerto; air pollution by Eucalyp; factory by monkik; analysis by Rflor; conflict
525 by Gregor Cresnar; mortality by Nithinan Tatah; health by Lakshisha].

526

527 **7 Funding**

528 Research described in this article was conducted under contract to the Health Effects Institute
529 (HEI), an organization jointly funded by the United States Environmental Protection Agency
530 (EPA) (Assistance Award No. R-82811201) and certain motor vehicle and engine
531 manufacturers. The contents of this article do not necessarily reflect the views of HEI, or its
532 sponsors, nor do they necessarily reflect the views and policies of the EPA or motor vehicle and
533 engine manufacturers. Mariska Bauwelinck is funded by an individual PhD grant supported by the

534 Research Foundation - Flanders (FWO) [grant number 11A9718N].

535

536 **8 Data statement**

537 The research data is confidential.

538

539 **9 Competing Financial Interests**

540 The authors declare they have no actual or potential competing financial interests.

541

542

543 **References**

- 544 Bivand, R., Keitt, T., and Rowlingson, B. (2017). *rgdal: Bindings for the Geospatial Data*
545 *Abstraction Library*. R package version 1.2-8.
- 546 Breslow, N. E. (1975). Analysis of survival data under the proportional hazards model.
547 *International Statistical Review/Revue Internationale de Statistique*, pages 45–57.
- 548 Chen, J. and Hoek, G. (2020). Long-term exposure to pm and all-cause and cause-specific
549 mortality: A systematic review and meta-analysis. *Environment international*, page 105974.
- 550 Crouse, D. L., Peters, P. A., Hystad, P., Brook, J. R., van Donkelaar, A., Martin, R. V.,
551 Villeneuve, P. J., Jerrett, M., Goldberg, M. S., Pope III, C. A., et al. (2015). Ambient pm2. 5,
552 o3, and no2 exposures and associations with mortality over 16 years of follow-up in the
553 canadian census health and environment cohort (canhec). *Environmental health*
554 *perspectives*, 123(11):1180.
- 555 Crouse, D. L., Peters, P. A., van Donkelaar, A., Goldberg, M. S., Villeneuve, P. J., Brion, O.,
556 Khan, S., Atari, D. O., Jerrett, M., Pope III, C. A., et al. (2012). Risk of nonaccidental and
557 cardiovascular mortality in relation to long-term exposure to low concentrations of fine
558 particulate matter: a canadian national-level cohort study. *Environmental health perspectives*,
559 120(5):708–714.
- 560 de Hoogh, K., Chen, J., Gulliver, J., Hoffmann, B., Hertel, O., Ketzler, M., Bauwelinck, M.,
561 van Donkelaar, A., Hvidtfeldt, U. A., Katsouyanni, K., et al. (2018). Spatial pm2. 5, no2, o3
562 and bc models for western europe—evaluation of spatiotemporal stability. *Environment*
563 *international*, 120:81–92.
- 564 de Hoogh, K., Gulliver, J., van Donkelaar, A., Martin, R. V., Marshall, J. D., Bechle, M. J.,
565 Cesaroni, G., Pradas, M. C., Dedele, A., Eeftens, M., et al. (2016). Development of west-
566 european pm2. 5 and no2 land use regression models incorporating satellite-derived and
567 chemical transport modelling data. *Environmental research*, 151:1–10.
- 568 Deboosere, P. and Gadeyne, S. (2002). Can regional patterns of mortality in belgium be
569 explained by individual socio-economic characteristics? *Reflats et perspectives de la vie*
570 *economique*, 41(4):87–103.
- 571 Di, Q., Wang, Y., Zanobetti, A., Wang, Y., Koutrakis, P., Choirat, C., Do-
572 minici, F., and Schwartz, J. D. (2017). Air pollution and mortality in the medicare population. *New England*

573 Journal of Medicine, 376(26):2513– 2522.

574 Dowle, M. and Srinivasan, A. (2017). data.table: Extension of ‘data.frame’. R package
575 version 1.10.4.

576 EEA (accessed 12 April 2021). Eea, 2014. airbase - the european air quality database, version
577 8. https://www.eea.europa.eu/ds_resolveuid/3c756b2021754f6bba40447397d67fdf.

578 EEA (accessed 12 April 2021). European environment agency 2019 - corine land cover -
579 nomenclature. https://www.eea.europa.eu/ds_resolveuid/d59f95e5a3f6a305ea04907cfd7ff5ad.

580 Eeftens, M., Beelen, R., de Hoogh, K., Bellander, T., Cesaroni, G., Cirach, M., Declercq, C.,
581 Dedele, A., Dons, E., de Nazelle, A., et al. (2012a). Development of land use regression
582 models for pm_{2.5}, pm_{2.5} absorbance, pm₁₀ and pm_{coarse} in 20 european study areas;
583 results of the escape project. Environmental science & technology, 46(20):11195–11205.

584 Eeftens, M., Tsai, M.-Y., Ampe, C., Anwander, B., Beelen, R., Bellander, T., Cesaroni, G.,
585 Cirach, M., Cyrus, J., de Hoogh, K., et al. (2012b). Spatial variation of pm_{2.5}, pm₁₀, pm_{2.5}
586 absorbance and pm_{coarse} concentrations between and within 20 european study areas and the
587 relation- ship with no₂—results of the escape project. Atmospheric Environment, 62:303–317.

588 Eisen, E., Agalliu, I., Thurston, S., Coull, B., and Checkoway, H. (2004). Smoothing in
589 occupational cohort studies: an illustration based on penalised splines. Occupational and
590 environmental medicine, 61(10):854– 860.

591 Eurostat (2018). Regions in the european union - nomenclature of territorial units for
592 statistics - nuts 2016/eu-28. doi:10.2785/475524.

593 Galobardes, B., Lynch, J., and Smith, G. D. (2007). Measuring socioeconomic position in
594 health research. British medical bulletin, 81(1):21.

595 Gariazzo, C., Carlino, G., Silibello, C., Tinarelli, G., Renzi, M., Finardi, S., Pepe, N.,
596 Barbero, D., Radice, P., Marinaccio, A., et al. (2021). Impact of different exposure models
597 and spatial resolution on the long-term effects of air pollution. Environmental Research,
598 192:110351.

599 Greenberg, J. A. and Mattiuzzi, M. (2015). gdalUtils: Wrappers for the Geospatial Data
600 Abstraction Library (GDAL) Utilities. R package version 2.0.1.7.

601 Hijmans, R. J. (2016). raster: Geographic Data Analysis and Modeling. R package version

602 2.5-8.

603 Hooyberghs, J., Mensink, C., Dumont, G., and Fierens, F. (2006). Spatial interpolation of
604 ambient ozone concentrations from sparse monitoring points in Belgium. *Journal of*
605 *Environmental Monitoring*, 8(11):1129–1135.

606 Huang, S., Li, H., Wang, M., Qian, Y., Steenland, K., Caudle, W. M., Liu, Y., Sarnat, J.,
607 Papatheodorou, S., and Shi, L. (2021). Long-term exposure to nitrogen dioxide and mortality:
608 A systematic review and meta-analysis. *Science of The Total Environment*, page 145968.

609 Huangfu, P. and Atkinson, R. (2020). Long-term exposure to NO₂ and O₃ and all-cause and
610 respiratory mortality: A systematic review and meta-analysis. *Environment International*,
611 144:105998.

612 Hvidtfeldt, U. A., Severi, G., Andersen, Z. J., Atkinson, R., Bauwelinck, M., Bellander, T.,
613 Boutron-Ruault, M.-C., Brandt, J., Brunekreef, B., Cesaroni, G., et al. (2020). Long-term
614 low-level ambient air pollution exposure and risk of lung cancer—a pooled analysis of 7
615 European cohorts. *Environment International*, 146:106249.

616 Jerrett, M., Turner, M. C., Beckerman, B. S., Pope III, C. A., van Donkelaar, A., Martin, R.
617 V., Serre, M., Crouse, D., Gapstur, S. M., Krewski, D., et al. (2016). Comparing the health
618 effects of ambient particulate matter estimated using ground-based versus remote sensing
619 exposure estimates. *Environmental Health Perspectives*, 125(4):552–559.

620 Klompmaker, J. O., Janssen, N., Andersen, Z. J., Atkinson, R., Bauwelinck, M., Chen, J., de
621 Hoogh, K., Houthuijs, D., Katsouyanni, K., Marra, M., et al. (2020). Comparison of
622 associations between mortality and air pollution exposure estimated with a hybrid, a land-use
623 regression and a dispersion model. *Environment International*, 146:106306.

624 Lefebvre, W. and Vranckx, S. (2013). Validation of the ifdm-model for use in urban
625 applications, study accomplished in the framework of the atmosys-project. In VITO report
626 2013/RMA/R/56.

627 Pinault, L., Tjepkema, M., Crouse, D. L., Weichenthal, S., van Donkelaar, A., Martin, R. V.,
628 Brauer, M., Chen, H., and Burnett, R. T. (2016). Risk estimates of mortality attributed to low
629 concentrations of ambient fine particulate matter in the Canadian community health survey
630 cohort. *Environmental Health*, 15(1):18.

631 R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R

632 Foundation for Statistical Computing, Vienna, Austria.

633 Roux, A.-V. D. (2007). Neighborhoods and health: where are we and where do we go from
634 here? *Revue d'épidémiologie et de sante publique*, 55(1):13– 21.

635 RStudio Team (2019). RStudio: Integrated Development Environment for R. RStudio, Inc.,
636 Boston, MA.

637 Temam, S., Burte, E., Adam, M., Ant´o, J. M., Basagana, X., Bousquet, J., Carsin, A.-E.,
638 Galobardes, B., Keidel, D., Kuñzli, N., et al. (2017). Socioeconomic position and outdoor
639 nitrogen dioxide (no₂) exposure in western europe: A multi-city analysis. *Environment*
640 *international*, 101:117–124.

641 Therneau, T. M. (2015). A Package for Survival Analysis in S. version 2.38. Therneau, T. M.
642 (2018). *coxme: Mixed Effects Cox Models*. R package version 2.2-7.

643 Van Hemelrijck, W. M., Willaert, D., and Gadeyne, S. (2016). The geographic pattern of
644 Belgian mortality: can socio-economic characteristics explain area differences? *Archives of*
645 *Public Health*, 74(1):22.

646 Vodonos, A., Awad, Y. A., and Schwartz, J. (2018). The concentration- response between
647 long-term pm_{2.5} exposure and mortality; a meta-regression approach. *Environmental*
648 *research*, 166:677–689.

649 W.H.O. (2004). *International statistical classification of diseases and related health problems*,
650 volume 1. World Health Organization.

651 Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer- Verlag New
652 York.