ONLINE SUPPLEMENTARY MATERIAL

Phthalates, Perfluoroalkyl Acids, Metals and Organochlorines and Reproductive Function: A Multi-Pollutant Assessment in Greenlandic, Polish and Ukrainian Men

Virissa Lenters[1], Lützen Portengen[1], Lidwien AM Smit[1], Bo AG Jönsson[2], Aleksander Giwercman[3], Lars Rylander[2], Christian H Lindh[2], Marcello Spanò[4], Henning S Pedersen[5], Jan K Ludwicki[6], Lyubov Chumak[7], Aldert H Piersma[8], Gunnar Toft[9], Jens Peter Bonde[10], Dick Heederik[1], Roel Vermeulen[1]

[1] Division of Environmental Epidemiology, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands
[2] Division of Occupational and Environmental Medicine, Lund University, Lund, Sweden
[3] Reproductive Medicine Centre, Skane University Hospital Malmö, Lund University, Malmö, Sweden
[4] Laboratory of Toxicology, Unit of Radiation Biology and Human Health, ENEA Casaccia Research Center, Rome, Italy
[5] Centre for Arctic Environmental Medicine, Nuuk, Greenland
[6] Department of Toxicology and Risk Assessment, National Institute of Public Health-National Institute of Hygiene, Warsaw, Poland
[7] Social Medicine and Organization of Public Health, Kharkiv National Medical University, Kharkiv, Ukraine
[8] Laboratory for Health Protection Research, National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[9] Danish Ramazzini Center, Department of Occupational Medicine, Aarhus University Hospital, Aarhus, Denmark
[10] Department of Occupational and Environmental Medicine, Bispebjerg Hospital, University of Copenhagen, Copenhagen, Denmark

Table of Contents

Supplementary Methods

Study populations

The present study is based on a previously established cohort, INUENDO (Toft et al. 2005a), and includes 602 men from Greenland, Warsaw (Poland) and Kharkiv (Ukraine) who filled in a questionnaire on lifestyle factors, provided a semen sample, and provided a blood sample. For the baseline study, 598 (90% participation rate), 472 (68%) and 640 (26%) expectant couples (pregnant women and their male partners) were enrolled from the eligible target populations in Greenland, Warsaw, and Kharkiv, respectively. To be eligible, it was required that both partners were ≥18 years of age and born in the country of study. The baseline study also included Swedish fisherman, whose partners were not necessarily pregnant; this sub-cohort was excluded from the present analysis to achieve a more uniform study population. The age distribution and number of children did not differ between participants, non-respondents and those who declined participation from Greenland and Kharkiv. A non-response analysis was not possible for the Polish subcohort as no data were available for those who did not explicitly accept or decline participation (Toft et al. 2005a). Of the eligible men, 201 (79% participation rate) from Greenland, 198 (29%) from Warsaw, and 208 (33%) from Kharkiv provided a semen sample.

Collection of blood samples

Blood samples were drawn from a cubital vein into 10 mL EDTA-containing vacuum tubes for collection without additives (Becton Dickinson, Meylan, France). The blood sample was collected on the same day as the semen sample for >97% of Polish and Ukrainian men, and within 3 days for the remaining men. For Greenlanders, 41% of samples were collected within 3 days, and for the rest, within a year (median 18 weeks, IQR 23–44). We did not collect samples in trace metal-free tubes, and therefore cannot exclude that there may have been some contamination in the analysis of metals. After cooling to room temperature the tubes were centrifuged at 4000 g for 15 min. Serum was transferred with ethanol rinsed Pasteur pipettes to ethanol rinsed brown glass bottles (Termometerfabriken, Gothenburgh, Sweden). A piece of aluminum foil was placed on top of the bottles which were then sealed. Samples were stored at -20°C until shipment, but it was accepted to keep it in refrigerator for up to four days (as originally described in Jönsson et al. (2005)). Samples were transported on dry ice to the Department of Occupational and Environmental Medicine, Lund University, Sweden, where all chemical analyses were performed. Samples were stored at -80°C until later analysis.

Exposure assessment

PCB-153 and *p,p´*-DDE were analyzed as previously described (Jönsson et al. 2005). Additional analytes (phthalates, metals, perfluoroalkyl acids, and hexachlorobenzene) were more recently analyzed. Perfluoroalkyl acids were analyzed (Lindh et al. 2012) along with phthalates using a triple quadrupole linear ion trap mass spectrometer equipped with a TurboIonSpray source (QTRAP 5500; AB Sciex, Foster City, CA, USA), coupled to a liquid chromatography system (UFLCXR, Shimadzu Corporation, Kyoto, Japan; LC/MS/MS). Aliquots of 100 µL serum were added with ^{2}H- ^{13}C- or ^{18}O-

labeled internal standards for all evaluated compounds. The samples were digested with glucoronidase and the proteins were precipitated with acetonitrile.

Only oxidized metabolites were analyzed. Serum has lipase activity, and if the monoesters should be analyzed it is necessary to deactivate the lipases with e.g. phospheric acid immediately at sampling collection to avoid contamination from phthalate diesters in the environment (Frederiksen et al. 2010; Högberg et al. 2008). While oxy-functional group metabolites were detected in only 40–50% of samples in our study, all phthalates metabolites were measured with relatively high precision; coefficients of variation between 7% and 19% were achieved.

For all analytes, the limits of detection (LOD) were determined as the concentrations corresponding to three times the standard deviation of the responses in chemical blanks.

Lipid assessment & adjustment

HCB, PCB-153 and *p,p´*-DDE were lipid adjusted, with the total lipid concentration in serum (g/L) calculated as total = 0.96 + 1.28*(triglycerides + cholesterol) (Rylander et al. 2006). Serum concentrations of triglycerides and cholesterol were determined by enzymatic methods using reagents from Roche Diagnostics (Mannheim, Germany). The inter-assay coefficients of variation for cholesterol and triglyceride determinations were 1.5–2.0%. The average molecular weights of triglycerides were assumed to be 807. For cholesterol we used an average molecular weight of 571, assuming that the proportion of free and esterified cholesterol in plasma was 1:2 (Jönsson et al. 2005).

Outcome assessment

Reproductive hormones were measured in male serum samples at Malmö University Hospital as previously described in detail (Giwercman et al. 2006) Measurements of follicle-stimulating hormone (FSH), luteinizing hormone (LH), and estradiol were made using a competitive binding immunoenzymatic assay (UniCel DxI 800 Beckman Access Immunoassay system, Chaska, MN, USA). Serum total testosterone levels were measured by means of a competitive immunoassay (Access; Beckman Coulter Inc., Fullerton, CA, USA). Sex hormone-binding globulin (SHBG) was measured using a fluoroimmunoassay (Immulite 2000; Diagnostic Products Corporation, Los Angeles, CA, USA). Inhibin B levels were assessed using a specific immunometric method (Groome et al. 1996). Free testosterone—the estimated bioactive fraction, unbound to SHBG or albumin—was calculated based on the measured total testosterone and SHBG levels (Vermeulen et al. 1999).

Conventional semen characteristics were assessed by centrally trained technicians as previously described (Toft et al. 2005b, 2006). Briefly, sperm concentration, motility and morphology were assessed according to WHO 1999 guidelines (World Health Organization 1999). Sperm concentration was determined in duplicate using an Improved Neubauer Hemacytometer (Paul Marienfeld, Bad Mergentheim, Germany). Sperm motility was determined by counting the proportion of a) rapid progressive spermatozoa; b) slow progressive spermatozoa; c) non-progressive motile spermatozoa; and d) immotile spermatozoa among 100 spermatozoa within each of two fresh

drops of semen, and progressively motile sperm were classified as a) plus b). For motility, samples with a delay of >1 hr from collection to analysis were excluded (n=28). Sperm morphology was assessed for at least 200 sperms in each sample by two technicians. Abnormalities were classified as head defects, midpiece defects, tail defects, cytoplasma drop and immature spermatozoa. For ejaculate volume and sperm counts, data was excluded if there had been spillage of the sample (n=67).

Sperm chromatin integrity. We evaluated two indices of DNA fragmentation index (DFI), as assessed by the sperm chromatin structure assay (SCSA) and the terminal deoxinucleotidyl transferase-driven dUTP Nick End Labelling (TUNEL) assay, as previously described (Spanò et al. 2005; Stronati et al. 2006). SCSA is the more frequently used, standardized method which detects sperms with abnormal chromatin packaging as characterized by susceptibility to acid-induced DNA denaturation *in situ* (Evenson et al. 2002). The TUNEL assay detects single- and double-strand DNA breaks, specifically free 3'-OH termini, present in spermatozoa. High DNA stainability (HDS), capturing incomplete chromatin condensation and considered a marker of immature sperm, was also determined via SCSA (Evenson et al. 2002).

Apoptotic markers. Apoptosis plays a crucial role in spermatogenesis. Proapoptotic (Fas) and anti-apoptotic (Bcl-xL) proteins present on ejaculated sperm were detected by means of indirect immunofluorescence (Stronati et al. 2006). Regarding the analysis of markers of sperm chromatin integrity and apoptotic markers, there were a high number of missing values due to a lost sample shipment from Ukraine and due insufficient number of sperm cells for some samples (Spanò et al. 2005; Stronati et al. 2006). A minimum of 10,000 sperm cells were measured by flow cytometry (Epics XL flow cytometer, Beckman Coulter-IL, Fullerton, Ca, USA). There were no significant differences in age and seminal parameters between participants with assessed versus missing data (Stronati et al. 2006).

Epididymal and accessory sex gland function. Motility of sperm is dependent maturation in the epididymis and interaction between prostatic and seminal vesicle secretions following ejaculation. Markers were assessed as previously described in detail (Elzanaty et al. 2006): neutral α-glucosidase (NAG) as a marker of epididymal function; prostate specific-antigen (PSA) and zinc as markers of prostatic function, and fructose as a marker of seminal vesicle function. The semen samples were first used to assess conventional semen characteristics, sperm chromatin integrity, apoptotic markers and the proportion of Y chromosome sperm cells (Spanò et al. 2005; Stronati et al. 2006; Tiido et al. 2006; Toft et al. 2005b). The epididymal and accessory sex gland function markers were subsequently assessed in semen samples with sufficient volume and in samples with no reported spillage (n=41–52 excluded). Samples were first analyzed for PSA, zinc and fructose, and the remaining amounts of seminal plasma were used to analyze NAG (Elzanaty et al. 2006).

Y chromosome sperm cells. The proportion of Y:X chromosome-bearing sperm was assessed in around 500 sperms per sample (range 253-743) using two-color fluorescence *in situ* hybridization

analysis (FISH), as previously described in detail (Tiido et al. 2006). Some samples were excluded from analysis because of low number of cells available or hybridization failure.

Statistical analysis

Partial least squares regression

To describe the univariate (single outcome) partial least squares (PLS or PLS1) regression model, for n observations, let $\boldsymbol{X}$ denote a matrix of mean-centered p predictors or exposures ($n \times p$), and $\boldsymbol{y}$ a vector of mean-centered continuous outcome data ($n \times q$ with q=1). Matrix or vector transposition is indicated by superscript $\top$, and the inverse of a matrix by superscript -1.

In both ordinary least squares (OLS) and PLS, $\boldsymbol{y}$ and $\boldsymbol{X}$ are related through a linear relationship $\boldsymbol{y} = \alpha + \boldsymbol{X\beta} + \in$ (with α=0 with centered inputs). For OLS, the least squares solution is $\widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}^{\top}\boldsymbol{X})^{-1}\boldsymbol{X}^{\top}\boldsymbol{y}$, and requires independent $\boldsymbol{X}$-variables (and $n > p$), whereas for PLS, the least squares solution is obtained via data compression into $\boldsymbol{K}$ latent components (latent variables; where $p \leq \boldsymbol{K}$, thus allowing for $n < p$), and PLS can accommodate multicollinear $\boldsymbol{X}$-variables. PLS decomposition is generalized (in matrix form) as (Indahl 2014; Wold et al. 2001):

$$\boldsymbol{y} = \boldsymbol{T}\boldsymbol{q}^{\top} + \boldsymbol{f} \;\; = \boldsymbol{X}\boldsymbol{W}\boldsymbol{q}^{\top} + \boldsymbol{f} \;\; = \boldsymbol{X}\boldsymbol{\beta}_{\boldsymbol{PLS}} + \boldsymbol{f}$$

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\top} + \boldsymbol{E}$$

where $\boldsymbol{T}$ represents the matrix ($n \times \boldsymbol{K}$) of latent components or 'scores' of orthogonal, linear combinations of $\boldsymbol{X}$ for the $\boldsymbol{K}$ number of model components; $\boldsymbol{q}$ and $\boldsymbol{P}$ represent the vector of $\boldsymbol{y}$- and matrix of $\boldsymbol{X}$-loading coefficients or 'loadings'; and $\boldsymbol{f}$ and $\boldsymbol{E}$ the random errors. $\boldsymbol{W}$ is a matrix ($p \times \boldsymbol{K}$) of direction vectors or 'loading weights' ($\boldsymbol{w}_{\boldsymbol{K}}$). Latent components are derived via successive optimizations (depending on the PLS1 algorithm), such that $\widehat{\boldsymbol{T}} = \boldsymbol{X}\widehat{\boldsymbol{W}}_{\boldsymbol{K}}$ or $\widehat{\boldsymbol{T}} = \boldsymbol{X}\widehat{\boldsymbol{W}}(\widehat{\boldsymbol{P}}^{\top}\widehat{\boldsymbol{W}})^{-1}$. (s)PLS models were fitted with the SIMPLS algorithm, described in detail elsewhere (Indahl 2014; Jong 1993).

As such, latent components are constructed ordered by the amount of explained variance in $\boldsymbol{y}$, so that the first component has the largest covariance with the outcome, the second component, the second largest covariance, and so on. PLS regression coefficients are computed as $\widehat{\boldsymbol{\beta}}_{\boldsymbol{PLS}} = \widehat{\boldsymbol{W}}_{\boldsymbol{K}}\widehat{\boldsymbol{q}}^{\top}$. For a $\boldsymbol{K}$=1 component model, PLS coefficients and weights are proportional to the univariable OLS coefficients; this is not the case for a PLS model with $\boldsymbol{K}$>1, in which coefficients are weighted across components.

Sparse partial least squares regression

Sparseness, in this context, means that a solution is obtained with a subset of the initial input variables. Noisy or uninformative variables are eliminated. To achieve sparsity, penalization (also called shrinkage) is introduced, in which regression coefficients are shrunk (down-weighted) via a penalty function towards zero or set to zero, depending on the penalty.

In sparse partial least squares (sPLS) regression, penalization is applied during the dimension reduction step. We applied the sPLS algorithm of Chun and Keleş (2010), as implemented in the R

spls package (Chung et al. 2013; Martens and Naes 1989; Mevik and Wehrens 2007). In brief, a penalty (η) is applied to a surrogate of the direction vector ($\boldsymbol{w}$, which is close to the original direction vector, as elaborated in Chun and Keleş (2010)). The sPLS sparsity penalty (η) approximates the L_1 penalty of LASSO (Tibshirani 1996): $\min_{\beta} \| \boldsymbol{y} - \boldsymbol{X}\beta\|^2 + \lambda\|\boldsymbol{\beta}\|_1$ where $\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^{p} |\boldsymbol{\beta}_j|$. [N.B.: L_1 corresponds, in the Bayesian setting, to a Laplace or double-exponential prior distribution (Cole et al. 2014).] The univariate sPLS penalization can be simplified to (Chun and Keleş 2010; Filzmoser et al. 2012):

$$\widehat{\boldsymbol{w}} = \max\left(0, |\widetilde{w}| - \eta \max_{1\le i\le p}|\widetilde{\boldsymbol{w}}_i|\right) \cdot sign(\widetilde{\boldsymbol{w}})$$

where $\widetilde{\boldsymbol{w}} = \left(\widetilde{w}_1 \ldots\ldots \widetilde{w}_p\right)^{\top}$ are the estimated PLS direction vectors with $\widetilde{\boldsymbol{w}}_1 = \boldsymbol{X}^{\top}\boldsymbol{y}/\| \boldsymbol{X}^{\top}\boldsymbol{y} \|_2$, and $0 \le \eta \le 1$ (sparsity increases as η approaches 1, and if $\eta=0$ then the model is equivalent to PLS). A fraction of each direction vector is retained. Thus, sPLS is a two-stage procedure; once sparsity has been applied on the direction vectors (and implicitly, a subset of $\boldsymbol{X}$-variables selected), coefficients are derived from ordinary PLS regression.

Imputation: exposure data

For the exposure data, we imputed values <LOD (0–18%) and, for sPLS-regression analyses only, values missing-at-random (12–16% for metals, 4% for PCB-153 and *p,p'*-DDE, and 2% for other compounds). Data was considered missing-at-random because some serum (n=13) and whole blood (n=71) samples were untraceable or depleted in the time since the baseline study. Further, regarding measurement of metals in whole blood, for some samples (n=26) there was insufficient volume to measure all three metals and a choice was made to analyze Hg and not Cd and Pb.

We used a maximum likelihood method to impute values <LOD based on the distribution estimated from detected values and conditional on the structure of the $\boldsymbol{X}$-matrix, and under the assumption that measurements follow a parametric (log-normal) underlying distribution. Specifically, we performed iterative imputation in which the mean of the imputation distribution for each missing exposure value was dependent on the study population (Greenland/Warsaw/Kharkiv) and levels of the other exposures, while the (residual) standard deviation was allowed to vary by study population. Each value <LOD was substituted with one imputed value (single imputation), which yields approximately unbiased estimates when measurements <LOD are less than 30% (Lubin et al. 2004).

Imputation: covariate data

We applied a minimal set of *a priori* selected confounders (specified in the main text, Table 1 and Table S3). As a substantial portion of data was missing for abstinence period (n=45) and time of blood sampling (n=98; all Greenlandic), we imputed missing data for these two covariates for the primary analyses. We performed single, fill-in imputation: for abstinence period, we assumed missing data followed the same distribution as the available data did; for time of blood sampling, we assumed the same proportion of Greenlandic participants were sampled prior to 12:00 hr as for the available data for Greenland (~20%), and randomly imputed a dichotomous (morning yes/no) value resulting in this proportion. In addition, missing values for age (n=5) and body mass index (BMI)

(n=7) were replaced with the population-specific median values. Missing values for cotinine (n=13) were replaced with the respective median cotinine value for smokers and non-smokers, based on self-reported smoking status; and values <LOD (0.7 ng/mL, 35%) were imputed based on a log-normal distribution, as described above.

Table S1. Analytical reproducibility of exposure and outcome biomarkers, and variability across study populations.

	Reproducibility		Inter population-variation: ICC[c]
	Coefficient of variation[a] (%)	Concentration[b] (ng/mL)	
Exposure			
Phthalate metabolites			
MEHHP	8	2.4	0.86
MEOHP	9	3.0	0.98
MECPP	18	1.3	0.83
MHiNP	8	2.2	0.91
MOiNP	7	2.0	0.85
MOiCP	19	3.5	1.00
Metals			
Hg	6	2.0	0.30
Cd	4	24	0.84
Pb	6	1.0	0.89
Perfluoroalkyl acids	(Lindh et al. 2012)		
PFHxS	8	1.5	0.19
PFOA	6	3.9	0.30
PFOS	5	26	0.15
PFNA	9	1.6	0.72
PFDA	9	0.6	0.33
PFUnDA	10	0.7	0.24
PFDoDA	22	0.08	0.32
Organochlorines	(Jönsson et al. 2005)		
HCB	37	0.1	0.24
PCB-153	10	0.5	0.26
p,p′-DDE	8	3	0.80
Outcome			
Reproductive hormones in serum	(Giwercman et al. 2006)		
FSH (IU/L)	3.5	5.5 IU/L	0.96
	4.1[d]	23.6 IU/L	
LH (IU/L)	5.2	4.0 IU/L	1.00
	2.3	19.3 IU/L	
Inhibin B (ng/L)	< 7	–	0.93
SHBG (nmol/L)	3.7	29 nmol/L	0.92
	6.7	85 nmol/L	
Total testosterone (nmol/L)	2.8	2.9 nmol/L	0.78
	302	8.1 nmol/L	
Free testosterone (nmol/L)	N/A	–	0.77
Estradiol (pmol/L)	17.4	44 pmol/L	0.84
	6.7	303 pmol/L	
Conventional semen characteristics	(Toft et al. 2005b)		
Semen volume (mL)	N/A	–	1.00
Sperm concentration (million/mL)	8.1	–	1.00
Total sperm count (million/ejaculate)	N/A	–	1.00

	Reproducibility		Inter population-variation: ICC[c]
	Coefficient of variation[a] (%)	Concentration[b] (ng/mL)	
Morphologically normal sperm (%)	N/A	–	1.00
Progressive sperm motility (%)	11	–	0.98
Sperm chromatin integrity	(Spanò et al. 2005)		
SCSA DFI (%)	6.0	–	0.92
High DNA stainability (%)	4.8	–	0.95
TUNEL DFI (%)	<5	–	0.60
Apoptotic markers	(Stronati et al. 2006)		
Fas positivity (%)	6	–	0.83
Bcl-xL positivity (%)	9	–	0.66
Epididymal and accessory sex gland function	(Elzanaty et al. 2006)		
NAG (mU/ejaculate)	N/A	–	0.93
PSA (μg/ejaculate)	N/A	–	0.93
Zinc (μmmol/ejaculate)	N/A	–	0.94
Fructose (μmmol/ejaculate)	N/A	–	0.99
Y:X chromosome sperm cells	(Tiido et al. 2006)		
Y chromosome (%)	2.3, 3.3[e]	–	0.93

Abbreviations: DFI, DNA fragmentation index; FSH, follicle-stimulating hormone; HCB, hexachlorobenzene; LH, luteinizing hormone; MECPP, mono-(2-ethyl-5-carboxypentyl) phthalate; MEHHP, mono-(2-ethyl-5-hydroxyhexyl) phthalate; MEOHP, mono-(2-ethyl-5-oxo-hexyl) phthalate; MOiCP, mono-(4-methyl-7-carboxyheptyl) phthalate; MHiNP, mono-(4-methyl-7-hydroxyloctyl) phthalate; MOiNP, mono-(4-methyl-7-oxooctyl) phthalate; N/A, not available or not applicable; NAG, neutral α-glucosidase; PCB-153, polychlorinated biphenyl 153; PFHxS, perfluorohexane sulfonic acid; PFNA, perfluorononanoic acid; PFOS, perfluorooctane sulfonic acid; PFOA, perfluorooctanoic acid; *p,p'*-DDE, 1,1-dichloro-2,2-bis(p-chlorophenyl)-ethylene; PSA, prostate-specific antigen; SCSA, sperm chromatin structure assay; SHBG, sex hormone-binding globulin; TUNEL, terminal deoxynucleotidyl transferase dUTP nick end-labeling.

[a] The coefficient of variation for the exposures was calculated as the standard deviation/mean ($|\sigma/\mu|*100$) of duplicate quality control samples worked-up and analyzed on different days (Jönsson et al. 2005; Lindh et al. 2012).

[b] The concentration(s) in quality control samples from which the reproducibility was determined.

[c] The intraclass correlation coefficient (ICC) was calculated from the within-population and the between-population variances from a one-way ANOVA. The higher the ICC value, the more similar the biomarker distributions across study populations.

[d] For most reproductive hormones, a coefficient of variation was determined for two different quality control concentrations.

[e] Interobserver and intraobserver coefficients of variation, respectively.

Table S2. Blood levels[a] of measured contaminants in male partners of pregnant women.

Exposure	LOD (ng/mL)	% >LOD	n[b]	All 3 populations (n=602) GM (5, 95 P)	Greenland (n=199) GM (5, 95 P)	Warsaw, Poland (n=197) GM (5, 95 P)	Kharkiv, Ukraine (n=206) GM (5, 95 P)	p-value[c]
Phthalate metabolites[d] (ng/mL)								
MEHHP	0.2	98	580	0.73 (0.26, 2.45)	1.00 (0.44-2.55)	0.62 (0.26, 1.54)	0.63 (0.21, 3.26)	<0.001
MEOHP	0.2	49	287	—[e] (<LOD, 0.52)	— (<LOD, 0.55)	— (<LOD, 0.40)	— (<LOD, 0.63)	—
MECPP	0.1	100	589	1.61 (0.58, 5.63)	1.17 (0.45, 3.72)	1.62 (0.75, 4.56)	2.16 (0.71, 8.93)	<0.001
ΣDEHPom	—	—	—	2.74 (1.22, 7.87)	2.54 (1.12, 5.85)	2.55 (1.39, 6.25)	3.17 (1.17, 10.55)	<0.001
ΣDEHPom (nmol/mL)	—	—	—	0.009 (0.004, 0.026)	0.008 (0.004, 0.019)	0.008 (0.005, 0.021)	0.010 (0.004, 0.035)	<0.001
MHiNP	0.1	93	549	0.24 (<LOD, 0.83)	0.30 (0.12, 0.86)	0.23 (0.10, 0.58)	0.20 (<LOD, 1.04)	<0.001
MOiNP	0.03	39	231	— (<LOD, 0.13)	— (<LOD, 0.11)	— (<LOD, 0.08)	— (<LOD, 0.34)	—
MOiCP	0.1	99	586	0.60 (0.19, 3.43)	0.57 (0.21, 1.61)	0.61 (0.29, 1.60)	0.61 (0.16, 5.72)	0.62
ΣDiNPom	—	—	—	0.91 (0.36, 4.11)	0.96 (0.40, 2.33)	0.90 (0.48, 2.25)	0.88 (0.27, 7.51)	0.48
ΣDiNPom (nmol/mL)	—	—	—	0.003 (0.001, 0.013)	0.003 (0.001, 0.007)	0.003 (0.002, 0.007)	0.003 (0.001, 0.024)	0.48
Metals (ng/mL)								
Hg	0.1	100	531	2.10 (0.38, 33.02)	8.66 (0.85, 49.12)	1.01 (0.39, 2.60)	0.84 (0.31, 2.24)	<0.001
Cd	0.02	100	505	0.50 (0.12, 2.59)	0.72 (0.13, 2.95)	0.33 (0.13, 2.16)	0.53 (0.10, 2.74)	<0.001
Pb	0.08	100	505	27.60 (14.47, 66.06)	29.95 (14.12, 84.90)	22.93 (14.00, 38.63)	31.15 (16.34, 69.21)	<0.001
Perfluoroalkyl acids (ng/mL)								
PFHxS	0.06	100	588	0.97 (0.21, 3.71)	2.39 (1.18, 6.15)	1.16 (0.68, 2.02)	0.35 (0.16, 0.72)	<0.001
PFOA	0.6	97	573	3.05 (0.78, 8.30)	4.60 (2.76, 7.36)	4.86 (2.54, 9.27)	1.33 (0.44, 3.74)	<0.001
PFOS	0.2	100	589	18.11 (4.52, 73.20)	47.39 (25.66, 103.02)	17.69 (9.61, 29.14)	7.32 (3.65, 14.13)	<0.001
PFNA	0.2	100	589	1.31 (0.59, 3.54)	1.85 (0.74, 4.65)	1.20 (0.66, 2.20)	1.02 (0.53, 2.13)	<0.001
PFDA	0.2	82	481	0.41 (<LOD, 1.66)	0.88 (0.33, 2.24)	0.39 (0.21, 0.73)	0.20 (<LOD, 0.47)	<0.001
PFUnDA	0.3	39	232	— (<LOD, 2.92)	— (<LOD, 4.08)	— (<LOD, 0.35)	— (<LOD, 0.37)	—
PFDoDA	0.07	29	180	— (<LOD, 0.31)	— (<LOD, 0.43)	— (<LOD, 0.08)	— (<LOD, <LOD)	—
Organochlorines (ng/g lipid)								
HCB	0.05	93	539	46.53 (6.48, 294.96)	58.83 (17.77, 211.83)	12.25 (4.57, 31.71)	135.22 (55.36, 469.21)	<0.001
PCB-153	0.05	95	551	55.68 (8.52, 579.82)	223.20 (50.00, 1092.55)	16.80 (6.36, 37.68)	44.69 (15.18, 138.61)	<0.001
p,p′-DDE	0.1	100	577	677.83 (196.18, 2223.44)	567.93 (108.66, 2188.36)	516.79 (224.18, 1093.09)	1051.35 (415.09, 2906.75)	<0.001

Cd, cadmium; GM, geometric mean; HCB, hexachlorobenzene; Hg, mercury; LOD, limit of detection; MECPP, mono-(2-ethyl-5-carboxypentyl) phthalate; MEHHP, mono-(2-ethyl-5-hydroxyhexyl) phthalate; MEOHP, mono-(2-ethyl-5-oxo-hexyl) phthalate; MOiCP, mono-(4-methyl-7-carboxyheptyl) phthalate; MHiNP, mono-(4-methyl-7-hydroxyloctyl) phthalate; MOiNP, mono-(4-methyl-7-oxooctyl) phthalate; P, percentile; Pb, lead; PCB-153, polychlorinated

biphenyl 153; PFDoDA, perfluorododecanoic acid; PFHxS, perfluorohexane sulfonic acid; PFNA, perfluorononanoic acid; PFOS, perfluorooctane sulfonic acid; PFOA, perfluorooctanoic acid; PFUnDA, perfluoroundecanoic acid; *p,p'*-DDE, 1,1-dichloro-2,2-bis(p-chlorophenyl)-ethylene.

[a] Values <LOD were imputed.

[b] Available for analysis (589 for phthalates, PFAAs and HCB; 531 for Hg; 505 for Cd and Pb; 578 for PCB-153 and p,p′-DDE) and measured value >LOD.

[c] Test for difference in levels between the three populations (ANOVA).

[d] The molar sums (nmol/mL) of the three oxidative DEHP and DiNP metabolites were calculated (ΣDEHPom and ΣDiNPom), and are also presented corrected for molecular weight, based on the weighted average molecular weight (ng/mL).

[e] GM and ANOVA p-value not calculated if >30% of data was below the LOD.

Table S3. Exposure-outcome associations identified from sPLS population-adjusted[a] and further adjusted[b] models: (1) sPLS and (2) OLS regression coefficients per ln-unit change in exposure, and corresponding percent changes in outcome per interquartile range increase in exposure.

			(1) Multi-pollutant sPLS models					(2) Single-pollutant OLS models				
			Pop.-centered inputs[a]		Pop.-centered and 'pre-standardized'[b]			Adjusted for pop.[a]	Further adjusted[b]			
Outcome (mean)	**Exposure[c] (IQR)**	**n**	***K*, *η* (Q^2 %)[d]**	**β_{sPLS}**	***K*, *η* (R^2, Q^2 %)[d]**	**β_{sPLS}**	**% Δ /IQR[e]**	**β_{OLS}[d]**	**β_{OLS}[d]**	**95% CI**	**% Δ /IQR[e]**	**95% CI**
LH[c] (IU/L)	*p,p′*-DDE (416.87–1143.00 ng/g)	456	—	—	1, 0.99 (2.13,1.27)	0.083	8.70	0.080[f]	0.083[f,g,h]	(0.031, 0.135)	8.73	(3.18, 14.59)
Inhibin B (182.3 ng/L)	Hg (0.698–4.852 ng/mL)	456	1, 0.99 (0.95)	10.580	1, 0.99 (2.05,1.12)	10.788	11.48	10.588[f]	10.816[f]	(3.899, 17.733)	11.51	(4.15, 18.86)
SHBG[c] (nmol/L)	MEHHP (0.440–1.237 ng/mL)	455	1, 0.53 (1.83)	-0.015	—	—	—	-0.046	-0.024[g]	(-0.068, 0.020)	-2.45	(-6.79, 2.09)
	MHiNP (0.146–0.353ng/mL)			-0.014		—	—	-0.043	-0.033[g]	(-0.076, 0.010)	-2.87	(-6.48, 0.88)
	MOiCP (0.354–0.868 ng/mL)			-0.017		—	—	-0.041	-0.035[g]	(-0.069, -0.001)	-3.09	(-6.00, -0.09)
	Cd (0.205–1.114 ng/mL)			0.015		—	—	0.033	0.005[g]	(-0.033, 0.043)	0.85	(-5.43, 7.55)
	Pb (20.55–35.41 ng/mL)			0.021		—	—	0.105	0.064[g]	(-0.011, 0.139)	3.54	(-0.60, 7.86)
	HCB (17.31–107.02 ng/g)			0.026		—	—	0.083[f]	0.057[g,h]	(0.012, 0.103)	10.94	(2.21, 20.64)
	PCB-153 (19.59–131.02 ng/g)			0.027		—	—	0.074[f]	0.045[g,h]	(0.005, 0.085)	8.93	(0.95, 17.53)
	p,p′-DDE			0.021		—	—	0.062	0.035[g,h]	(-0.007, 0.078)	3.59	(-0.70, 8.19)
Total testosterone (15.81 nmol/L)	MECPP (1.020–2.265 ng/mL)	456	1, 0.58 (2.80)	-0.329	1, 0.90 (3.13,1.05)	—	—	-0.811	-0.727[g]	(-1.357, -0.097)	-3.67	(-6.85, -0.49)
	MHiNP			-0.535		-1.141	-6.36	-1.166[f]	-1.153[f,g]	(-1.741, -0.565)	-6.43	(-9.70, -3.15)
	MOiCP			-0.526		—	—	-0.746	-0.684[g]	(-1.162, -0.206)	-3.88	(-6.59, -1.17)
	Cd			0.653		—	—	0.759[f]	0.515	(-0.015, 1.045)	5.51	(-0.16, 11.19)
	ΣDiNPom (0.0018–0.0039 nmol/mL)			*NT*		*NT*	*NT*	-0.976[f]	-0.929[f]	(-1.459, -0.399)	-4.70	(-7.39, -2.02)
Free testosterone (0.339 nmol/L)	MECPP	455	1, 0.61 (0.95)	-0.0055	1, 0.64 (2.87,0.13)	—	—	-0.013	-0.013[g]	(-0.025, 0.000)	-3.06	(-5.88, 0.00)
	MHiNP			-0.0084		-0.0113	-2.93	-0.019[f]	-0.019[f,g]	(-0.032, -0.007)	-4.93	(-8.31, -1.82)
	MOiCP			-0.0059		-0.0091	-2.41	-0.011	-0.010[g]	(-0.020, 0.000)	-2.64	(-5.28, 0.00)
	Cd			0.0083		0.0091	4.52	0.014[f]	0.012	(0.001, 0.023)	5.98	(0.50, 11.47)
Semen volume[c] (mL)	MEHHP	535	1, 0.99 (0.82)	-0.110	1, 0.99 (2.12,1.21)	-0.106	-10.35	-0.110[f]	-0.106[f]	(-0.167, -0.045)	-10.38	(-15.86, -4.55)
Progressive sperm (57%)	PCB-153	565	1, 0.99 (1.18)	-3.488	1, 0.99 (1.70,1.00)	-3.365	-11.22	-3.488[f]	-3.365[f,h]	(-5.484, -1.246)	-11.22	(-18.29, -4.16)
TUNEL DFI[c] (%)	MEHHP	462	1, 0.63 (2.72)	-0.102	1, 0.99 (3.00,2.25)	—	—	-0.177[f]	-0.185[f,g]	(-0.303, -0.068)	-17.41	(-26.90, -6.79)
	MHiNP			-0.133		-0.218	-17.48	-0.217[f]	-0.218[f,g]	(-0.332, -0.104)	-17.47	(-25.36, -8.76)
	Cd			-0.161		—	—	-0.130[f]	-0.090	(-0.189, 0.008)	-14.13	(-27.37, 1.36)
NAG[c] (mU/ejaculate)	MEHHP	448	2, 0.99 (3.76)	-0.170	1, 0.77 (2.73,0.62)	-0.163	-15.52	-0.178[f]	-0.164[f]	(-0.255, -0.073)	-15.60	(-23.18, -7.27)
	Cd			-0.118		—	—	-0.123[f]	-0.109[g]	(-0.188, -0.030)	-16.85	(-27.25, -4.95)

Pop., study population; *NT*, not tested; —, indicates association was not selected in sPLS model.

[a] The 'unadjusted' models were only adjusted for study population: exposure and outcome variables were mean-centered by study population prior to sPLS modeling, and study population was included as a covariate in OLS models.

[b] 'Adjusted' models included additional potential confounders. sPLS models were constructed with outcome and exposure variables 'pre-standardized' by confounders, inputting the residuals of linear regression models of each outcome versus confounders, and each exposure versus confounders. Confounders were included as covariates

in OLS models. All models were adjusted for study population and cotinine, and additionally for age, BMI and time of blood sampling (morning, yes/no) for the reproductive hormones; for ln-abstinence period for all conventional semen characteristics except proportion normal sperm; and for age and ln-abstinence period for markers of sperm chromatin integrity, apoptotic markers, and markers of epididymal and accessory sex gland function.

[c] All exposures and some outcomes, as indicated, were ln-transformed in statistical analyses.

[d] *K* and η represent the tuning parameters for the sPLS model; *K*, the number of components used to construct the model, and η , the degree of sparsity (with η approaching 1 yielding a sparser model). R^2 is the explained variance of ***y*** by ***X***. It represents the partial variance explained by the exposure(s) only, as input ***X***-exposure and ***y***-outcome data were pre-standardized for covariates. Q^2 represents the cross-validated fraction of predicted ***y***-variation (or predictive ability of the model); $Q^2 = 1 - PRESS / SS$, where $PRESS = \sum(\hat{y}_i - y_i)^2$ is the predictive residual error sum of squares, and $SS = \sum(y_i - \bar{y}_i)^2$ is the sum of squares of ***y*** corrected for the mean.

[e] sPLS and OLS regression coefficients derived per ln-unit exposure were transformed to represent the percent change in outcome associated with the interquartile range in exposure (IQR; the 75th compared to the 25th percentile in ln-exposure). For ln-transformed outcomes, this is the proportional change: $\left(e^{\beta_{ij} * IQR_{\ln(x_i)}} - 1\right) * 100$. For untransformed outcomes, this is the absolute change in the outcome relative to the arithmetic mean outcome level: $\left(\beta_{ij} * IQR_{\ln(x_i)}\right) / mean_{y_j} * 100$. Mean outcome values for the untransformed outcomes are presented. We used IQRs for the full population (n=602), and present untransformed values.

[f] Significant after adjustment for multiple comparisons (FDR <10%): 330 tests in the primary analysis; 374 tests in the additional analysis with ΣDEHPom and ΣDiNPom.

[g] Interaction p-value <0.10 for the cross-product term between exposure and study population (see supplementary figure S3 for population-stratified regression plots).

[h] Sensitivity analysis: adjusted β_{OLS} (95% CI) for models with organochlorines unadjusted for lipids (ng/mL), and with total lipids (g/L) included as an additional covariate: LH and *p,p′*-DDE, 0.070 (0.017, 0.123); SHBG and HCB, 0.036 (-0.010, 0.082); PCB-153, 0.043 (0.002, 0.083); *p,p′*-DDE, 0.018 (-0.024, 0.061); progressive sperm and PCB-153: -3.375 (-5.543, -1.207).

Figure S1. Pearson correlation coefficients, also represented as a heat map, between the reproductive function biomarkers.

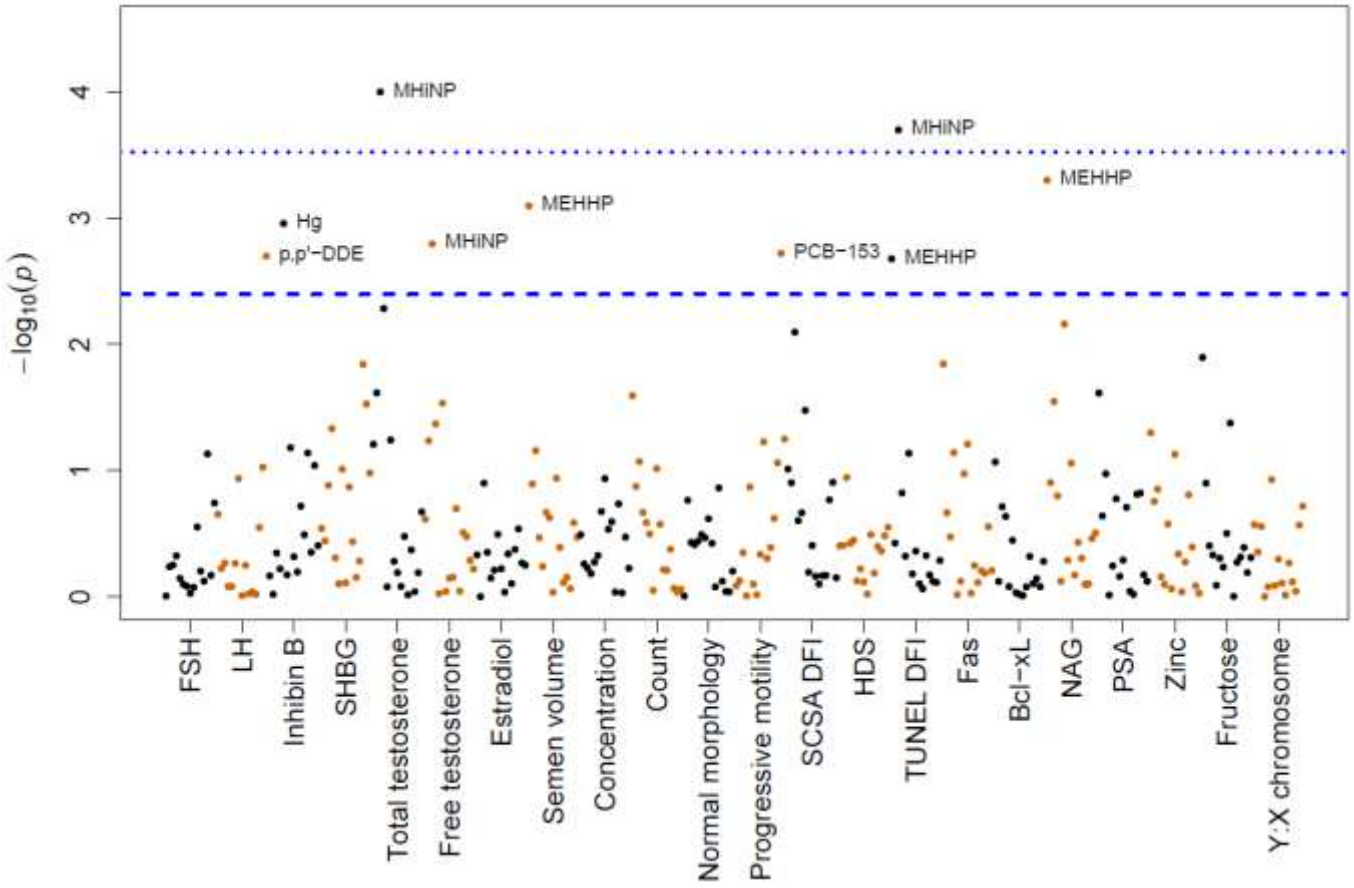

Figure S2. The p-values ($-\log_{10}$ scale) from single exposure-outcome OLS regression analyses, plotted per outcome.

Analyses are adjusted for study population and cotinine, and variably adjusted for age, BMI, abstinence period and time of blood sampling as indicated in Tables 1 and S3. The dotted and dashed lines demarcate a false discovery rate of <5% and <10%, respectively. Each dot corresponds to the p-value from a single exposure-outcome association, and alternating black and orange colors delineate outcomes.

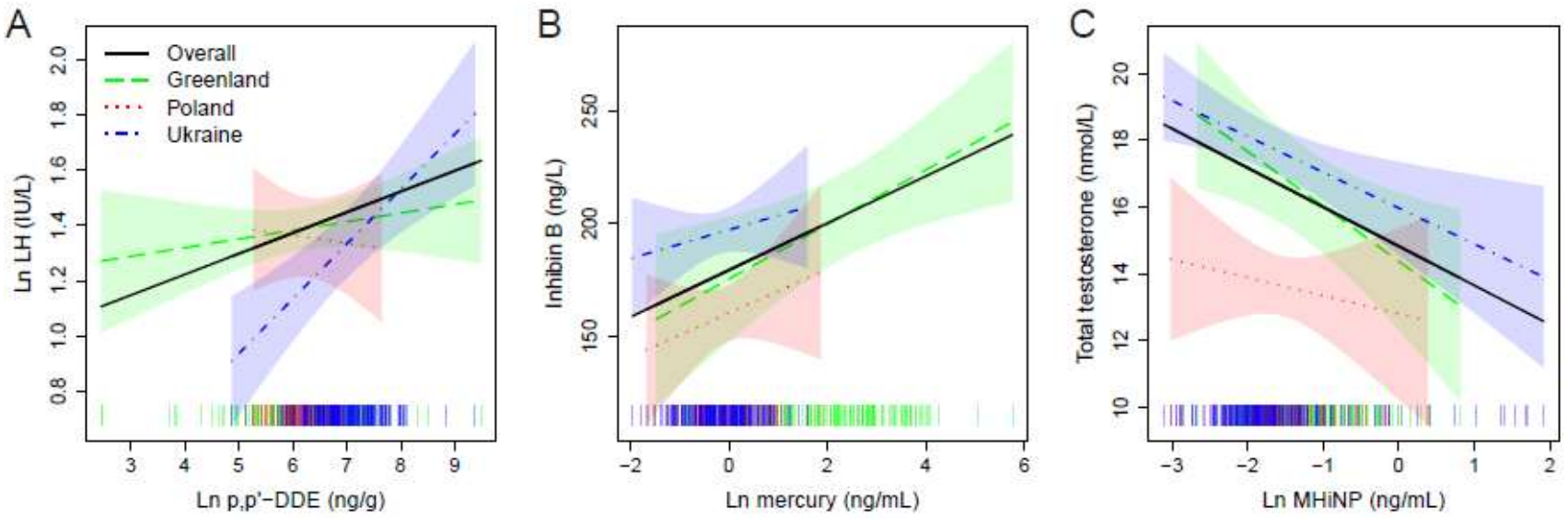

Figure S3. Exposure-outcome associations, plotted as linear regressions across all and stratified by study population (A-L).

Models are adjusted for study population and cotinine, and variably adjusted for age, BMI, abstinence period and time of blood sampling as indicated in the footnotes of table S3. Predicted functions, with confounders set at the mean of continuous confounders and morning time of blood sampling are presented: population-specific exposure-outcome relationships (dashed lines) and 95% confidence intervals (shaded), and an overall exposure-outcome relationship for the pooled analysis, plotted at the Greenland-specific intercept (solid black line). Rug plots display the density of the exposure data.

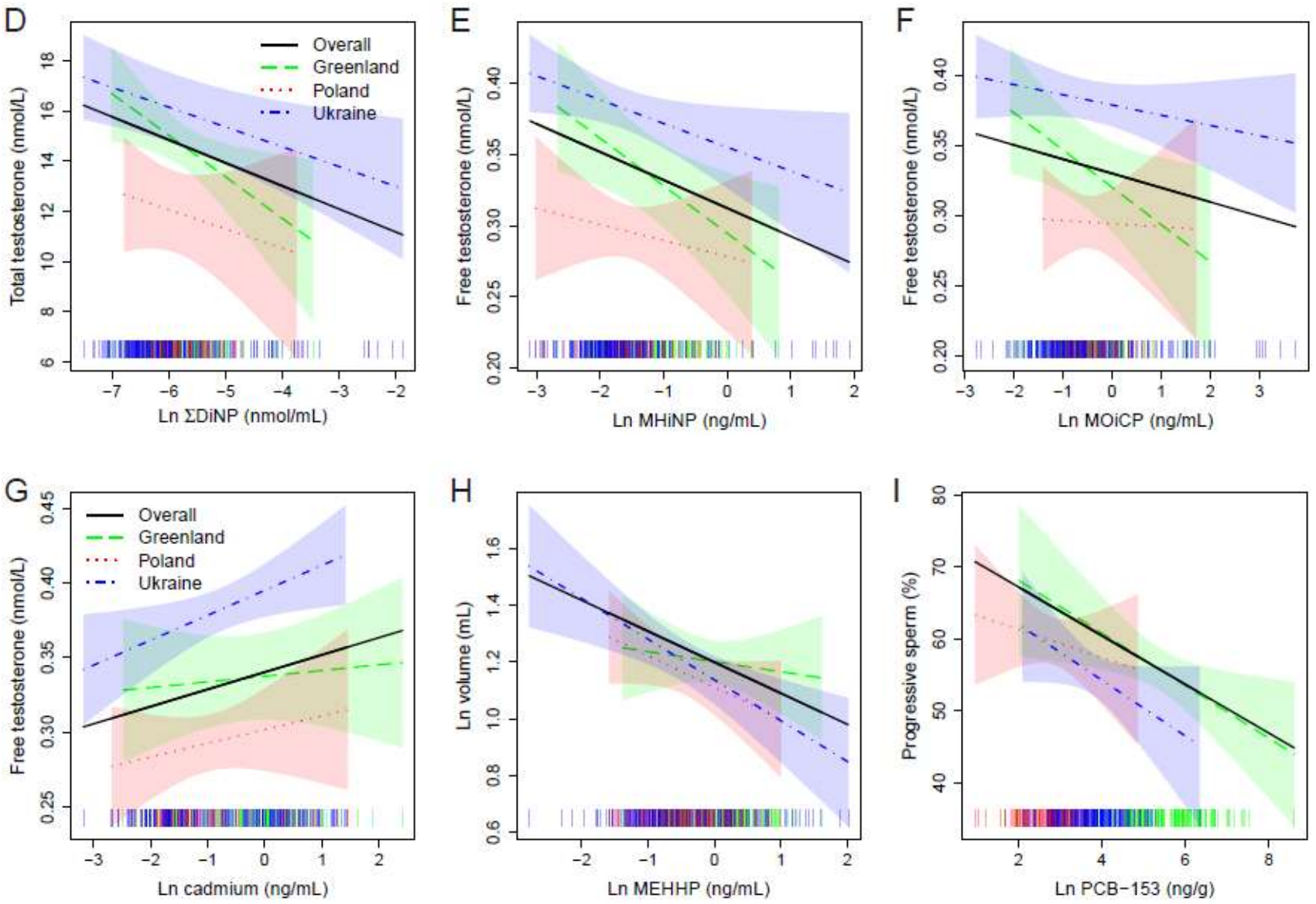

Figure S3. *Continued*

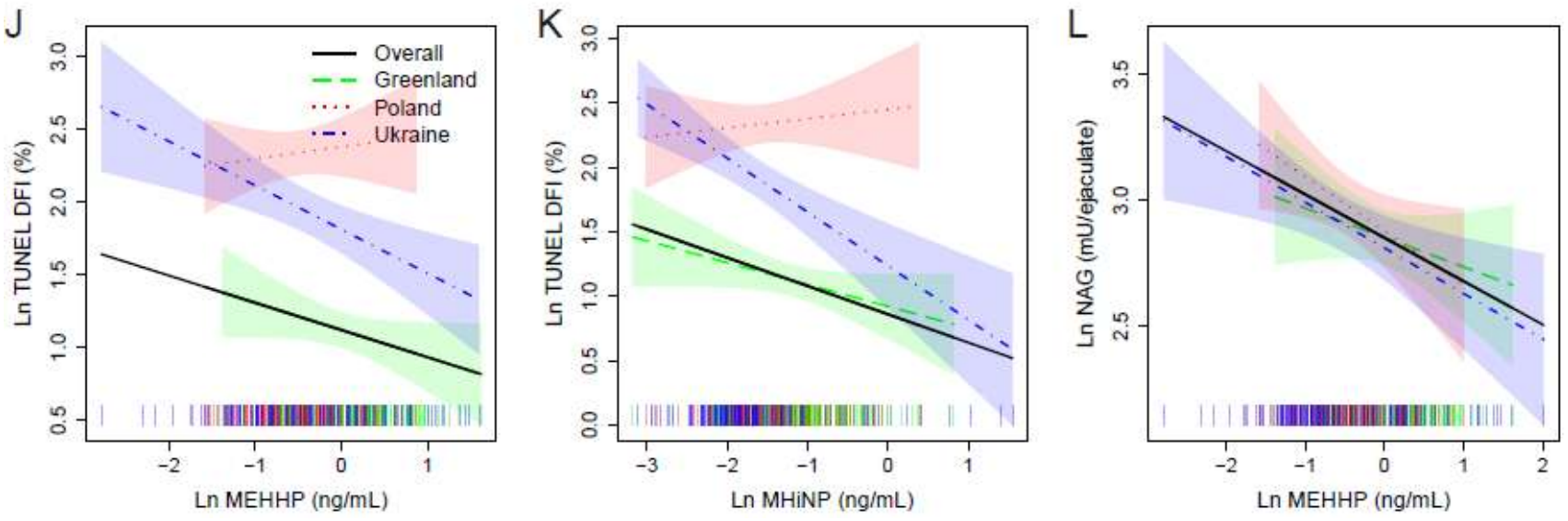

Figure S3. *Continued*

References

Chun H, Keleş S. 2010. Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *J R Stat Soc Series B Stat Methodol* 72:3–25.

Chung D, Chun H, Keles S. 2013. spls: Sparse Partial Least Squares (SPLS) Regression and Classification. R Package version 22-1. Available: http://cran.r-project.org/package=spls.

Cole SR, Chu H, Greenland S. 2014. Maximum likelihood, profile likelihood, and penalized likelihood: a primer. *Am J Epidemiol* 179:252–60.

de Jong S. 1993. SIMPLS: an alternative approach to partial least squares regression. *Chemom Intell Lab Syst* 18: 251–263.

Elzanaty S, Rignell-Hydbom A, Jönsson BAG, Pedersen HS, Ludwicki JK, Shevets M, et al. 2006. Association between exposure to persistent organohalogen pollutants and epididymal and accessory sex gland function: multicentre study in Inuit and European populations. *Reprod Toxicol* 22:765–73.

Evenson DP, Larson KL, Jost LK. 2002. Sperm chromatin structure assay: its clinical use for detecting sperm DNA fragmentation in male infertility and comparisons with other techniques. *J Androl* 23:25–43.

Filzmoser P, Gschwandtner M, Todorov V. 2012. Review of sparse methods in regression and classification with application to chemometrics. *J Chemom* 26:42–51.

Frederiksen H, Jørgensen N, Andersson A-M. 2010. Correlations between phthalate metabolites in urine, serum, and seminal plasma from young Danish men determined by isotope dilution liquid chromatography tandem mass spectrometry. *J Anal Toxicol* 34: 400–10.

Giwercman AH, Rignell-Hydbom A, Toft G, Rylander L, Hagmar L, Lindh C, et al. 2006. Reproductive hormone levels in men exposed to persistent organohalogen pollutants: a study of inuit and three European cohorts. *Environ Health Perspect* 114:1348–53.

Groome NP, Illingworth PJ, O'Brien M, Pai R, Rodger FE, Mather JP, et al. 1996. Measurement of dimeric inhibin B throughout the human menstrual cycle. *J Clin Endocrinol Metab* 81: 1401–5.

Högberg J, Hanberg A, Berglund M, Skerfving S, Remberger M, Calafat AM, et al. 2008. Phthalate diesters and their metabolites in human breast milk, blood or serum, and urine as biomarkers of exposure in vulnerable populations. *Environ Health Perspect* 116:334–9.

Indahl U. 2014. The geometry of PLS1 explained properly: 10 key notes on mathematical properties of and some alternative algorithmic approaches to PLS1 modelling. *J Chemom* 28:168–80.

Jönsson BAG, Rylander L, Lindh C, Rignell-Hydbom A, Giwercman A, Toft G, et al. 2005. Inter-population variations in concentrations, determinants of and correlations between 2,2',4,4',5,5'-hexachlorobiphenyl (CB-153) and 1,1-dichloro-2,2-bis (p-chlorophenyl)-ethylene (p,p'-DDE): a cross-sectional study of 3161 men and women from Inuit and European populations. *Environ Health* 4:27.

Lindh CH, Rylander L, Toft G, Axmon A, Rignell-Hydbom A, Giwercman A, et al. 2012. Blood serum concentrations of perfluorinated compounds in men from Greenlandic Inuit and European populations. *Chemosphere* 88:1269–75.

Lubin JH, Colt JS, Camann D, Davis S, Cerhan JR, Severson RK, et al. 2004. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ Health Perspect* 112:1691–6.

Martens H, Naes T. 1989. Multivariate calibration. Chichester: Wiley.

Mevik B, Wehrens R. 2007. The pls package: Principal component and partial least squares regression in R. *J Stat Softw* 18: 1–24.

Rylander L, Nilsson-Ehle P, Hagmar L. 2006. A simplified precise method for adjusting serum levels of persistent organohalogen pollutants to total serum lipids. *Chemosphere* 62:333–6.

Spanò M, Toft G, Hagmar L, Eleuteri P, Rescia M, Rignell-Hydbom A, et al. 2005. Exposure to PCB and p, p'-DDE in European and Inuit populations: impact on human sperm chromatin integrity. *Hum Reprod* 20:3488–99.

Stronati A, Manicardi GC, Cecati M, Bordicchia M, Ferrante L, Spanò M, et al. 2006. Relationships between sperm DNA fragmentation, sperm apoptotic markers and serum levels of CB-153 and p,p'-DDE in European and Inuit populations. *Reproduction* 132:949–58.

Tibshirani R. 1996. Regression shrinkage and selection via the lasso. *J R Stat Soc Series B Stat Methodol* 58: 267–288.

Tiido T, Rignell-Hydbom A, Jönsson BAG, Giwercman YL, Pedersen HS, Wojtyniak B, et al. 2006. Impact of PCB and p,p'-DDE contaminants on human sperm Y:X chromosome ratio: studies in three European populations and the Inuit population in Greenland. *Environ Health Perspect* 114: 718–24.

Toft G, Axmon A, Giwercman A, Thulstrup AM, Rignell-Hydbom A, Pedersen HS, et al. 2005a. Fertility in four regions spanning large contrasts in serum levels of widespread persistent organochlorines: a cross-sectional study. *Environ Health* 4:26.

Toft G, Rignell-Hydbom A, Tyrkiel E, Shvets M, Giwercman A. 2005b. Quality control workshops in standardization of sperm concentration and motility assessment in multicentre studies. *Int J Androl* 28:144–9.

Toft G, Rignell-Hydbom A, Tyrkiel E, Shvets M, Giwercman A, Lindh CH, et al. 2006. Semen quality and exposure to persistent organochlorine pollutants. *Epidemiology* 17:450–8.

Vermeulen A, Verdonck L, Kaufman JM. 1999. A critical evaluation of simple methods for the estimation of free testosterone in serum. *J Clin Endocrinol Metab* 84: 3666–72.

Wold S, Sjöström M, Eriksson L. 2001. PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab* 58:109–130.

World Health Organization. 1999. WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction. 4th ed. Published on behalf of the World Health Organization by Cambridge University Press, Cambridge, UK; New York, NY.