Main methodological issues in a prospective study on plasma concentrations of persistent organic pollutants and pancreatic cancer risk within the EPIC cohort

Authors: fully provisional / to be decided.

Magda Gasull[1,2,3]*, José Pumarega[1,3]*, Hannu Kiviranta[4], Panu Rantakokko[4], Ole Raaschou-Nielsen[5], Ingvar A. Bergdahl[6,7], Torkjel Manning Sandanger[8], Fernando Goñi[2,9], Lluís Cirera[2,10], Juan Alguacil[2,11], Mar Iglesias[12], Anne Tjønneland[5], Kim Overvad[13], ... {France, Germany, Greece}… Alessio Naccarati[14] {Italy, The Netherlands}…, Elisabete Weiderpass[8,15,16,17], … {Asturias} ... Miguel Rodríguez-Barranco[2,18], Amaia Molinuevo[2,9], María-Dolores Chirlaque[2,10,19], … {Navarra} …, Malin Sund[20], … {UK}, …

¿Tim Key, Weimin Ye, Mazda Jenab, Dominique Michaud, Giuseppe Matullo, Rudolf Kaaks, Federico Canzian, Alexandra Nieters, Marie-Christine Boutron, Ute Nöthlings, Suzanne Jeurnink, Veronique Chajes, Marco Matejcic, ...?

→ other EPIC co-authors ... ?

¿... Marc Gunter, Elio Riboli, Antoni Agudo, Carlos Alberto Gonzalez? Eric Duell[21], Bas Bueno-de-Mesquita[22,23,24], Paolo Vineis[14,22], Miquel Porta[1,2,3]

Affiliations of authors:

1 Hospital del Mar Institute of Medical Research (IMIM), Barcelona, Catalonia, Spain.

2 Universitat Autònoma de Barcelona, Barcelona, Catalonia, Spain.

3 CIBER de Epidemiología y Salud Pública (CIBERESP), Madrid, Spain.

4 National Institute for Health and Welfare, Department of Health Security, Kuopio, Finland.

5 Danish Cancer Society Research Center, Copenhagen, Denmark.

6 Department of Biobank Research, Umeå University, Umeå, Sweden.

7 Occupational and Environmental Medicine, Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden.

8 Department of Community Medicine, UiT-The Arctic University of Norway, Tromsø, Norway.

9 Subdirección de Salud Pública de Gipuzkoa, Gobierno Vasco, San Sebastian, Spain.

10 Department of Epidemiology, Murcia Regional Health Council, IMIB - Arrixaca, Murcia, Spain.

11 Universidad de Huelva, Huelva, Spain.

12 Department of Pathology, Hospital del Mar (PSMar), Barcelona, Spain.

13 Section for Epidemiology, Department of Public Health, Aarhus University, Aarhus, Denmark.

14 Molecular and Genetic Epidemiology Unit, Italian Institute for Genomic Medicine (IIGM), Turin, Italy.

15 Cancer Registry of Norway, Institute of Population-Based Cancer Research, Oslo, Norway.

16 Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

17 Genetic Epidemiology Group, Folkhälsan Research Center, Faculty of Medicine, University of Helsinki, Helsinki, Finland.

18 Escuela Andaluza de Salud Pública. Instituto de Investigación Biosanitaria, Granada. Hospitales Universitarios de Granada/Universidad de Granada, Granada, Spain.

19 Department of Health and Social Sciences, University of Murcia, Murcia, Spain.

20 Department of Surgical and Perioperative Sciences, Umeå University, Umeå, Sweden.

21 Unit of Nutrition and Cancer, Catalan Institute of Oncology (ICO-Idibell), Barcelona, Spain.

22 Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, United Kingdom.

23 Department for Determinants of Chronic Diseases (DCD), National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands.

24 Department of Social & Preventive Medicine, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia.

. . . . . .

*These authors contributed equally.

Corresponding author: Miquel Porta, Clinical & Molecular Epidemiology of Cancer Unit, Hospital del Mar Institute of Medical Research (IMIM), Universitat Autònoma de Barcelona, Carrer del Dr. Aiguader 88, E-08003 Barcelona, Catalonia, Spain. mporta@imim.es
Telephone: +34 93 316 0700 and +34 93 316 0400. Fax: +34 93 316 0410.

ABBREVIATIONS
BMI, body mass index; CI, confidence interval; DDE, dichlorodiphenyldichloroethene; DDT, dichlorodiphenyltrichloroethane; EPIC, European Prospective Investigation into Cancer and Nutrition; HCB, hexachlorobenzene; HCH, hexachlorocyclohexane; PBDE, polybrominated diphenyl ethers; PCBs, polychlorinated biphenyls; PeCB, pentachlorobenzene; POPs, persistent organic pollutants.

ABSTRACT

The use of biomarkers of environmental exposure to explore new risk factors of pancreatic cancer presents clinical, logistic, and methodological challenges that may also be relevant in research on other complex diseases. The objectives of the present report are two-fold: first, to summarize the main design features of a prospective case-control study –nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort– on plasma concentrations of persistent organic pollutants (POPs) and pancreatic cancer risk; and second, to assess the main methodological challenges posed by associations among characteristics and habits of study participants, fasting status, time from blood draw to cancer diagnosis, disease progression bias, basis of cancer diagnosis, and plasma concentrations of lipids and POPs.

Study subjects were 1,533 participants (513 cases and 1,020 matched controls) enrolled between 1992 and 2002. Plasma concentrations of 22 POPs were measured by gas chromatography - high triple quadrupole mass spectrometry (GC-MS/MS). To estimate the magnitude of the associations we calculated multivariate-adjusted odds ratios by unconditional logistic regression, and adjusted geometric means by General Linear Regression Models.

There were differences among countries for subjects' characteristics (as age, gender, smoking, lipid and POP concentrations), and for study characteristics (as time from blood collection to index date, year of last follow-up, length of follow-up, basis of cancer diagnosis, fasting status). Adjusting for centre and time of blood collection, men and obese participants were more likely to had fasted more hours, whilst age, smoking, alcohol intake and physical activity were not significantly associated with fasting status. As expected, no relationship in fasting status was observed between cases and controls, since fasting status was one of the matching factors. Lipid concentrations were related to age, BMI, fasting, country, and smoking. We assessed the possible occurrence of disease progression bias (DPB) in eight situations defined by lipid and POP measurements, on one hand, and by four factors: interval from blood draw to index date, tumour site, tumour stage, and grade of differentiation. In seven of the eight situations results argued against the occurrence of a DPB. We detected and quantified 16 of the 22 POPs in more than 90% of individuals. All 22 POPs were detected in some participants, and the smallest number of POPs detected in one person was 15 (median, 19) with few differences by country.

The coexistence of differences across study centres in some design features and participant characteristics is of relevance to other multicentre studies. Associations among subjects' characteristics and between such characteristics and design features may play important roles in the forthcoming analyses on the association between plasma concentrations of POPs and pancreatic cancer risk.

INTRODUCTION

Increasing current knowledge on environmental causes of pancreatic cancer remains elusive. This might partly be due to the difficulties that such a biologically and clinically aggressive disease poses to obtain biological specimens, to use valid biomarkers, to elicit accurate information from severely ill patients, and sometimes to achieve a precise anatomo-pathological diagnosis [1]. Yet some of such clinical, logistic, and methodological challenges are also common in studies on other diseases [2- 8]. Notably, biomarkers of lipophilic contaminants are prone to disease progression bias, a mechanism of reverse causation bias through which the pathophysiological progression of the disease alters body concentrations of the contaminants in blood and fatty tissues; as a consequence, disease-altered exposure estimates lack etiologic significance [9,10]. Prior to analysing data from a cohort-nested case-control study on pancreatic cancer and persistent organic pollutants (POPs) we carefully considered these methodological issues.

In addition to the above, a related important issue is the influence of fasting status at blood collection on blood concentrations of a variety of lipophilic substances (e.g., some vitamins and other nutrients, most organochlorine compounds). An also related and unresolved issue that affects many etiologic studies is how to approach conceptually and analytically the concomitant confounding and mediating effects of blood lipids, fasting, and body mass index when estimating possible causal effects of biomarkers of such lipophilic substances [11].

Diseases whose diagnostic accuracy and precision in clinical practice depend on age, gender, lifestyle or other factors [1,12,13,14] offer the opportunity to assess whether the diagnostic basis, and the corresponding diagnostic certainty, contribute to disease misclassification, and hence to bias causal estimates.

While prospective longitudinal designs as cohort-nested case-control studies can overcome several of the previously sketched difficulties, control of biases associated with disease progression cannot rest exclusively on such studies [10,15,16,17].

The objectives of the present report are two-fold: first, to summarize the main design features of a prospective case-control study –nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort– on plasma concentrations of persistent organic pollutants (POPs) and pancreatic cancer risk; and second, to assess the main methodological challenges posed by potential associations among characteristics and habits of study participants, fasting status, time from blood draw to cancer diagnosis, disease progression bias, basis of cancer diagnosis, and plasma concentrations of lipids and POPs.

METHODS AND RESULTS

Study population

We performed a case-control study nested within the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort. The EPIC cohort has been previously described in detail [18]. Briefly, 521,457 subjects (153,447 men) aged 35–70 years old were recruited between 1992 and 2000 by 23 collaborating centres from 10 European countries. Three bio-repositories from the EPIC study contributed samples for the present study: the repository from Denmark, which centralized samples from the collaborating centres of Aarhus and Copenhagen; the repository from the collaborating centre of Umeå, in Sweden; and the IARC central repository, which centralized the biospecimens of 8 countries (Germany, United Kingdom, Netherlands, Italy, Spain, Greece, France, and Norway) (Supplemental Table 1).

Over 98% of the 1,533 participants in the present study (cases plus controls) were enrolled between 1992 and 1998. They were followed until cancer diagnosis, death, migration, or the end of the follow-up period (2007, 2010 and 2014 for Denmark, IARC, and Umeå, respectively), whichever occurred first (Supplemental Table 1). The mean length of follow up for the study participants was 11.4 years (standard deviation [SD], 3.4 years).

Pancreatic cancer cases, coded as C25 (C25.0-25.3, 25.7-25.9) according to the International Classification of Diseases-Oncology (ICD-O) 3rd edition, were identified and included in the study. Exclusion criteria were: a) cases of endocrine pancreatic cancer; b) occurrence of other malignant tumours preceding the diagnosis of pancreatic cancer, except for non-melanoma skin cancer; c) cases with pancreatic cancer diagnosed during the first 2 years of blood draw (5 years for cases from Denmark); and d) 12 cases with less than 2 straws of plasma remaining available. Of the final 513 cases included in the present study, 135 (26.3%) came from Denmark, 79 (15.3%) from Umeå (Sweden), and 299 from the IARC central repository (main contributing countries: Germany, 13.6% of the 513 cases; United Kingdom, 11.7%; The Netherlands, 9.2%; Italy, 8.4%; Spain, 7.2%; and Greece, 5.5%) (Supplemental Table 1).

For each case, two control subjects alive and free of cancer at the time of diagnosis of the index case were selected using an incidence density sampling procedure [19]; only 6 cases had just one control. Thus, a total of 1020 matched controls were included. Matching factors were study centre, sex, age at blood collection (± 1 year), date (± 6 months) and time (± 2 h) at blood collection, fasting status (<3 h, 3-6 h, >6 h after last meal), and, for women, use of exogenous hormones (yes, no). The age at blood collection of 11 controls differed by more than ± 1 year (up to 4.8 years) to the age of their matched case; the date at blood collection of 11 controls (4

of the previous) differed by more than ± 6 months (up to 12 months); and the time at blood collection of 83 controls differed by more than ± 2 h (up to 3.5 h) (information on time at blood collection was missing for all 235 individuals from Umeå and for 27 individuals of 9 casesets from IARC repository). The fasting status of 55 controls was different from the fasting status of their matched case, and was missing for 21 participants. Finally, use of exogenous hormones of 8 controls was different from use of their matched case, and was missing for 92 women.

The EPIC study was approved by the Ethical Review Board of the International Agency for Research on Cancer (IARC, Lyon) and by the local Ethical Committees.

Main variables, and collection of blood samples

A questionnaire collected baseline information about sociodemographic characteristics, lifestyles (especially those related to cancer aetiology, such as lifetime history of diet, alcohol and smoking), and medical history. Anthropometric measures and blood samples were taken at recruitment. Over 80% of participants in the present study underwent blood extraction the same day of the recruitment, and 11% during the previous or the following 7 days after recruitment; only 3% of participants had their blood collected more than one year after being enrolled.

For a subset of pancreatic cancer cases information on characteristics of the tumour was also collected. Information on the pancreatic subsite of the tumour was available for 348 of the 513 cases included in the study (68%). Information on the stage of the tumour was available for 177 cases (35%) as it was registered just in some centres. The category 'overlapping' of site of the tumour included tumours registered as an overlapping lesion of the pancreas (C25.8, N = 18), as localised in a pancreatic duct (C25.3, N = 2), and in other specified parts of the pancreas (C25.7, N = 3). The EPIC classification for stage of the tumour included: in situ (none), localised (N = 51), metastatic (N = 49), metastatic regional (N = 33), and metastatic distant (N = 44). For the present study the three last categories were joined in one category named 'metastatic'. A minimum of one basis of the cancer diagnosis was recorded for 506 cases (98.6%): 50 cases had three bases of cancer diagnosis recorded, 156 cases had two bases, and 300 cases had one basis. Only 86 cases (17%, all from IARC bio-repository) had information on the grade of the tumour: 11, 45 and 29 cases were classified as well, moderately, and poorly differentiated, respectively, while 1 case was classified as undifferentiated.

Statistical analyses

Univariate statistics were computed as customary [20,21]. To assess differences on participants' characteristics by gender, case-control status, fasting status, basis of cancer diagnosis, and lipid and persistent organic pollutant (POP) concentrations, Student's t-test, ANOVA, Kruskal-Wallis, and Mann-Whitney's U tests were used. Fisher's exact test for homogeneity was applied to assess the relationship between two categorical variables [20,21]. Spearman's rank correlation coefficient (ρ) was computed to evaluate correlations among pairs of POPs.

To estimate the magnitude of the associations between participants' characteristics and *a)* fasting status, and *b)* basis of cancer diagnosis, multivariate-adjusted odds ratios (ORs) and their corresponding 95% confidence intervals (CI) were calculated by unconditional logistic regression [19]. The main effects of all predictors were independently explored in base models. Final models for fasting status were adjusted for centre and time at blood collection, while final models for basis of cancer diagnosis were adjusted for age at diagnosis, sex, and centre.

General Linear Regression Models (GLM) were applied to study the relationships of participants' characteristics with lipid concentrations [20]. GLM were also used to study the relation between lipid or POP concentrations in plasma samples collected during recruitment and some variables related to the disease, such as the time elapsed between blood collection and the date of pancreatic cancer diagnosis, or some characteristics of the tumour (among cases with available information). Results are expressed as adjusted geometric means (aGMs) with the corresponding 95% CIs. The Kolmogorov-Smirnov test for normality was used to check the distributions of lipid and POP concentrations, and of the time between blood collection and cancer diagnosis; as none was normal, log-transformed values were used in regression analyses. The following potential confounders were included in the final models: age, sex, BMI, fasting status, and centre (repository in models including tumour characteristics).

We assessed associations between plasma concentrations of POPs and study and participant characteristics. POP concentrations were entered in the models as quartile categories (defined using the concentrations in controls).

To assess exposure to multiple compounds, we computed *a)* the sum of PCBs for each participant by adding plasma concentrations of all ten PCBs, and then assigning each participant to one quartile of the new variable (sum of PCBs) [22]; *b)* the sum of four PCBs for each participant by adding plasma concentrations of 4 prevalent PCBs (congeners 118, 138, 153, and 180), and then assigning each participant to one quartile of the new variable (sum of 4 PCBs); and *c)* to compute the sum of orders of the 6 most prevalent organochlorine pesticides,

each compound was categorized in quartiles and the category number of each compound was summed, producing a value ranging between 6 (when concentrations of all 6 organochlorine pesticides were in the lowest quartile) and 24 (when concentrations of all 6 compounds were in the top quartile) [22,23,24].

For the 16 compounds quantified in more than 90% of individuals (see below) we calculated the number of POPs detected in each person at high concentrations (nPhc) as follows: for each subject we added the number of POPs whose plasma concentrations were equal to or greater than a selected cut-off point, as percentile 90 (P90) (the upper decile), or percentile 75 (P75) (the upper quartile) [22,25].

Based on historical factors related to exposure to POPs (essentially, before and after World War II) [26], the following birth cohorts were defined: participants born from 1919 to 1938 (N = 817), from 1939 to1945 (N = 503), and from 1946 to 1964 (N = 213).

The level of statistical significance was set at 0.05 and all tests were two tailed. Analyses were conducted using SPSS version 18 (SPSS, Armonk, NY, USA, 2009) and R version 3.1.3 (R Core Team, Vienna, Austria, 2015).

Baseline characteristics of study participants

At blood collection, the age of the 1,533 individuals included in the study ranged between 29 and 76 years (mean, 56.8 years). Over 58% of participants had overweight or obesity, 26% were current smokers, and 28% consumed more than 18 g of alcohol per day (Table 1). Before reporting etiologic analyses in a future paper, it is useful to summarize the essential characteristics of the entire population of participants. Nevertheless, some baseline characteristics of cases and controls will be presented separately below.

787 participants (51%) were women and 746 (49%) were men. Significant differences between men and women were observed for all main characteristics. The proportion of women was higher among participants with the lowest educational level (less than primary completed) (70% of such individuals were women), among never smokers (67%), among never drinkers and among those with an intake of 0-6 g of alcohol per day (77% and 65%, respectively), and among participants with a moderately active physical activity (63%). The percentage of women was also higher among participants with a fasting status at blood collection between 3 and 6 hours (60%) (Table 1).

A higher proportion of men was observed among overweight participants (58% of men), participants with the highest educational level (65%), former and current smokers (around 60%), participants with the highest alcohol consumption (more than 18 g alcohol/day) (72%), participants physically inactive (71.5%), and among participants with diabetes (56%) (Table 1).

Differences in baseline characteristics were also observed by country (Supplemental Table 1). The percentage of women was significantly lower in Denmark (38%) and Germany (37%), and higher in the United Kingdom (65%) and The Netherlands (79%), while in France and Norway all participants were women (N = 24 and 18, respectively). Denmark had the highest proportion of current smokers (36%), followed by Italy (34%).

Participants from Sweden were younger at blood collection and had a lower body mass index (BMI) than participants from Denmark and the IARC central repository. As mentioned, the end of the follow-up period for participants from the collaborating centre of Umeå (Sweden) was in 2014, while for the rest of centres follow-up ended at least 4 years earlier (2010), and up to 9 years earlier (2005 in France). However, the period of blood collection in Sweden (1992-1996) was similar to other countries. The age at cancer diagnosis of cases from Sweden (mean = 65.4 ± 8.7 years old) was similar to that of cases from the rest of the countries (mean = 65.8 ± 7.6), as expected (Supplemental Table 1). For controls, the index date was the date of cancer diagnosis of the case they were matched with.

As a result, participants from Sweden:
a) had the longest follow-up (they may have been followed for up to 22 years, from 1992 to 2014), and the mean follow-up for Sweden was 15.3 years, while for the rest of the countries it was 10.7 years;
b) had the longest time from blood collection to index date: mean of 12.5 years for Sweden, and 8.3 for the other countries; and
c) were younger at blood collection: mean of 52.9 years for participants from Sweden, and 57.5 for the other countries (Supplemental Table 1).

The above results are of potential relevance if the pancreatic cancer risk function was not linear over measures of time or age; e.g., if risks associated with certain exposures or subject characteristics were essentially or only apparent after 10 years since blood collection, then the characteristics of the study in Sweden would have a stronger influence in the detection of such risks than the other countries. As we shall see later, there were also other differences by study centre, and the coexistence of differences across study centres in some design features and participant characteristics is of relevance to other multicentre studies.

The distribution of the time from blood collection to index date is shown in Figure 1. It was between 2 and <5 years for 14% of participants, between 5 and <10 years for 51% of participants, and ≥10 years for 35%. As mentioned, we included no cases (and thus no matched controls) with a diagnosis of pancreatic cancer within <2 years of blood draw. Only 5% of participants, all from Sweden, had such time interval ≥15 years. In multivariate models (for cases and for controls separately), repository was the only study characteristic statistically significantly associated with time from blood collection to index date, adjusting for age, sex, BMI, and centre; the minor exception were overweight and obese controls, who had slightly longer intervals than normal weight controls (aGM for obese = 8.7 years, 95% CI: 8.3-9.3 vs. aGM for normal weight = 7.8 years, 95% CI: 7.5-8.1; p-value = 0.008). No associations were observed either when subjects from Sweden were excluded.

Baseline characteristics of cases and controls

As expected, cases and controls showed no differences in the variables they were matched by (results not shown). No differences between cases and controls were also observed for BMI, education, marital status, alcohol consumption, and physical activity (Table 2). In contrast, a higher proportion of cases than controls were current smokers at study entry (33% vs. 23%, respectively), and had diabetes mellitus (6% of cases vs. 3% of controls), as expected.

Basis of cancer diagnosis

We next assessed whether the diagnostic basis of the pancreatic cancer, and its corresponding potential disease misclassification, were independent or instead associated with sociodemographic, lifestyle and clinical characteristics of cases. Of the 506 cases (98.6% of all 513 cases) with information on the basis of cancer diagnosis, 382 (75.5%) had been microscopically confirmed, while the remaining 124 (24.5%) had been diagnosed by imaging results, laboratory tests, clinical symptoms, or physical examination (Table 3).

Significant differences in the percentage of cases with microscopic confirmation were observed across countries: more than 95% of cases from Denmark and Sweden were microscopically confirmed, while this figure was 39% in Greece, and 22% in the United Kingdom (Supplemental Table 1).

Microscopic confirmation was higher in men than in women (83% and 68%, respectively). As expected and important, it was lower at older ages: it was about 83% in cases <60 years old, 78% in cases 60-69 years, and 68% in cases ≥70 years (Table 3). Microscopic confirmation was

not related to site or stage of the tumour, neither to BMI. By contrast, in univariate analyses it was associated to smoking, alcohol, and physical activity. In models adjusting for age, sex and centre, only differences by age remained statistically significant. There was also a positive, monotonic, but statistically non-significant relation between education and microscopic basis. Age was a strong confounder of some associations with diagnostic basis. For instance, 82% of current smokers and 72% of never smokers had been diagnosed by microscopic methods; however, since current smokers were younger than never smokers (e.g., 49% vs. 34% were <55 years old, respectively), when adjusting for age the observed relation between smoking and diagnostic basis became much weaker (Table 3).

Fasting status

In descriptive analyses, differences in fasting status were observed for sex, age, study centre, time at blood collection, smoking status, alcohol consumption, and physical activity. Fasting status was one of the variables showing more differences among countries: while a large proportion of participants from Sweden, Italy and Spain had more than six hours of fasting (94%, 83% and 68%, respectively), the corresponding percentage for the rest of countries was <15% (except France, with 7 out of its 24 participants with >6 hours of fasting) (Supplemental Table 1).

Adjusting for centre and time of blood collection, men and obese participants were more likely to had fasted >6 hours, while age, smoking, alcohol intake or physical activity were not statistically significantly associated with fasting status (Table 4). As expected, no relationship was observed with case-control status, since fasting status was one of the matching factors.

Lipid concentrations in plasma

Measurements of total cholesterol and triglycerides were carried out enzymatically by Abbott Architect reagents (Abbott Laboratories, Abbott Park, IL, USA) in plasma obtained at study entry [18]. Total lipids (TL) were calculated by the Standard formula 2, based on total cholesterol and triglycerides [27,28,29].

Among the entire study population (1,533 individuals), mean plasma concentrations of total cholesterol, triglycerides and TL were, respectively, 201, 140 and 658 mg/dL (Supplemental Table 2). As expected because of the study design (i.e., because blood samples were obtained at least 2 years before cancer diagnosis), there were no significant associations between lipid concentrations (total cholesterol, triglycerides, and TL) and case-control status (Figure 2 and

Supplemental Table 2). The absence of association between lipids and case-control status was also evident when adjusting for age, sex, centre, BMI and fasting status (p-value = 0.441) (Table 5).

Total lipids tended to be significantly higher in subjects with higher age and BMI. Men also had slightly higher values than women (Supplemental Table 2). Except for sex, these relationships held when models were mutually adjusted for age, sex, centre, BMI, and fasting status (p-values ≤ 0.036) (Table 5). A statistically significant association was also observed between lipid concentrations and smoking status in adjusted models: participants who reported to be current smokers had higher concentrations of triglycerides and TL (Table 5 and Supplemental Table 3). Adjusted models also show that triglycerides and TL were associated with fasting status: individuals with more than 6 hours of fasting had lower concentrations of triglycerides and of TL.

Differences in lipid concentrations were also observed by country: while participants from Sweden were younger, had a lower BMI, and were more often in fasting status (Supplemental Table 1), they had the highest mean and median concentrations of total cholesterol and total lipids (Supplemental Table 2); participants from Sweden also showed the highest concentrations of triglycerides once adjusting for age, sex, BMI and fasting status (Table 5 and Supplemental Table 3). In adjusted models the lowest concentrations of TL were observed in participants from Spain and Greece: the aGM of TL was 736 mg/dL, 585 mg/dL and 587 mg/dL for subjects from Sweden, Spain and Greece, respectively (Table 5).

Concentrations of persistent organic pollutants (POPs)

POP concentrations assays
POP concentrations were measured in 200µL plasma samples at the National Institute for Health and Welfare (THL), Finland [30]. Twenty-two POPs were measured: three polybrominated diphenyl ethers (PBDEs 47, 99, 153), eight non-dioxin like polychlorinated biphenyls (PCB congeners 74, 99, 138, 153, 170, 180, 183, and 187), two dioxin like PCBs (congeners 118 and 156), and nine organochlorine pesticides or their metabolites: p,p'-DDT (dichlorodiphenyltrichloroethane), p,p'-DDE (dichlorodiphenyldichloroethene), α-HCH (hexachlorocyclohexane), β-HCH, γ-HCH, PeCB (pentachlorobenzene), HCB (hexachlorobenzene), trans-nonachlor, and oxychlordane.

Pretreatment of the samples was as follows: ethanol and ^{13}C-labelled internal standards of each compound in toluene were added to samples (200 µL) in test tubes and mixed to precipitate the proteins and equilibrate internal standards. Dichloromethane-hexane (1:4) was added for extraction followed by activated silica to bind the sample water, ethanol, and precipitate.

Samples were mixed, and layers were allowed to separate. The upper dichloromethane-hexane layer was poured to a solid phase extraction cartridge (SPE cartridge) containing from bottom to top 10% AgNO3 impregmented silica and a mixture of Na2SO4 and silica. The lower layer in the test tube was extracted again with dichloromethane-hexane, which was also poured to SPE-cartridge. Elution of SPE-cartridges was continued with dichloromethane-hexane, and the eluate was concentrated to 15-20 µL for gas chromatography - high triple quadrupole mass spectrometry (GC-MS/MS) analysis. The instrument used was an Agilent 7010 GC-MS/MS system (Wilmington, DE, USA), GC column DB-5MS UI (J&W Scientific, 20m, ID 0.18 mm, 0.18 µm). Limits of detection ranged from 2 pg/mL for PCB congeners and trans-nonachlor to 16 pg/mL for p,p'-DDE. Limits of quantification ranged from 5 pg/mL for PCB congeners and trans-nonachlor to 40 pg/mL for p,p'-DDE (Supplemental Table 4). When a sample had a concentration of a compound below the detection threshold, it was assigned the mid-value of this limit; when a compound was detected but under the quantification threshold, the mid-value between detection and quantification limits was used. The THL laboratory participates three times a year in AMAP interlaboratory comparisons (Ring Test for Persistent Organic Pollutants in human serum, National Institute of Public Health, Quebec, Canada) [30-32]. POP concentrations were individually converted as lipid-based concentrations (i.e., corrected or normalized for TL) by dividing the crude plasma POP concentration by TL (see 'Lipid concentrations' above).

The results that follow refer to the entire study population of cases and controls. The results were selected for inclusion in the present article based on their importance for the primary study on POPs and pancreatic cancer, and on the methodological challenges that the associations among POP concentrations, lipid concentrations, and other study variables and features pose for the primary study and for other studies on POPs and cancer aetiology. Therefore, the sections below do not include results on POPs and pancreatic cancer risk, nor on all determinants of POP and lipid concentrations in plasma in the study subjects.

Percentages of detection and quantification
We detected and quantified 16 of the 22 compounds in more than 90% of individuals (henceforth, 'most prevalent POPs') (Figure 3, Table 6, and Supplemental Table 4). Seven of the 10 PCBs, as well as HCB and trans-nonachlor were detected and quantified in 100% of participants. α-HCH and PeCB were detected in more than 80% of participants and quantified in 49% and 35%, respectively. The percentage of detection for the other 4 compounds (γ-HCH, and PBDEs 47, 99 and 153) ranged between 8% and 47%. Thus, all 22 POPs were detected in some participants, and no individual was free from POPs: the smallest number of POPs detected in one person was 15. The median number of POPs detected per person was 19, with no differences by subjects' individual characteristics, and very few by country (only Sweden and United Kingdom had medians of 18 and 20 POPs detected per person, respectively). Twenty or

more compounds were detected in 25% of participants; this was so in 28% of participants born in 1919–1938 and in 21% of subjects born in 1939–1964. The number of POPs quantified per person ranged between 21 and 11, and only 6% of participants had less than 16 compounds quantified.

Correlations among POPs

The highest Spearman's correlation coefficients (ρ) among crude concentrations of the 16 most detected POPs were observed between pairs of PCBs; e.g., the ρ between PCB 170 and 180 was 0.985, and between PCB 138 and 153 it was 0.981). The ρ between oxychlordane and trans-nonachlor was 0.809. The ρ between DDT and DDE was 0.800, and between HCB and β-HCH, 0.744. The rest of correlations between pairs of organochlorine pesticides ranged from 0.159 (for β-HCH and trans-nonachlor) to 0.711 (for DDT and β-HCH) (all p-values <0.001). Finally, correlation coefficients between PCBs and organochlorine pesticides ranged from 0.017 (for PCB 156 and β-HCH, p-value = 0.496) to 0.616 (for PCB 74 and oxychlordane, p-value < 0.001). When TL-corrected concentrations of POPs were analysed, ρ's were only slightly attenuated, and the results just mentioned remained virtually unaltered.

Concentrations of POPs

The highest concentrations were found for DDE, PCBs 153 and 180 (median concentration: 3371, 1023, and 810 pg/mL, respectively). For the rest of PCBs median concentrations ranged from 66 pg/mL for PCB 74 to 635 pg/mL for PCB 138; and for the rest of organochlorine pesticides the corresponding values ranged from 55 pg/mL for oxychlordane to 393 pg/mL for HCB (Table 6). 39% of participants had one or more of the 16 most prevalent POPs at concentrations above their respective P90 (i.e., 61% of participants had each of the 16 POPs at concentrations below their respective P90). The corresponding figures for P75 were 69% and 31%. Figures were similar for TL-corrected POPs (e.g., 42% of participants had one or more of the 16 TL-corrected POPs above their respective P90). For both cut-offs (P90 and P75), the geometric means of the nPhc were 2.9 and 4.3, respectively (2.8 and 4.3 for TL-corrected POPs).

Women had statistically significant higher concentrations of HCB (median 29% higher than men's) and β-HCH (median 49% higher), while men had higher concentrations of trans-nonachlor (median 34% higher than women's) and of PCBs (except PCBs 74 and 118) (Table 6). No significant differences between men and women were observed for DDT, DDE and oxychlordane. Differences in univariate analyses were also observed by age, BMI, fasting status, and country.

Median POP concentrations were higher in subjects with greater age and BMI; e.g., median concentrations of HCB were 1.7 times higher in the oldest group than in the youngest, and 2.5

times higher in obese participants than in normal-weight individuals (Table 6). However, as often reported in the literature [23,26], obese participants had lower concentrations of some PCBs (congeners 153, 156, 170, and 180) than overweight and normal-weight participants. Differences in POP concentrations were also observed by fasting status; higher median concentrations were found in individuals who had fasted more than six hours. Differences observed by sex, age, BMI and fasting status remained significant when POP concentrations were corrected by TL (Supplemental Table 5).

Participants from Spain had the highest concentrations of DDT, HCB, β-HCH and PCBs 183 and 187; remarkably, their median concentration of HCB (4401 pg/mL) was over 18 times higher than the corresponding value of participants from the United Kingdom, the country with the lowest concentrations of HCB (237 pg/mL); it was also two times higher than the corresponding value from the second country with the highest concentrations of HCB (Greece, 2187 pg/mL) (Table 6). Subjects from Italy had the highest median concentrations of oxychlordane, and of PCBs 74, 99 and 118. Germany had the highest concentrations of PCBs 153, 156, 170, and 180. The highest plasma levels of trans-nonachlor and PCB 138 were found in participants from Sweden. Subjects from Greece had the highest concentrations of DDE, but also the lowest concentrations of all PCBs. The lowest concentrations of DDT and DDE were detected in participants from Denmark. Again, similar results were obtained when TL-corrected POP concentrations were analysed (Supplemental Table 5): the highest concentrations (in ng/g of lipid) of DDT, HCB, β-HCH and PCBs 180, 183 and 187 were observed in subjects from Spain, while participants from Germany had the highest concentrations of PCBs 138, 153, 156 and 170.

Relation between lipid and POP concentrations
Spearman's correlation coefficients (ρ) among crude concentrations of POPs and the three lipid measures (total cholesterol, triglycerides and TL) were all positive and statistically significant; the highest values of ρ were 0.464 for oxychlordane and TL, 0.447 for trans-nonachlor and TL, and 0.412 for trans-nonachlor and triglycerides (p-values < 0.001). When models adjusting for age, sex, centre, and BMI were applied the observed associations remained significant, with no differences according to fasting status. However, most associations between lipid and POP concentrations disappeared when concentrations of POPs corrected by TL were used; only oxychlordane and trans-nonachlor remained positively and statistically significantly associated with triglycerides and TL in adjusted models, mainly in participants with less than six hours of fasting.

Disease progression bias

The design of the present case–control study nested within the EPIC cohort is a valid and efficient way to cope with possible biases caused by changes in lipophilic biomarkers of exposure induced by the disease (subclinical or clinical disease); i.e., disease progression bias (DPB) [9,10,33,34]. Moreover, to further control for potential DPB, we did not include in the study cases of pancreatic cancer diagnosed within 2 years of blood draw (5 years for cases from Denmark). Nevertheless, the effectiveness of such design features to control DPB has seldom been empirically tested. Therefore, we studied the relation between lipid and POP concentrations in plasma samples collected during recruitment, and some variables related to the disease. Specifically, we studied: *a)* the relation between lipid concentrations at blood collection and the time elapsed between such collection and the date of pancreatic cancer diagnosis (please see section *Baseline characteristics of study participants*), and *b)* the relation between lipid concentrations and some characteristics of the tumour, such as stage, tumour site within the pancreas, and tumour grade (among cases with available information). We also analysed the same relationships using POP concentrations instead of lipid concentrations.

If subclinical pancreatic cancer or its precursors were already influencing concentrations of lipids in plasma at the time of blood draw (e.g., through metabolic changes, weight loss and lipid mobilization from fatty tissues to blood), we would expect to observe that cases with higher lipid concentrations were diagnosed closer to the time of blood draw than cases with lower concentrations of lipids; an inverse relationship between lipid concentrations at blood draw and the interval from blood draw to cancer diagnosis would suggest that DPB would need to be considered in the etiologic analyses (e.g., latency analyses stratified by time since blood draw would be warranted).

No inverse association was observed between total lipid concentrations of cases and time from blood collection to diagnosis, either in descriptive analyses (Figure 4) or in multivariate regression models adjusting for study centre, age, sex, BMI, and fasting status (Table 7). No inverse associations were observed either for total cholesterol and for triglycerides (results not shown). Actually, in multivariate models we observed a slightly positive association between TL and the time interval: the interval for cases in the upper quartile of TL concentrations was around one year longer than for cases in the lowest TL quartile (aGM 8.6 and 7.4, respectively, Table 7). A similar association was observed when cases with time intervals >15 years and <5 years (only registered in Umeå and in the IARC repository, by design) were excluded. However, when stratifying by fasting status, the association between TL and the interval was only observed in non-fasting individuals (Table 7); there were no associations in cases who had fasted ≥3 hours (p-value for interaction = 0.132). Although similar results were observed in non-fasting controls, again no statistically significant associations between lipid concentrations and

the time from blood collection to the index date were observed in controls. These results argue against the occurrence of a DPB.

The interval from blood collection to cancer diagnosis was also slightly longer in cases with higher concentrations of some POPs (both crude and TL-corrected). For instance, cases with TL-corrected concentrations of HCB in the upper quartile had such interval more than two years longer than cases in the lower quartile once adjusting for study centre, age, sex, BMI, and fasting status (aGM of the interval from blood collection to cancer diagnosis for HCB upper and lower quartiles: 9.6 and 7.3 years, respectively; p-value = 0.003) (Supplemental Table 6); this association remained statistically significant among cases who had fasted ≥3 hours. A similar association was found for the sum of PCBs 118, 138, 153 and 180 (aGM for upper quartile = 8.8 years vs. aGM for the lower quartile = 7.5 years, p-value = 0.025); this association was not present in cases who had fasted ≥3 hours (data not shown). No associations were found between the mentioned interval and other compounds, such as DDT, DDE and β-HCH. Although attenuated, the interval from blood collection to index date was also slightly longer in controls with higher concentrations of some POPs (both crude and TL-corrected). These results argue against the occurrence of a DPB.

We also studied the relation between lipid or POP concentrations and some characteristics of the tumour and, based on previous work [1,10,33,35], we hypothesised that if subclinical pancreatic cancer were already causing subtle pathophysiologic changes at the time of blood draw (e.g., through moderate weight loss, lipid mobilisation or other metabolic changes), we would observe that cases later diagnosed with a more advanced disease (e.g., a metastatic tumour) or a more aggressive disease (e.g., a poorly differentiated tumour) would have higher lipid and POP concentrations (again, at blood draw). Among the limited number of cases with the available information (see section *Main variables, and collection of blood samples*), stage and pancreatic site of the tumour were not associated with any of the three lipid measures (total cholesterol, triglycerides and total lipids, Table 8). These results suggest that a DPB did not occur.

However, in the 85 cases with data on grading of the tumour and on adjusting variables, descriptive and multivariate analyses showed that cases with less differentiated tumours had higher concentrations of triglycerides and TL at study entry than cases diagnosed with more differentiated tumours (aGM of TL for poorly differentiated tumours = 636 vs. aGM for well differentiated tumours = 539; p-value = 0.025) (Table 8). These associations held when cases with longer times from blood extraction to cancer diagnosis (e.g., ≥10 years) were excluded. These results suggest that DPB might occur.

Similar results for the presence or absence of the associations between all three tumour characteristics and the three lipid measures were found when cases in non-fasting status were excluded from the analyses.

Finally, lower TL-corrected concentrations of some POPs (as always, at baseline, >2 years prior to the diagnosis of pancreatic cancer) were observed in cases with metastatic tumours (N = 126) than in cases with localised tumours (N = 51), once again adjusting for centre, age, BMI and fasting status. Differences were statistically significant only for the sum of PCBs 118, 138, 153, and 180, and for the sum of all PCBs: aGM of TL-corrected concentrations of all PCBs for metastatic tumours = 548 ng/g (95% CI: 503 - 597) vs. aGM for localised tumours = 679 (95% CI: 578 - 798) (p-value = 0.050). There were no differences in TL-corrected concentrations of POPs by tumour site. Contrary to the hypothesis supporting the existence of DPB, TL-corrected concentrations of POPs were slightly lower in cases with poorly differentiated tumours than in cases with moderately differentiated and with well differentiated tumours. These results argue against the occurrence of a DPB.

DISCUSSION

We observed a remarkable number of associations among variables that may play roles in analyses on the possible influence of POPs on pancreatic cancer risk and therefore need to be considered in the upcoming study of how POPs and pancreatic cancer risk are related. Thus, lipid concentrations were related to age, BMI, fasting status, country, and smoking; while fasting status, once adjusted for centre and time at blood collection, was related to sex and BMI. Differences among countries were observed for subjects' characteristics (as sex, smoking, alcohol consumption, physical activity, and diabetes, but not age or BMI), and for study characteristics (as year of last follow-up, length of follow-up and, hence, time from blood collection to index date, basis of cancer diagnosis, fasting status, and for lipid and POP concentrations). We found no differences in lipid concentrations between pancreatic cancer cases and controls. Although expected because of the study design, this fact needed confirmation.

The fact that lipid concentrations were so similar in cases and controls (Figure 2) shows the methodological progress that the present design represents with respect to previous studies on POPs and pancreatic cancer risk (Supplemental Tables 7 and 8). Other methodological characteristics and findings of previous studies are also summarized in the mentioned Tables.

Adjusted models showed increasing concentrations of total lipids with increasing age and BMI. Total lipids were also higher in smokers, in participants with less than six hours of fasting at blood collection, and in Sweden. The specific way the study was conducted in each country determined fasting status.

Therefore, when studying the effects of POP concentrations corrected or normalized by lipids [14,29,33], it may be necessary to take into account the associations observed between lipid concentrations and age, BMI, fasting status, country, and smoking; e.g., because some of the possible effects of these latter variables could partly be adjusted by the lipid correction. Most associations between lipid and POP concentrations waned when concentrations of POPs corrected by TL were used, as it can partly be expected from the work of Phillips et al. [27,28].

The main reason for the correction of POP concentrations by lipids is to remove any variation between fasting and non-fasting individuals [27,28]. Thus, lipid correction may be unnecessary or unwarranted in conditional analyses, since fasting status is one of the matching factors and, therefore, no differences in fasting status exist in most casesets. However, lipid correction of POP concentrations may be preferable than adjusting by fasting if the latter is deemed a less accurate measure than plasma concentrations of lipids. If uncorrected POP concentrations are

used in unconditional analyses, then fasting status and BMI may need to be included as covariates. Models with POPs uncorrected for lipids (and perhaps unadjusted for BMI as well) may be informative and valid too because such variables may be both confounders and mediators in the hypothetical causal chain between POPs and pancreatic cancer [36].

For reasons explained above, we excluded cases diagnosed of pancreatic cancer within 2 years of blood draw. We can envision no plausible biases that such exclusion might create in the analyses on POPs and pancreatic cancer risk.

We assessed the possible occurrence of DPB in eight situations defined by lipid and POP measurements, on one hand, and by four factors: interval from blood draw to index date, tumour site, tumour stage, and grade of differentiation. In seven of the eight situations results argued against the occurrence of a DPB, the possible exception being that cases with less differentiated tumours had higher concentrations of triglycerides and TL at study entry than cases diagnosed with more differentiated tumours. Analyses of the eight situations were possible to different extents due to different numbers of subjects with the necessary information available.

In both pancreatic cancer cases and controls, no inverse association was observed between total lipid concentrations and time from blood collection to diagnosis, either in descriptive analyses (Figure 4) or in multivariate models. These findings argue against the existence of DPB. A slightly positive association between total lipids and the interval from blood collection to diagnosis was observed in non-fasting individuals (i.e., the less reliable subgroup); there were no associations in subjects who had fasted ≥3 hours at blood draw. The interval from blood collection to cancer diagnosis was also slightly longer in cases with higher concentrations of some POPs, both crude and corrected by TL.

Also concerning DPB: among 177 cases with the available data, tumour stage was not associated with any of the three lipid measures; therefore, concentrations of lipids at blood draw were not increased by subclinical, more disseminated tumours.

Among only 85 cases with data on tumour grade, cases with less differentiated tumours had higher concentrations of triglycerides and TL at study entry than cases diagnosed of more differentiated tumours. There were only 11 cases in the reference category of well differentiated tumours. The small numbers do not warrant the conclusion that concentrations of lipids at blood draw were already increased by subclinical poorly differentiated tumours. However, future studies may consider this approach to assessing DPB.

Similarly, the category 'overlapping' for tumour subsite was in our study too heterogeneous; studies with more complete information on subsite may consider assessing whether subclinical tumours in different pancreatic subsites might have altered differently lipids and lipophilic substances at blood draw [10,33].

Lower TL-corrected concentrations of POPs were observed in cases with metastatic tumours than in cases with localised tumours. The direction of the association is the opposite that would be expected in case of DPB. Furthermore, differences were statistically significant only for the sums of PCBs, and the number of cases with information on stage was small. Also not suggesting DPB were analyses of the associations between TL-corrected concentrations of POPs and tumour subsite, grade of differentiation and, as previously mentioned, the interval from blood draw to index date.

Cohort studies, nested case-control studies and other longitudinal designs are often most efficient to cope with DPB [9]. However, biases associated with disease progression still need to be assessed in such studies [10]. First, because in some cohort studies the interval between blood draw and outcome is short for at least a subset of cases [15,16,17]. Second, because such designs sometimes suffer from selection biases due to partial availability of blood samples or limited retrieval of disease-related samples (e.g., tumour tissue) [1]. And third, because they may not have collected relevant etiologic data. Therefore, the empirical tests of DPB reported here are relevant beyond the present study.

We used a classification of diagnostic basis to perform future etiologic analyses stratified on diagnostic basis, as a proxy for diagnostic certainty; these analyses are sometimes referred to as sensitivity analyses. Over 75% of cases had been diagnosed through microscopic methods, which is a common figure in large studies and case series [12,13]. Significant differences in the percentage of cases with microscopic confirmation were observed across countries (>95% of cases from Denmark and Sweden, <40% in other countries). Microscopic confirmation was higher in men, and in younger and more educated cases. It was not related to tumour site or stage, neither to other potentially important variables as BMI, smoking, alcohol, or physical activity. Thus, assessing potential biases due to diagnostic certainty and disease misclassification will be warranted in etiologic analysis not only as a general precaution, but also because diagnostic certainty could be related to exposures associated with factors as age, gender, or education [13].

Finally, exposure to the POPs selected for analysis was quite widespread in the study population, with substantial variability. This is a main strength of EPIC [18]. All 22 POPs analysed were detected in some participants, and the lowest number detected in one person was 15. We detected and quantified 16 compounds (all but 6) in more than 90% of individuals.

42% of participants had one or more of such 16 prevalent POPs at concentrations (corrected by TL) above their respective P90. In a study based on the general population of Catalonia (N = 919) [22], the corresponding figure was 32% (with 8 most prevalent POPs out of 19 POPs analysed). In a study based on the US general population (N >4,000) [25], the corresponding figure was 67%, in part, probably, because the number of POPs analysed was higher (37 most prevalent POPs out of 91 POPs analysed). For participants in the present study and in the studies in Catalonia and the US, the geometric mean of the nPhc (TL-corrected POPs; nPhc cut-off: P90) was 2.8, 2.0 and 3.4, respectively.

The observed correlations between pairs of POPs were expected [22,23,26]. Correlations among certain POPs are often strong worldwide; it is a feature of human contamination by POPs that many studies must address, and so will ours.

Significant differences in plasma concentrations of several POPs (uncorrected and corrected by total lipids) were observed by age, sex, BMI, fasting status, and country. They will need to be considered in the subsequent analyses on POPs and pancreatic cancer risk.

Participants from Spain had the highest concentrations of DDT, β-HCH, PCBs 183 and 187, and particularly of HCB, as also observed in previous studies [37]. Germany had the highest concentrations of some PCBs.

To conclude, the present article not only summarizes the main methodological features of the study, but it also reports a number of associations among study and subjects' characteristics that may play important roles in the forthcoming analyses on the association between plasma concentrations of POPs and pancreatic cancer risk.

REFERENCES

1. Porta M. Role of organochlorine compounds in the etiology of pancreatic cancer: A proposal to develop methodological standards. Epidemiology 2001; 12: 272–276.

2. Wolff MS, Anderson HA, Britton JA, Rothman N. Pharmacokinetic variability and modern epidemiology–the example of dichlorodiphenyltrichloroethane, body mass index, and birth cohort. Cancer Epidemiol Biomarkers Prev 2007; 16: 1925–1930.

3. Wolff MS, Zeleniuch-Jacquotte A, Dubin N, Toniolo P. Risk of breast cancer and organochlorine exposure. Cancer Epidemiol Biomarkers Prev 2000; 9: 271–277.

4. Hoppin JA, Tolbert PE, Holly EA, et al. Pancreatic cancer and serum organochlorine levels. Cancer Epidemiol Biomark Prev 2000; 9: 199–205.

5. Porta M, Malats N, Jariod M, et al. Serum concentrations of organochlorine compounds and K-ras mutations in exocrine pancreatic cancer. Lancet 1999; 354: 2125–2129.

6. De Roos AJ, Hartge P, Lubin JH, et al. Persistent organochlorine chemicals in plasma and risk of non-Hodgkin's lymphoma. Cancer Res 2005; 65: 11214–11226.

7. Vo TT, Gladen BC, Cooper GS, et al. Dichlorodiphenyldichloroethane and polychlorinated biphenyls: intraindividual changes, correlations, and predictors in healthy women from the southeastern United States. Cancer Epidemiol Biomarkers Prev 2008;17:2729–2736.

8. Baris D, Kwak LW, Rothman N, et al. Blood levels of organochlorines before and after chemotherapy among non-Hodgkin's lymphoma patients. Cancer Epidemiol Biomarkers Prev 2000;9:193–197.

9. Porta M, Greenland S, Hernán M, dos Santos Silva I, Last M, eds. A dictionary of epidemiology. 6th edition. New York: Oxford University Press; 2014. p. 78, 276.

10. Porta M, Pumarega J, López T, Jariod M, Marco E, Grimalt JO. Influence of tumor stage, symptoms and time of blood draw on serum concentrations of organochlorine compounds in exocrine pancreatic cancer. Cancer Causes Control 2009; 20: 1893–1906.

11. Gallo V, Egger M, McCormack V, Farmer PB, Ioannidis JPA, Kirsch-Volders M, et al. STrengthening the Reporting of OBservational studies in Epidemiology – Molecular Epidemiology (STROBE-ME). An extension of the STROBE statement. PLoS Medicine 2011; 8 (10): e1001117.

12. Porta M, Malats N, Piñol JL, Rifà J, Andreu M, Real FX, for the PANKRAS I Project Investigators. Diagnostic certainty and potential for misclassification in exocrine pancreatic cancer. J Clin Epidemiol 1994; 47: 1069–1079.

13. Porta M. Commentary on Chapter 7 – Epidemiology. In: Von Hoff DD, Evans DB, Hruban RH, eds. Pancreatic cancer. Boston: Jones and Bartlett; 2005: 113–117.

14. López T, Pumarega J, Pollack AZ, Lee DH, Richiardi L, Jacobs DR Jr, et al. Adjusting serum concentrations of organochlorine compounds by lipids and symptoms: a causal framework for the association with K-ras mutations in pancreatic cancer. Chemosphere 2014; 114: 219–225.

15. Wolff MS, Toniolo PG, Lee EW, et al. Blood levels of organochlorine residues and risk of breast cancer. J Natl Cancer Inst 1993; 85: 648–652.

16. Hunter DJ, Hankinson SE, Laden F, et al. Plasma organochlorine levels and the risk of breast cancer. N Engl J Med 1997; 337: 1253–1258.

17. Dorgan JF, Brock JW, Rothman N, et al. Serum organochlorine pesticides and PCBs and breast cancer risk: results from a prospective analysis (USA). Cancer Causes Control 1999; 10: 1–11.

18. Riboli E, Hunt KJ, Slimani N, Ferrari P, Norat T, Fahey M, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. Public Health Nutr. 2002; 5:1113–1124.

19. Rothman KJ, Greenland S, Lash TL, eds. Modern Epidemiology. 3rd. ed. Philadelphia: Lippincott-Raven; 2008.

20. Armitage P, Berry G, Matthews JNS. Statistical methods in medical research. 4th ed. Oxford, Blackwell; 2002.

21. Kleinbaum DG, Kupper LL, Muller KE, et al. Applied regression analysis and other multivariable methods. 3rd ed. Duxbury, CA: Pacific Grove; 1998.

22. Porta M, Pumarega J, Gasull M. Number of persistent organic pollutants detected at high concentrations in a general population. Environ Int. 2012; 44: 106–111.

23. Gasull M, Pumarega J, Téllez-Plaza M, et al. Blood concentrations of persistent organic pollutants and prediabetes and diabetes in the general population of Catalonia. Environ Sci Technol. 2012; 46: 7799–7810.

24. Gasull M, Castell C, Pallarès N, Miret C, Pumarega J, Téllez-Plaza M, López T, Salas-Salvadó J, Lee DH, Goday A, Porta M. Blood concentrations of persistent organic pollutants and unhealthy metabolic phenotypes in normal-weight, overweight and obese individuals. Am J Epidemiol. 2018; 187: 494–506.

25. Pumarega J, Gasull M, Lee DH, et al. Number of persistent organic pollutants detected at high concentrations in blood samples of the United States population. PLoS One. 2016; 1: e0160432. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0160432 Accessed on 25 April 2018.

26. Porta M, Puigdomènech E, Ballester F, Selva J, Ribas-Fitó N, Llop S, López T. Monitoring concentrations of persistent organic pollutants in the general population: the international experience. Environ Int 2008; 34: 546–561.

27. Phillips DL, Pirkle JL, Burse VW, Bernert JT Jr, Henderson LO, Needham LL. Chlorinated hydrocarbon levels in human serum: effects of fasting and feeding. Arch Environ Contam Toxicol 1989; 18: 495–500.

28. Bernert JT, Turner WE, Patterson DG Jr, Needham LL. Calculation of serum ''total lipid'' concentrations for the adjustment of persistent organohalogen toxicant measurements in human samples. Chemosphere 2007; 68: 824–831.

29. Porta M, Jariod M, López T, Pumarega J, Puigdomènech E, Marco E, et al. Correcting serum concentrations of organochlorine compounds by lipids: alternatives to the organochlorine / total lipids ratio. Environ Int. 2009; 35: 1080–1085.

30. Koponen J, Rantakokko P, Airaksinen R, Kiviranta H. Determination of selected perfluorinated alkyl acids and persistent organic pollutants from a small volume human serum sample relevant for epidemiological studies. J Chromatogr A. 2013; 1309: 48–55.

31. Krauskopf J, de Kok TM, Hebels DG, Bergdahl IA, Johansson A, Spaeth F, et al. MicroRNA profile for health risk assessment: Environmental exposure to persistent organic pollutants strongly affects the human blood microRNA machinery. Scientific Reports. 2017; 23: 1–9.

32. Vafeiadi M, Roumeliotaki T, Chalkiadaki G, Rantakokko P, Kiviranta H, Fthenou E,. Persistent organic pollutants in early pregnancy and risk of gestational diabetes mellitus. Environment Int. 2017; 98: 89–95.

33. Porta M, Ferrer-Armengou O, Pumarega J, López T, Crous-Bou M, Alguacil J, et al. Exocrine pancreatic cancer clinical factors were related to timing of blood extraction and influenced serum concentrations of lipids. J Clin Epidemiol. 2008; 61:695–704.

34. Porta M, Fabregat X, Malats N, Guarner L, Carrato A, de Miguel A, et al. Exocrine pancreatic cancer: symptoms at presentation and their relation to tumour site and stage. Clin Transl Oncol. 2005; 7:189–197.

35. Porta M, Pumarega J, Ferrer-Armengou O, López T, Alguacil J, Malats N, et al. Timing of blood extraction in epidemiologic and proteomic studies: Results and proposals from the PANKRAS II Study. Eur J Epidemiol. 2007; 22: 577–588.

36. O'Brien KM, Upson K, Cook NR, Weinberg CR. Environmental chemicals in urine and blood: improving methods for creatinine and lipid adjustment. Environ Health Perspect. 2016; 124: 220–227.

37. Aylward LL, Green E, Porta M, et al. Population variation in biomonitoring data for persistent organic pollutants (POPs): an examination of multiple population-based datasets for application to Australian pooled biomonitoring data. Environ Int. 2014; 68: 127–138.