Environment International 192 (2024) 109056

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

Short communication

How well does a single blood sample represent long-term exposure for epidemiological studies of PFOA among men in the general population?

Scott M. Bartell [a,*], Mark P. Purdue [b], Jongeun Rhee [b], Therese H. Nøst [c,d], Jennifer Rusiecki [e], Kyle Steenland [f]

[a] Department of Environmental and Occupational Health, Department of Statistics, and Department of Epidemiology and Biostatistics, University of California, Irvine, Irvine, CA, USA
[b] Occupational and Environmental Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD, USA
[c] Department for Community Medicine, UiT The Arctic University of Norway, Tromsø, Norway
[d] HUNT Research Centre, NTNU Norwegian University of Science and Technology, Trondheim, Norway
[e] Department of Preventive Medicine and Biostatistics, Uniformed Services University of the Health Sciences, Bethesda, MD, USA
[f] Dept. of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, GA, USA

ARTICLE INFO

Editor: Adrian Covaci

Keywords:
exposure misclassification
bias
biomarker
PFAS
average exposure

ABSTRACT

Many epidemiological studies use a single blood sample per participant to assess exposure, but it is unclear how well a single sample represents longer term exposure. We performed a simulation study using summary statistics for repeated serum PFOA measurements from several previous studies in men to generate plausible serum concentrations over time, taking within-subject correlations into account. Simulated serum concentrations for controls were categorized into quintiles at each time point, and used to determine the extent of misclassification at each time point compared to the "true" long-term average exposure. We then generated case counts by quintile needed to produce an odds ratio (OR) of 1.5 for the highest vs. lowest quintile categorized based on long term exposure, and used the same misclassification rates observed in the controls to simulate misclassified exposure quintiles for cases. Comparing long term vs. single baseline exposure measures for repeated serum samples collected within about 5–13 years of each other revealed similar effect estimates, although there was a small bias to the null. Trend tests across quintiles were mostly significant using either baseline or long-term exposure. For the general population sample of men in Norway, with 5 repeated measurements over 28 years, serum PFOA was substantially lower prior to 1987, and using either of the two earliest samples as the exposure metric, compared to the long term average, produced larger bias to the null and non-significant trend tests; however using later time points as the exposure metric resulted in only a small bias. Using data based on studies of men, single baseline serum samples represented rather well the mean of repeated samples collected up to 13 years apart, but were not always reliable surrogates for average exposure over 3 decades, during which time PFOA exposure levels in the general population have changed substantially.

1. Introduction

Many epidemiologic studies of per- and polyfluoroalkyl substances (PFAS) are conducted in general populations with background levels of exposure. Such studies include ecological studies, cross-sectional studies, cohort studies, and case control studies. In cohort studies and case-control studies nested in cohorts, exposure is often ascertained by a single serum sample taken at baseline. These cohort-based study designs with baseline serum samples are particularly well suited for studying diseases with long latency, because they ensure that exposure differences preceded disease development, and that the controls arose from the same study base as the cases. Examples of such studies of potential associations of PFAS with cancer in Europe and in the United States include studies of testicular cancer (Purdue et al., 2023), kidney cancer (Shearer et al., 2021; Rhee et al., 2023), breast cancer (Mancini et al., 2020; Chang et al., 2023), prostate cancer (Rhee et al., 2023), and other cancer types (Eriksen et al., 2009).

Although perfluorooctanoate (PFOA) has a long half-life of about 2–3 years in human serum (Bartell et al. 2010, Gomis et al. 2016, Li et al. 2018), suggesting that a single serum sample might reasonably represent

* Corresponding author.
E-mail address: sbartell@uci.edu (S.M. Bartell).

https://doi.org/10.1016/j.envint.2024.109056
Received 6 June 2024; Received in revised form 22 September 2024; Accepted 3 October 2024
Available online 5 October 2024

long-term exposure for a member of the general population, background serum PFOA concentrations have declined substantially over time in recent decades in the U.S. (CDC, 2023) and in Norway (Berg et al., 2021). Such decreases have been attributed to the phase-out of PFOA and PFOS production, increasing regulation, and shift to shorter-chain PFAS and other replacements (Brennan et al., 2021).

For epidemiologists, the question arises to what extent single blood samples can represent longer term exposures, which might be more toxicologically relevant in studies of chronic disease development. Ideally one might collect repeated serum samples over time for study subjects, in order to characterize long term exposure prior to the development of disease. However, caution is needed, as the timing of disease initiation versus clinical symptoms and diagnosis is often unclear. Moreover, serum samples obtained after clinical symptoms or diagnoses may be more prone to physiological confounding or reverse causation due to disease development (Dhingra et al., 2017; Weisskopf and Webster, 2017).

There is one study of repeated serum PFAS measures over a long time period, from Norway, based on 42 men from the general population sampled in 1979, 1986, 1994, 2001 and 2007, and analyzed for 10 PFAS (Nøst et al., 2014). That study showed a markedly increasing trend for both PFOA and perfluorooctane sulfonate (PFOS) from 1979 until 2001, after which levels decreased until 2007. This trend corresponds to the increased manufacture and environmental presence world-wide of PFAS since the 1940 s, and then a decrease in "legacy" PFAS chemicals after about 2005 when concerns about toxicity of PFOA and PFOS in particular led to production decreases and more restricted use.

There are two other recent studies with repeated serum samples from subsets of participants. Rhee et al. (2023) studied 750 cases and 750 controls for aggressive prostate cancer in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO). These authors collected repeated serum samples from a subset of 60 controls at baseline from 1993 to 2001 (T0), one year later (T1), and 5 years later (T5). Purdue et al. studied 530 cases of testicular cancer and 530 controls in a case-control study of Air Force servicemen nested within the U.S. Department of Defense Serum Repository. Baseline serum samples in the Air Force study were collected between 1988 and 2017, and there were 187 controls with a second serum sample (median 4.7 years after the first serum sample).

Summary statistics for serum PFOA concentrations at each time point in each study are provided in Tables 1-3, and plotted in Fig. 1. The median participant age at the time of the first serum sample was 43 years

Table 1
Descriptive statistics for controls with repeated samples (n = 60) from Rhee et al. (2023) PLCO study of prostate cancer *.

sample	Mean	Median	Std. Dev.	Mean ln PFOA	s.d. ln PFOA
T0	3.88	3.63	1.84	1.24	0.52
T1	3.87	3.53	1.99	1.25	0.46
T5	4.69	4.53	2.43	1.42	0.51
Pearson Corr ln PFOA					
T0,T1	0.89				
T0,T5	0.68				
T1,T5	0.73				
simulated					
T0	3.97	3.47	2.21	1.24	0.52
T1	3.88	3.48	1.89	1.25	0.46
T5	4.73	4.15	2.60	1.42	0.51
Pearson Corr ln PFOA					
T0,T1	0.89				
T0,T5	0.68				
T1,T5	0.73				

* T0, T1, and T5 are samples at baseline, one year later, and five years later. Simulated data use the summary statistics for observed data for ln PFOA at five time points (mean, std.dev., Pearson correlations) to generate 1000 simulations of 650 controls for each year of sampling. Statistics for simulated data are the averages across the 1000 simulations for each sampling year.

Table 2
Descriptive statistics for controls with repeated samples (n = 84) from Purdue et al. Air Force study of testicular cancer*.

Sample 1					Sample 2					
Mean	Median	SD	Mean ln PFOA	std dev ln PFOA	Mean	Median	SD	mean ln PFOA	std dev ln PFOA	Pearson correlation Samples 1 and 2
Observed data										
6.8	6.10	3.0	1.82	0.47	5.50	5.10	2.30	1.62	0.44	0.36
Simulateddata										
6.88	6.14	3.45	1.82	0.47	5.56	5.06	2.57	1.62	0.44	0.36

* Observed data restricted to those with 2nd sample more than 4.7 years (the median for controls sampled twice) after the first one. The mean time between first and second samples for these 84 subjects was 7.8 years. Simulated data use the summary statistics for observed data for ln PFOA at five time points (mean, std.dev., Pearson correlations) to generate 1000 simulations of 650 controls for each year of sampling. Statistics for simulated data are the averages across the 1000 simulations for each sampling year.

Table 3a
Descriptive statistics for participants with repeated samples (n = 42) from Nøst et al. study of Norwegian men*.

Observed data					
year	**mean**	**sd**	**median**	**mn of log PFOA**	**s.d. log PFOA**
1979	0.97	0.42	0.86	−0.11	0.39
1986	2.52	0.68	2.51	0.89	0.28
1994	4.42	1.64	3.93	1.42	0.36
2001	4.3	1.78	4.02	1.38	0.39
2007	3.26	1.24	3.25	1.12	0.37
Simulated data					
1979	0.97	0.39	0.90	−0.11	0.39
1986	2.52	0.71	2.43	0.89	0.28
1994	4.42	1.63	4.15	1.42	0.36
2001	4.30	1.75	3.98	1.38	0.39
2007	3.27	1.25	3.05	1.11	0.37

* observed data are for 42 Norwegian men (controls) sample repeatedly in 1979, 1986, 1994, 2001, and 2007. Simulated data use the summary statistics for observed data for ln PFOA at five time points (mean, std.dev., Pearson correlations) to generate 1000 simulations of 650 controls for each year of sampling. Statistics for simulated data are the averages across the 1000 simulations for each sampling year.

Table 3b
Pearson correlations for pairs of sampling years in Nøst et al. study of Norwegian men*.

	year				
Observed data	1979	1986	1994	2001	2007
1979	1	0.52	0.33	0.32	0.38
1986		1	0.70	0.51	0.46
1994			1	0.72	0.67
2001				1	0.83
2007					1
Simulated data	1979	1986	1994	2001	2007
1979	1	0.52	0.33	0.32	0.38
1986		1	0.70	0.51	0.46
1994			1	0.72	0.67
2001				1	0.83
2007					1

* Statistics for simulated data are the averages across the 1000 simulations for each sampling year.

for the Norwegian study (range 29–54), 62 years in the PLCO study (range 55–74), and between 25 and 29 years in the Air Force study (ranging up to 39), for which ages were obtained in broad categories only to preserve anonymity. For the Air Force study, approximately 78 % of participants were Non-Hispanic White, 14 % were Hispanic, 2 % were Non-Hispanic Black, 2 % were Asian, and 4 % were in other race/ethnicity categories. For the PLCO study, approximately 87 % of participants were Non-Hispanic White, 7 % were Non-Hispanic Black, 5 % were Asian, and 2 % were Hispanic. Race/ethnicity was not obtained in the Norwegian study, but all participants were born in Norway and resided in Tromsø, North Norway throughout the entire study period.

Based on summary statistics for serum PFOA concentrations among men provided by the authors of all three studies, including within-subject correlations over time, we perform simulations investigating potential bias in epidemiological effect estimates from exposure misclassification using single serum samples instead of long-term exposure, measured by individual-specific mean serum PFOA concentration across sampling time points. This corresponds to the common practice of using a single serum sample per participant to characterize exposures in epidemiological studies, e.g., the case-control studies of Rhee et al. 2023 (PLCO) and Purdue et al. 2023 (Air Force).

Existing studies with repeated serum measurements typically either do not include the health outcomes under study, or only obtain repeated serum measurements for a small subset of participants. Our approach uses existing repeated measured serum data to inform simulations of epidemiology studies where both repeated serum measurements and health outcomes are collected. By leveraging the correlation structures observed using repeated measurements in a smaller number of participants, we were able to design realistic simulations examining the likely impacts of exposure misclassification in larger epidemiologic studies.

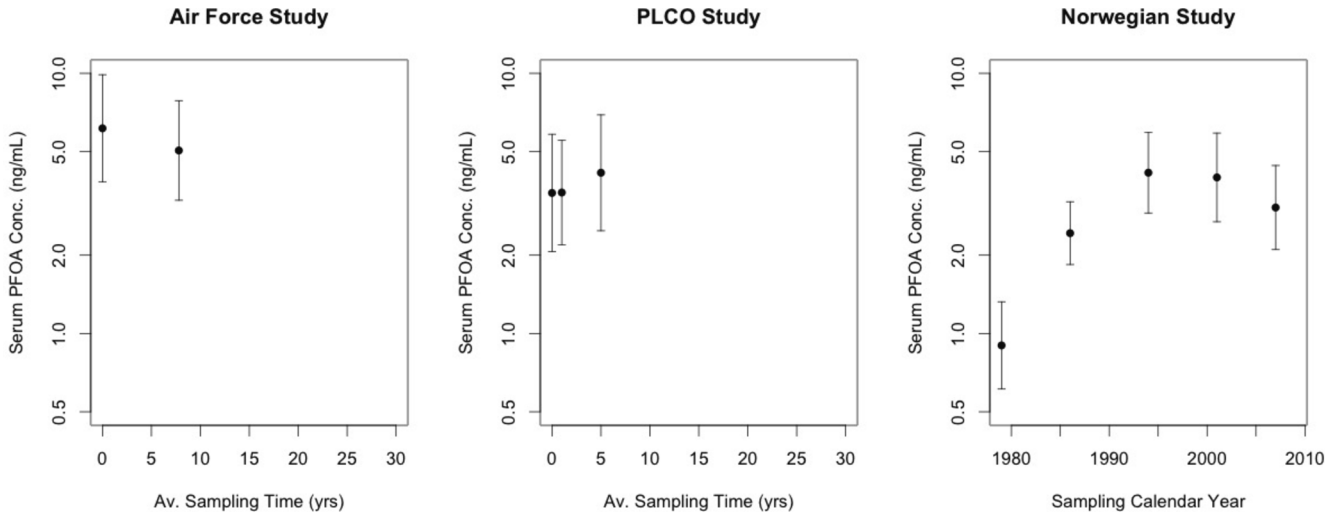

Fig. 1. Temporal patterns in serum PFOA concentrations for the 3 studies (PLCO, Air Force, and Norwegian). Geometric mean serum PFOA concentration (points) with 1 geometric standard deviation (error bars) are shown for each time point.

2. Methods

Serum PFOA measurements were reported at 3 time points for the 60 controls in the PLCO data, at 2 time points for the Air Force data, and at 5 time points for the Norwegian population-based data from the Tromsø Study. All three data sets were limited to adult males. Summary statistics were based on all 60 controls for the PLCO data, 84 controls from the Air Force data for whom there was > the median 4.7 years between samples, and 42 subjects who were sampled at all five time points from the Norwegian data.

We used reported summary statistics for the controls in each case-control study and for the 42 subjects with measurements at all five time points in the Norwegian study to generate serum PFOA data for a hypothetical nested case-control study of a chronic disease such as cancer, with similar properties as the original data sets. We assumed that long term average exposure is the best metric for chronic disease development for the purposes of this exercise. This is a common assumption for many studies in environmental epidemiology, including air pollution and cardiopulmonary disease, smoking and cancer, PFAS and cancer, and a variety of occupational illnesses (Bonnie et al., 2015; Pope et al., 2002; Bartell and Vieira, 2021).

We assumed that the distribution of serum PFOA concentrations at each time point was lognormal. We first obtained the mean m_i and standard deviation s_i of the natural log serum PFOA concentration at each time point i, and the Pearson correlations r_{ij} across all pairs of time points, $i \neq j$. We then constructed the covariance matrix for the log serum PFOA concentrations for each study, with elements ($r_{ij}s_is_j$), and generated multivariate normal random variates according to the means and covariance matrix for each time point in each study using the *mvtnorm* library in R (Genz and Bretz, 2009). We simulated 1000 data sets with 650 controls for each study at each time point, generating serum PFOA measurements at each time point according to the summary statistics for each study.

We confirmed that the simulated data for controls conformed to the means and standard deviations of log PFOA at each time point, as well as the Pearson correlations between pairs of time points as reported by the authors.

For each of our 1000 simulated data sets, we computed long-term exposures as the mean of simulated serum PFOA concentrations over time for each of our 650 controls (across 3 samples for PLCO, across 2 samples for the Air Force data, and across 5 samples for the Norwegian data). We then categorized these single and long-term average serum PFOA concentrations into quintiles for each simulated data set. We calculated discrepancies in quintiles based on single measurements and long-term exposure, considering the quintiles based on single serum measurements at the first time point as 'misclassified' when they did not match the quintiles based on long-term exposure.

We generated hypothesized case counts in each of the long-term exposure quintiles for the three studies (650 cases in each simulated data set, with 1000 simulated data sets), such that there was a resulting positive monotonic trend across increasing quintiles, with quintiles having odds ratios of 1.000, 1.125, 1.250, 1.375, and 1.500. This trend across quintiles was chosen to represent a realistic moderate and positive trend, compared to odds ratios reported in case-control analyses. The case counts yielding any particular set of odds ratios was determined by solving algebraically the system of equations for the usual Wald odds ratio point estimates:

$$n_{21} = \frac{n_{2\bullet}}{\sum_{i=1}^{5} \frac{n_{1i}}{n_{11}} OR_i}$$

and

$$n_{2i} = \frac{n_{21} n_{1i}}{n_{11}} OR_i$$

where $n_{2\bullet}$ is the total number of cases, set to 650 in our simulations, n_{1i} is the number of controls in exposure quintile i (within quintile 1 as the referent), n_{2i} is the number of cases in exposure quintile i, and OR_i is the odds ratio for exposure quintile i versus quintile 1, for $i \in \{1, 2, 3, 4, 5\}$. We rounded the results of each these calculations to the nearest whole number, to ensure simulated case counts were realistic.

We then in turn created a group of hypothetical 'misclassified' cases for each study, following the same pattern of misclassification observed in the controls using first samples only (i.e., non-differential misclassification with regard to the disease outcome). This was accomplished by computing a 5x5 matrix of misclassification probabilities from the control data, reflecting the proportion of controls in each long-term exposure quintile classified correctly or incorrectly into each of the 5 possible quintiles based on their exposure quintile for the first time point only, and moving those same proportions of cases into the appropriate correct and incorrect quintiles.

In order to assess the impact of potential exposure misclassification on epidemiological effect estimates when using only the one serum sample per participant, we analyzed the exposure-outcome associations across the quintile categories of exposure for both the long-term (mean exposure across all time points), and misclassified baseline exposure data (using a single serum sample only) in each of our 1000 simulated data sets. For the Norwegian data, which spanned about 3 decades, we analyzed the exposure-outcome associations using serum concentrations at each of the 5 time points alone vs. mean serum concentrations across all 5 time points.

Chi-square trend tests for a trend in odds ratios across the 5 exposure quintiles were calculated for long term and single sample data using the *prop.trend.test* function in R, with the natural log of the mid-point of each exposure category (determined from the quintiles of long-term exposure) as scores. We generated histograms of 1) the p-values for the tests of trend in the misclassified data, and 2) the misclassified odds ratios for quintile 5 vs quintile 1 (the referent).

3. Results

Tables 1-3 present means, medians, and standard deviations, and within-subject correlations at each time point in the original data for comparisons with the analogous data for our 650 simulated controls for each study (using average means, standard deviations, and correlation across 1000 simulations). In general, the means and standard deviations were very similar between observed and simulated data, especially for the log transformed data which we used as the basis for our simulations of associations with health outcomes. The Pearson correlations were also similar in the observed and simulated data.

Figs. 2-6 display the distribution of p-values of tests for trend using the misclassified data, i.e., the first sample for PLCO and the Air Force, and the first, second, and third sampling times for the Norwegian data, across 1000 simulations using 650 cases and 650 controls. We also evaluated the fourth and the fifth sample vs. the average of all five for the Norwegian data, but results are not presented as they were similar to the results using the third sample. Also presented in Figs. 2-6 are the distributions of odds ratios for the fifth exposure quintile vs. the first exposure quintile across the 1000 simulations, using the misclassified data. Median values for these trend test p-values and odds ratios are provided in Table 4.

The PLCO and Air Force simulations based on the misclassified data yielded mostly significant tests for trends, and ORs for the fifth exposure quintile compared to the first quintile were only slightly attenuated from the true OR of 1.50. For the Norwegian data, using the first and second sampling times resulted in p-values not significant at the 0.05 level, and ORs for the fifth quintile that were considerably lower than the true OR of 1.50. It was only for the third time point (1994) or later that the p-values and ORs approximated what is expected based on the long-term average exposure.

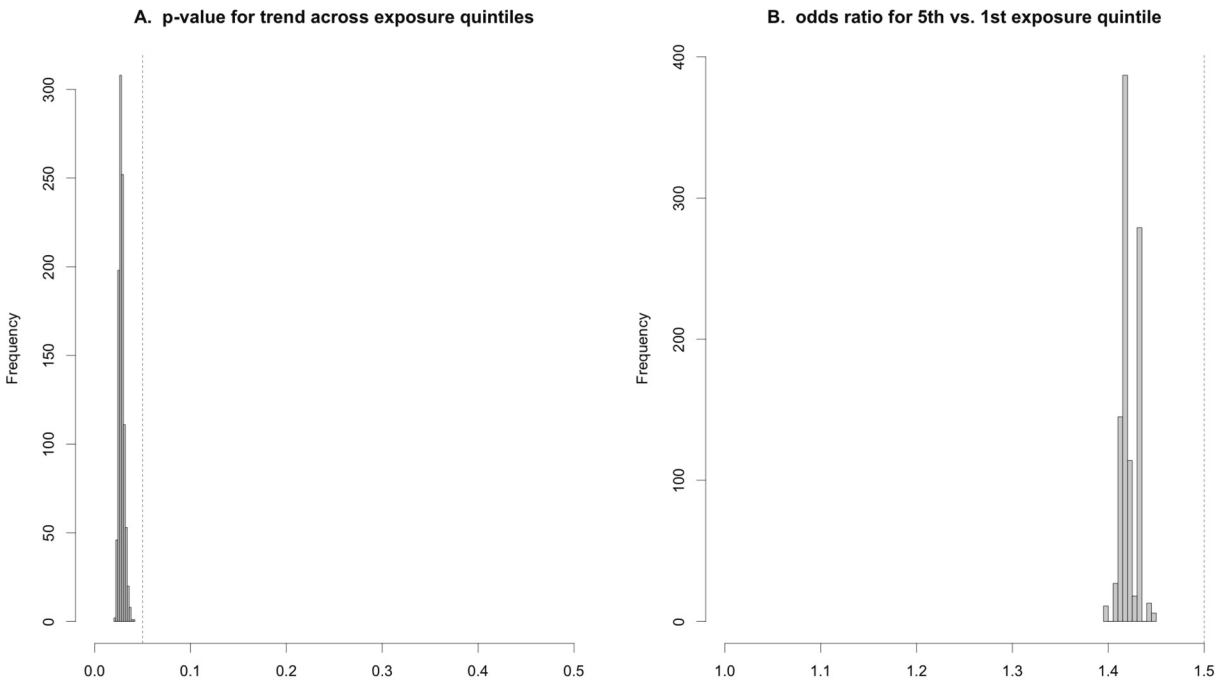

Fig. 2. Distribution of A.) p-values for trend test across quintiles and B.) odds ratios for 5th vs. 1st quintile, for the PLCO data, using misclassified (first sample only) data, across 1000 simulations, 650 cases and controls. p = 0.05 and OR = 1.5 are marked by dashed red lines.

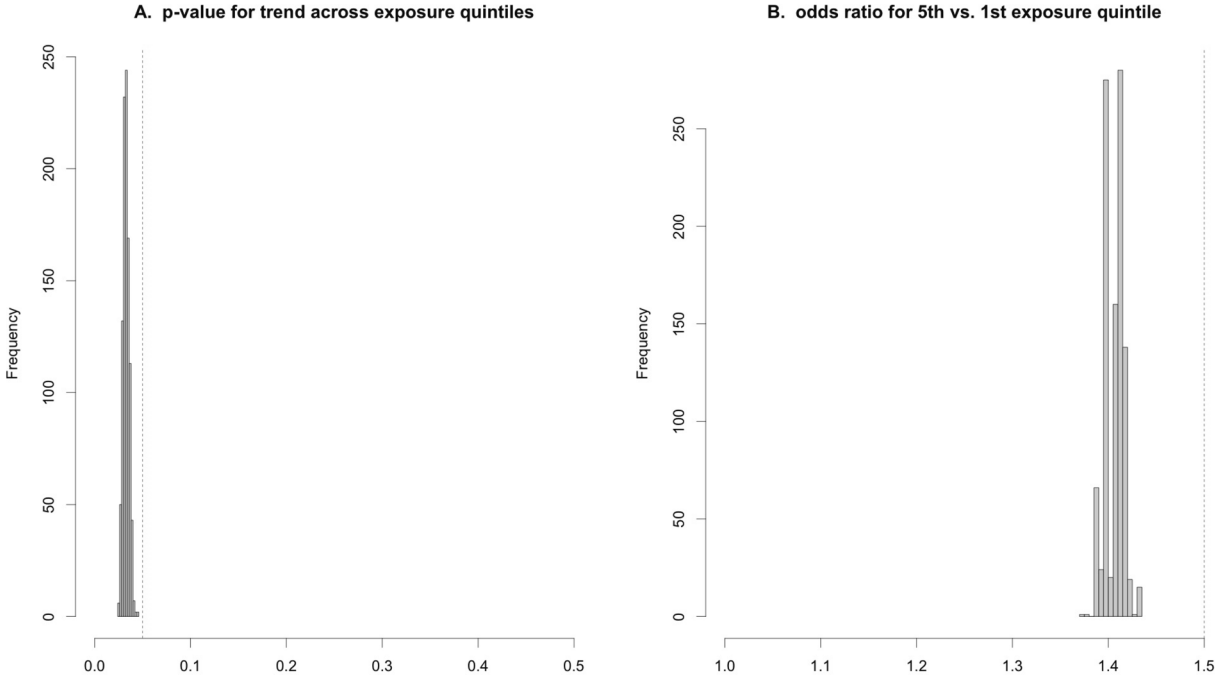

Fig. 3. Distribution of A.) p-values for trend test across quintiles and B.) odds ratios for 5th vs. 1st quintile, for the Air Force data, using misclassified (first sample only) data, across 1000 simulations, 650 cases and controls. p = 0.05 and OR = 1.5 are marked by dashed red lines.

4. Discussion

We observed little difference in exposure–response trends when using a single first serum sample vs. using the mean of repeated serum samples across periods of a decade or less for the data on men from PLCO and the Air Force study. In these two data sets, there was a relatively small bias to the null when using the single first serum sample, considered 'misclassified', vs. using the average across all samples for each participant. A bias to the null is expected when the true exposure response is positive and monotonic, and exposure misclassification is non-differential and non-extreme (Weinberg et al., 1994; Dosemeci et al., 1990). We have used PFOA in the simulations presented here, but expect our results could be generalizable to other legacy PFAS with relatively long half-lives.

The bias in the results for the Norwegian study were dependent on which sampling round was used for the single time point, perhaps because of the relatively rapid increase in serum PFOA concentrations during the early rounds of sampling. The bias to the null using on the

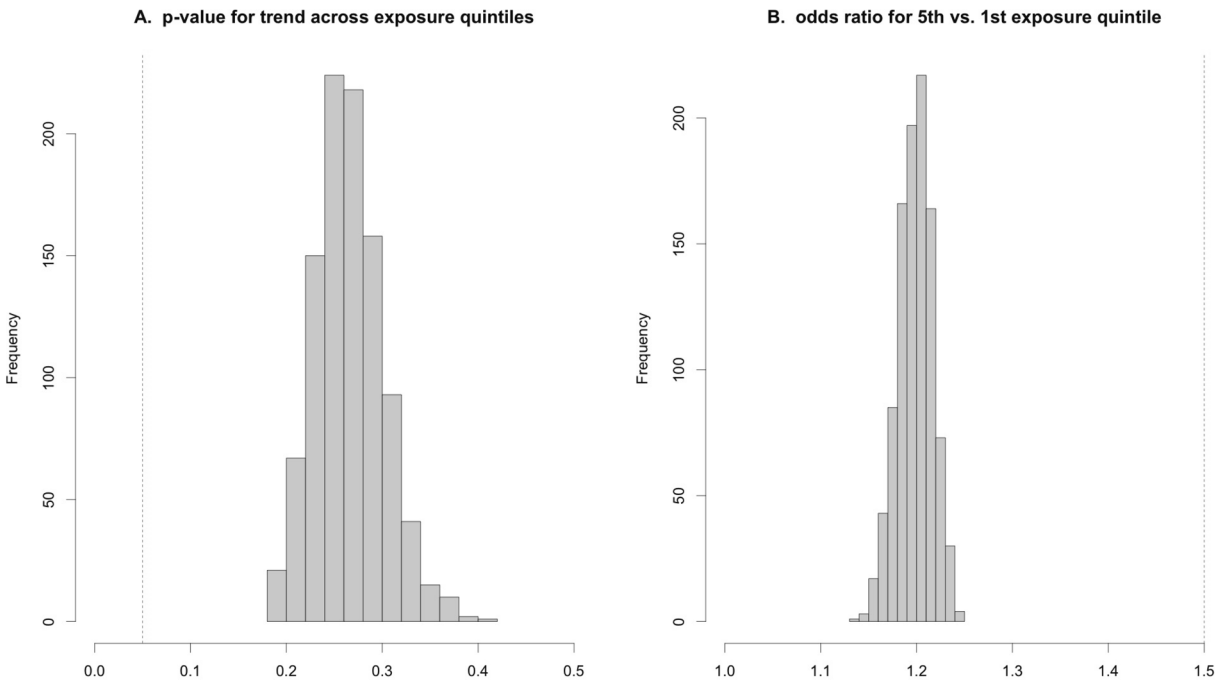

Fig. 4. Distribution of A.) p-values for trend test across quintiles and B.) odds ratios for 5th vs. 1st quintile, for the Norwegian data, using misclassified (first sample, 1979) data, across 1000 simulations, 650 cases and controls. p = 0.05 and OR = 1.5 are marked by dashed red lines.

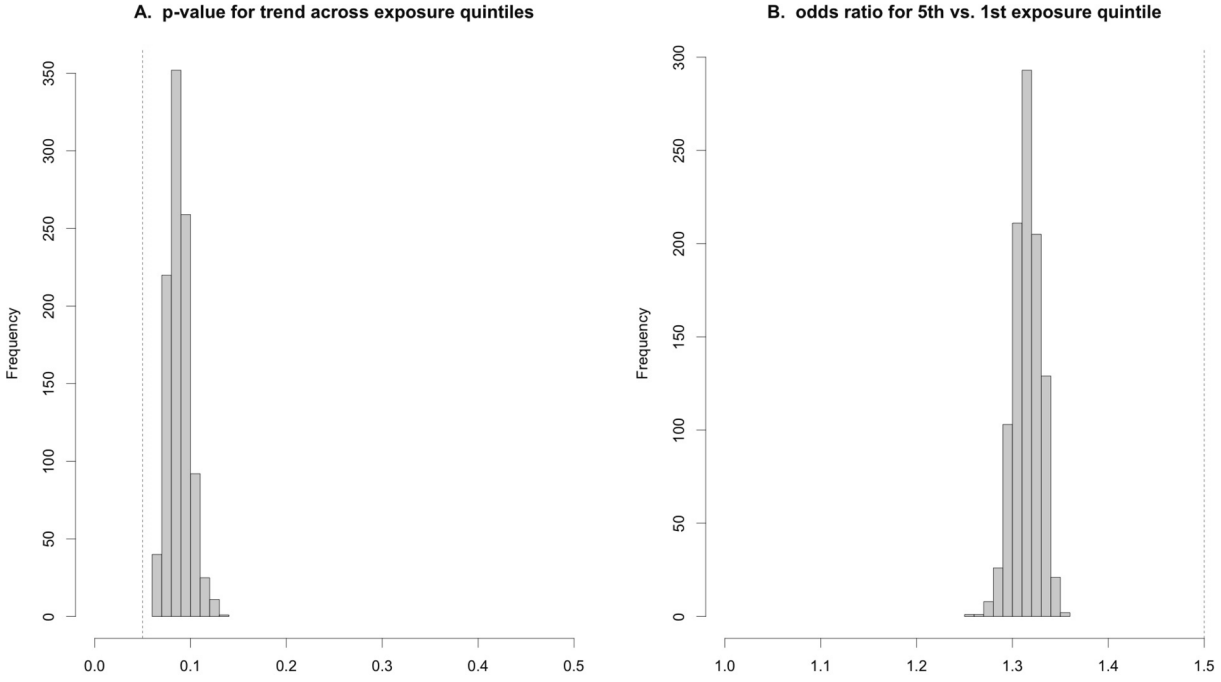

Fig. 5. Distribution of A.) p-values for trend test across quintiles and B.) odds ratios for 5th vs. 1st quintile, for the Norwegian data, using misclassified (second sample, 1986) data, across 1000 simulations, 650 cases and controls. p = 0.05 and OR = 1.5 are marked by dashed red lines.

first (1979) or second (1986) sample vs. the long term average in the Norwegian study was considerable, while the bias to the null based on the third (1994) or later sample vs. the long term average was minimal. PFOA concentrations were very low for the earliest sample in 1979, and also relatively low in 1986, while from 1994 to 2007 they were more typical of concentrations in many countries in those years, after typical human serum concentrations had peaked at around 3–4 ng/ml. Thus, the bias was less in the latter 13 years of the sampling period, during which a single sample approximated longer term mean exposure fairly well. Up until the 1990 s, concentrations were markedly increasing over time, and perhaps reflected different exposure patterns or sources than in later decades. This pattern of a relatively rapid increase in serum PFOA concentrations during the earlier time period followed by declining exposure, combined with a much longer time span and a greater number of repeated samples, is much different from the context of the serum samples in the PLCO and Air Force studies, which covered average time spans of about 5 years.

Although serum concentrations may change over time for the same

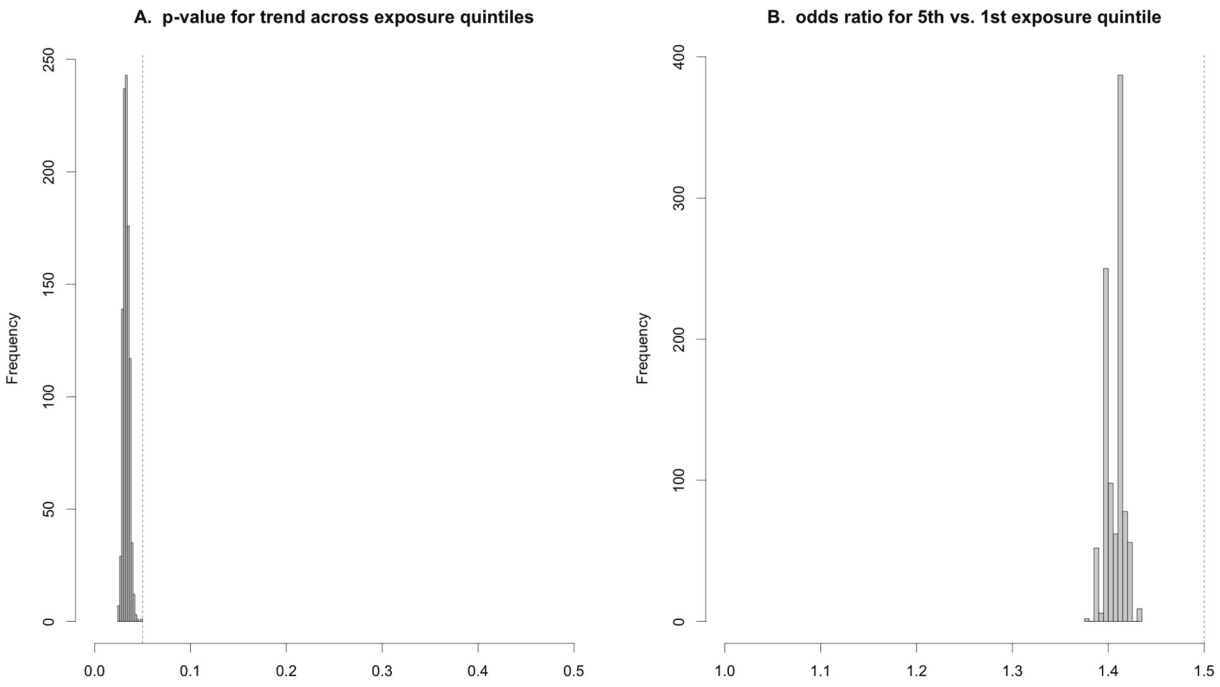

Fig. 6. Distribution of A.) p-values for trend test across quintiles and B.) odds ratios for 5th vs. 1st quintile, for the Norwegian data, using misclassified (third sample, 1994) data, across 1000 simulations, 650 cases and controls. p = 0.05 and OR = 1.5 are marked by dashed red lines.

Table 4
Median p-values for trend across exposure quintiles and median odds ratios for 5th vs. 1st quintile across 1000 simulations, for 650 cases and controls.

Study and time point	Median p-value for trend	Median odds ratio, 5th vs. 1st quintile
PLCO, 1st time point	0.028	1.42
Air Force, 1st time point	0.039	1.39
Norweigan, 1st time point	0.26	1.20
Norweigan, 2nd time point	0.087	1.32
Norweigan, 3rd time point	0.033	1.41

people, the relative exposure ranking among participants in a cohort could be rather stable, and because odds ratio estimates and trend tests are based on these relative rankings, they are not necessarily affected as much by absolute changes. Given that the serum concentrations for men are moderately or highly correlated over time, suggesting a relatively stable exposure ranking, tests for trend of disease risk within such a population which relied on only one sample, might not differ markedly from tests for trend using an estimate of average exposure over time, i.e., the average of repeated samples across many years. However, as noted, we found this to be likely only for a period of about 13 years or less; with longer time periods between serum samples in years covering different time trends in Norway, there were lower correlations over time, and biases to the null were strongest when using single serum samples furthest back in time when human exposures were substantially lower.

Although we arbitrarily chose the sample size to be 650 each for cases and controls in this simulation, similar to the sample sizes for the two motivating case-control studies (750 cases in PLCO, and 530 cases in the Air Force study), it should be noted that these sample sizes are relatively large compared to many case-control studies in other settings. The p-values for the trend tests depend heavily on sample size; we expect that those p-values would be higher for in simulations with smaller samples sizes that are otherwise similar to our setting.

Our findings are reassuring as to the validity of exposure classification in cohort studies of men in general populations with relatively stable or only slowly changing exposures, in which only one baseline sample of PFOA (or likely other legacy PFASs with long half-lives such as PFOS and perfluorononanoate) is available to characterize exposure.

Other researchers have also found that non-differential errors in estimating PFOA exposure have little effect on exposure–response trends in an area with very high exposures, when the relative ranks of different individuals did not markedly change (Avanasi et al., 2016). Our simulations here suggest that the same might also hold when exposures are more typical of the general population and reasonably stable over time. Intuitively, the possibility of changing rankings when exposure contrasts are small and environmental exposures are rapidly increasing would seem much greater than when exposure contrasts are large and exposures are much higher. In this sense, our simulation findings that people retain their relative ranking in a low exposure setting, may be counter-intuitive. However, the within-subject correlations between time points was moderate or large (ranging from 0.32 to 0.89) in all three studies used as the basis for our simulations, suggesting that a single serum sample can be informative regarding longer term exposure ranking within a cohort. Such relative rankings might remain relatively constant in the case of the legacy PFAS, such as PFOA, because they have relatively long half-lives, and their local sources in the environment (e.g., drinking water, consumer products, and diet) may only change slowly over time. However, few longitudinal data sets with repeated serum measurements are available for PFAS and many other persistent organic pollutants. Such studies would allow researchers to assess the extent of within-person correlation in serum concentrations and its potential impacts on epidemiological findings.

Strengths of our study include the use of empirical data from the only three general population studies we are aware of with repeated sampling of serum PFOA among the same participants over many years. Limitations include the use of only men in the study populations; no data were available for women or children. Breastfeeding and menstruation, for example, have distinct effects on PFAS excretion rates (Verner et al., 2016; Li et al., 2018), which could affect the reliability of single blood samples for representing long-term average exposure. Children have more frequent and rapid changes in diet and physiology, particularly

during infancy and puberty, resulting in less stable serum PFAS concentrations (Shin et al., 2011; Verner et al., 2016) so it is less likely that single serum samples would reflect their long-term PFOA exposure than it is for adults. The other major limitation is that, with the exception of the Norwegian study, the other two studies we used had repeated sampling over a relatively short period (about 15 years or less), during time periods in which PFOA serum concentrations were only changing slowly.

In conclusion, for men, single baseline serum samples represented rather well the mean of repeated samples collected up to 13 years apart, but were not always reliable surrogates for epidemiologic analyses using average exposure over 3 decades, during which time PFOA exposure levels in the general population have changed more substantially. More research is needed to evaluate the reliability of single blood sample for representing long-term exposure for epidemiological studies of PFOA among women and children, which were not represented in our analysis. Future research could also compare single blood samples to alternative exposure metrics such as cumulative or peak exposure, rather than average exposure.

CRediT authorship contribution statement

Scott M. Bartell: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Formal analysis. **Mark P. Purdue:** Writing – review & editing, Formal analysis. **Jongeun Rhee:** Writing – review & editing, Formal analysis. **Therese H. Nøst:** Writing – review & editing, Formal analysis. **Jennifer Rusiecki:** Writing – review & editing, Formal analysis. **Kyle Steenland:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Formal analysis.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Scott M. Bartell has served as a compensated expert witness for PFOA medical monitoring lawsuits in New Hampshire. The terms of this arrangement were reviewed and approved by the University of California, Irvine in accordance with its conflict-of-interest policies. The other authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The prostate cancer study was supported by the the Intramural Research Program of the National Cancer Institute (NCI), National Institutes of Health (NIH). The testicular cancer study was supported by the Office of the Assistant Secretary of Defense for Health Affairs through the Peer-Reviewed Cancer Research Program (W81XWH-19-1-1044) and the Intramural Research Program of the NCI, NIH. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not to be construed as official or necessarily endorsed by the Department of Defense, Uniformed Services University of the Health Sciences, or the NIH. The use of trade names is for identification only and does not imply endorsement by the Department of Health and Human Services or the Department of Defense.

References

Avanasi R, Shin HM, Vieira VM, Savitz DA, Bartell SM. Impact of Exposure Uncertainty on the Association between Perfluorooctanoate and Preeclampsia in the C8 Health Project Population. Environ Health Perspect. 2016 Jan;124(1):126-32. doi: 10.1289/ehp.1409044. Epub 2015 Jun 19. PMID: 26090912; PMCID: PMC4710599.

Bartell, S.M., Calafat, A.M., Lyu, C., Kato, K., Ryan, P.B., Steenland, K., 2010. Rate of decline in serum PFOA concentrations after granular activated carbon filtration at two public water systems in Ohio and West Virginia. Environ. Health Persp. 118 (2), 222–228.

Bartell, S.M., Vieira, V.M., 2021. Critical review on PFOA, kidney cancer, and testicular cancer. J. Air Waste Manag. Assoc. 71 (6), 663–679. https://doi.org/10.1080/10962247.2021.1909668.

Berg, V., Sandanger, T.M., Hanssen, L., et al., 2021. Time trends of perfluoroalkyl substances in blood in 30-year old Norwegian men and women in the period 1986–2007. Environ. Sci. Pollut. Res. 28, 43897–43907. https://doi.org/10.1007/s11356-021-13809-6.

Bonnie RJ, Stratton K, Kwan LY, editors. Public Health Implications of Raising the Minimum Age for Purchasing Tobacco Products. Board on Population Health and Public Health Practice; Institute of Medicine; Washington (DC): National Academies Press (US); 2015 Jul 23. Available from: https://www.ncbi.nlm.nih.gov/books/NBK310413/.

Brennan, N.M., Evans, A.T., Fritz, M.K., Peak, S.A., von Holst, H.E., 2021 Oct 17. Trends in the regulation of per- and polyfluoroalkyl substances (PFAS): A scoping review. Int. J. Environ. Res. Public Health. 18 (20), 10900. https://doi.org/10.3390/ijerph182010900.

Centers for Disease Control and Prevention (CDC), 2023. National Report on Human Exposure to Environmental Chemicals, Biomonitoring Data Tables for Environmental Chemicals. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention.

Chang, V.C., Rhee, J., Berndt, S.I., Moore, S.C., Freedman, N.D., Jones, R.R., et al., 2023. Serum perfluorooctane sulfonate and perfluorooctanoate and risk of postmenopausal breast cancer according to hormone receptor status: An analysis in the prostate, lung, colorectal and ovarian cancer screening trial. Int. J. Cancer. 153 (4), 775–782. https://doi.org/10.1002/ijc.34487. PMID:36843273.

Dhingra, R., Winquist, A., Darrow, L.A., Klein, M., Steenland, K., 2017. A study of reverse causation: Examining the associations of perfluorooctanoic acid serum levels with two outcomes. Environ. Health Perspect. 125 (3), 416–421. https://doi.org/10.1289/EHP273.

Dosemeci, M., Wacholder, S., Lubin, J.H., 1990. Does nondifferential misclassification of exposure always bias a true effect toward the null value? Am. J. Epidemiol. 132 (4), 746–748. https://doi.org/10.1093/oxfordjournals.aje.a115716.

Eriksen, K.T., Sørensen, M., McLaughlin, J.K., Lipworth, L., Tjønneland, A., Overvad, K., et al., 2009. Perfluorooctanoate and perfluorooctanesulfonate plasma levels and risk of cancer in the general Danish population. J. Natl. Cancer Inst. 101 (8), 605–609. https://doi.org/10.1093/jnci/djp041. PMID:19351918.

Genz A, Bretz F (2009). *Computation of Multivariate Normal and t Probabilities*, series Lecture Notes in Statistics. Springer-Verlag, Heidelberg. ISBN 978-3-642-01688-2.

Gomis, M.I., Vestergren, R., Nilsson, H., Il, C., 2016. Contribution of direct and indirect exposure to human serum concentrations of perfluorooctanoic acid in an occupationally exposed group of ski waxers. Environ. Sci. Technol. 50, 7037–7046.

Jongeun Rhee, Vicky C. Chang, Iona Cheng, Antonia M. Calafat, Julianne Cook, Botelho3, Joseph J. Shearer, Joshua N. Sampson5, Veronica Wendy Setiawan, Lynne R. Wilkens, Debra T. Silverman1, Mark P. Purdue, Jonathan N. Hofmann, Serum concentrations of per- and polyfluoroalkyl substances and risk of renal cell. Env Int 2023.

Li, Y., Fletcher, T., Mucs, D., et al., 2018. Half-lives of PFOS, PFHxS and PFOA after end of exposure to contaminated drinking water. Occup. Environ. Med. 75, 46–51.

Mancini, F.R., Cano-Sancho, G., Gambaretti, J., Marchand, P., Boutron-Ruault, M.C., Severi, G., et al., 2020. Perfluorinated alkylated substances serum concentration and breast cancer risk: Evidence from a nested case-control study in the French E3N cohort. Int. J. Cancer. 146 (4), 917–928.

Nøst, T.H., Vestergren, R., Berg, V., Nieboer, E., Odland, J.Ø., Sandanger, T.M., 2014. Repeated measurements of per- and polyfluoroalkyl substances (PFASs) from 1979 to 2007 in males from Northern Norway: assessing time trends, compound correlations and relations to age/birth cohort. Environ. Int. 67, 43–53. https://doi.org/10.1016/j.envint.2014.02.011. Epub 2014 Mar 20 PMID: 24657493.

Pope III, C.A., Burnett, R.T., Thun, M.J., et al., 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. JAMA 287 (9), 1132–1141. https://doi.org/10.1001/jama.287.9.1132.

Purdue, M.P., Rhee, J., Denic-Roberts, H., McGlynn, K.A., Byrne, C., Sampson, J., et al., 2023. A nested case-control study of serum per- and polyfluoroalkyl substances and testicular germ cell tumors among US Air Force servicemen. Environ. Health Perspect. 131 (7), 77007. PMID:37458713.

Rhee, J., Barry, K.H., Huang, W.Y., Sampson, J.N., Hofmann, J.N., Silverman, D.T., et al., 2023. A prospective nested case-control study of serum concentrations of per- and polyfluoroalkyl substances and aggressive prostate cancer risk. Environ. Res. 228, 115718.

Shearer, J.J., Callahan, C.L., Calafat, A.M., Huang, W.Y., Jones, R.R., Sabbisetti, V.S., et al., 2021. Serum concentrations of per- and polyfluoroalkyl substances and risk of renal cell carcinoma. J. Natl. Cancer Inst. 113 (5), 580–587. https://doi.org/10.1093/jnci/djaa143. PMID:32944748.

Shin, H.M., Vieira, V.M., Ryan, P.B., Steenland, K., Bartell, S.M., 2011. Retrospective exposure estimation and predicted versus observed serum perfluorooctanoic acid

concentrations for participants in the C8 Health Project. Environ. Health Perspect. 119 (12), 1760–1765. https://doi.org/10.1289/ehp.1103729.

Weinberg, C.R., Umbach, D.M., Greenland, S., 1994. When will nondifferential misclassification of an exposure preserve the direction of a trend? Am. J. Epidemiol. 140 (6), 565–571. https://doi.org/10.1093/oxfordjournals.aje.a117283. PMID: 8067350.

Weisskopf MG, Webster TF. Trade-offs of Personal Versus More Proxy Exposure Measures in Environmental Epidemiology. Epidemiology. 2017 Sep;28(5):635-643. doi: 10.1097/EDE.0000000000000686.Verner MA, Ngueta G, Jensen ET, Fromme H, Völkel W, Nygaard UC, Granum B, Longnecker MP. A Simple Pharmacokinetic Model of Prenatal and Postnatal Exposure to Perfluoroalkyl Substances (PFASs). Environ Sci Technol. 2016 Jan 19;50(2):978-86. doi: 10.1021/acs.est.5b04399.