ORIGINAL ARTICLE

Phthalates, perfluoroalkyl acids, metals and organochlorines and reproductive function: a multipollutant assessment in Greenlandic, Polish and Ukrainian men

Virissa Lenters,[1] Lützen Portengen,[1] Lidwien A M Smit,[1] Bo A G 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]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/oemed-2014-102264).

For numbered affiliations see end of article.

Correspondence to
Dr Roel Vermeulen, Institute for Risk Assessment Sciences, Utrecht University, P.O. Box 80.178, Utrecht 3508 TD, The Netherlands;
R.C.H.Vermeulen@uu.nl

Received 8 April 2014
Revised 2 August 2014
Accepted 20 August 2014
Published Online First
10 September 2014

ABSTRACT

Objectives Numerous environmental contaminants have been linked to adverse reproductive health outcomes. However, the complex correlation structure of exposures and multiple testing issues limit the interpretation of existing evidence. Our objective was to identify, from a large set of contaminant exposures, exposure profiles associated with biomarkers of male reproductive function.

Methods In this cross-sectional study (n=602), male partners of pregnant women were enrolled between 2002 and 2004 during antenatal care visits in Greenland, Poland and Ukraine. Fifteen contaminants were detected in more than 70% of blood samples, including metabolites of di(2-ethylhexyl) and diisononyl phthalates (DEHP, DiNP), perfluoroalkyl acids, metals and organochlorines. Twenty-two reproductive biomarkers were assessed, including serum levels of reproductive hormones, markers of semen quality, sperm chromatin integrity, epididymal and accessory sex gland function, and Y:X chromosome ratio. We evaluated multipollutant models with sparse partial least squares (sPLS) regression, a simultaneous dimension reduction and variable selection approach which accommodates joint modelling of correlated exposures.

Results Of the over 300 exposure–outcome associations tested in sPLS models, we detected 10 associations encompassing 8 outcomes. Several associations were notably consistent in direction across the three study populations: positive associations between mercury and inhibin B, and between cadmium and testosterone; and inverse associations between DiNP metabolites and testosterone, between polychlorinated biphenyl-153 and progressive sperm motility, and between a DEHP metabolite and neutral α-glucosidase, a marker of epididymal function.

Conclusions This global assessment of a mixture of environmental contaminants provides further indications that some organochlorines and phthalates adversely affect some parameters of male reproductive health.

What this paper adds

- Some ubiquitous environmental contaminants have been linked to changes in male reproductive health; however, interpretation of the mostly single-pollutant assessments is difficult given the complex mixture of daily exposures.
- In a study of around 600 fertile men from Greenland, Poland and Ukraine, we assessed associations between 15 contaminants measured in serum and 22 biomarkers of male reproductive function using sparse partial least squares regression for variable selection.
- We identified 8 perturbed biomarker outcomes, including a decrease in serum testosterone levels with increasing diisononyl phthalate levels, and confirmed a previously reported decrease in progressive sperm motility with increasing polychlorinated biphenyl-153 levels.

Linked

▶ http://dx.doi.org/10.1136/oemed-2014-102546

CrossMark

To cite: Lenters V, Portengen L, Smit LAM, *et al*. *Occup Environ Med* 2015;**72**:385–393.

INTRODUCTION

There has been considerable scientific and public concern, particularly over the past two decades, that exposure to environmental contaminants may impede development in utero, perturb hormonal homoeostasis and regulation, and contribute to subfertility in humans. This concern stems in part from evidence of geographic variability in semen quality,[1] secular trends of decreasing testosterone[2] and reports, although inconsistent and contested, of secular declines in sperm concentrations and semen quality.[3 4]

Despite the surge in studies on environmental contaminants and male reproductive health, interpretation of the evidence base is hampered by the piecemeal approach to many investigations. In observational epidemiology, often one class of compounds is correlated with a subset of outcomes, disregarding correlations among background low-level exposures. We chose to simultaneously evaluate a large set of environmental contaminant exposures and biomarkers of male reproductive function in a pregnancy-based study of men from Greenland, Poland and Ukraine.[5 6] The set of contaminants comprises high priority[7] legacy and emerging

contaminants of concern, suspected to interfere with the endocrine system: high-molecular-weight phthalates, plasticisers used in a wide variety of consumer products, including polyvinyl chloride; perfluoroalkyl acids (PFAAs), surfactants used in many applications, such as water and stain-repellent coatings; three non-essential metals; one polychlorinated biphenyl and two organochlorine pesticides. These leach from products or are deposited directly into the environment, leading to near-ubiquitous human exposure,[8 9] and all, except phthalates, are persistent and bioaccumulate.

We sought an approach to identify exposure–outcome associations between the 19 exposures, representing four classes of xenobiotic compounds, and 22 biomarker outcomes, reflecting various aspects of male reproductive function. Given the relatively weak prior information for the many possible associations, we decided to take a more agnostic, data-driven approach. That the exposures are highly correlated represented a methodological challenge. For this reason, we used sparse partial least squares (sPLS) regression modelling, which relates linear combinations of a subset of the most predictive exposures to an outcome via linear regression.[10] Unlike conventional ordinary least squares (OLS) regression, this dimension reduction approach does not suffer from instability in estimates and failure to converge due to multicollinearity, and enabled a joint assessment of the full set of measured biomarkers of low-level environmental contaminant exposures and reproductive function.

MATERIALS AND METHODS

Study populations

Study design and data collection procedures have been described previously.[6] Briefly, pregnant women and their male partners were recruited during routine antenatal care visits from 2002 through 2004 inclusive at a large central hospital in Warsaw, Poland, at three hospitals and eight antenatal clinics in Kharkiv, Ukraine, and at local hospitals in 19 municipalities and settlements across Greenland. Of the 1710 couples enrolled (45% participation rate), male partners were consecutively invited to participate in a semen study until approximately 200 men at each of the three locations had agreed (38% participation rate). Participants provided a semen sample, a venous blood sample and information on lifestyle factors (further details provided in the online supplementary methods).

Exposure and outcome assessment

Phthalate metabolites and PFAAs were analysed in serum samples by liquid chromatography-tandem mass spectrometry[9] (expanded in the online supplementary methods). Phthalates included secondary oxidised metabolites of di(2-ethylhexyl) phthalate (DEHP) [mono(2-ethyl-5-hydroxyhexyl) phthalate (MEHHP, alternatively 5OH-MEHP), mono(2-ethyl-5-oxohexyl) phthalate (MEOHP, alternatively 5oxo-MEHP) and mono(2-ethyl-5-carboxypentyl) phthalate (MECPP, alternatively 5cx-MEPP)], and of diisononyl phthalate (DiNP) [mono(4-methyl-7-hydroxyoctyl) phthalate (MHiNP, alternatively 7OH-MMeOP), mono(4-methyl-7-oxooctyl) phthalate (MOiNP, alternatively 7oxo-MMeOP) and mono (4-methyl-7-carboxyheptyl) phthalate (MOiCP, alternatively 7cx-MMeHP)]. Analysed PFAAs [a subset of per- and polyfluoroalkyl substances (PFASs), also referred to as perfluorinated compounds (PFCs)] included perfluorooctane sulfonic acid (PFOS), perfluorooctanoic acid (PFOA), perfluorhexane sulfonic acid (PFHxS), perfluorononanoic acid (PFNA), perfluorodecanoic acid (PFDA), perfluoroundecanoic acid (PFUnDA) and perfluorododecanoic acid (PFDoDA). Metals were measured in whole blood samples; cadmium (Cd) and lead (Pb) by inductively coupled plasma-mass spectrometry (Thermo X7, Thermo Elemental, Winsford, UK) and mercury (Hg) by cold vapour atom fluorescence spectrophotometry. 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB-153) and the dichlorodiphenyltrichloroethane metabolite, 1,1-dichloro-2,2-bis(*p*-chlorophenyl)-ethylene (*p,p*′-DDE), were analysed by gas chromatography-mass spectrometry,[11] as was hexachlorobenzene (HCB) according to a modified method by Otero *et al*.[12]

Twenty-two markers of reproductive function (listed in table 1) were assessed following standardised protocols[13–17]: reproductive hormones, conventional semen characteristics, markers of sperm chromatin integrity and apoptosis, markers of epididymal and accessory sex gland function, and the proportion of Y:X chromosome-bearing sperm. Analytical details are provided in the online supplementary methods and table S1.

Statistical analysis

We used sPLS regression modelling to assess associations between the exposure profiles and each outcome. In PLS regression, an X-matrix of exposures and Y outcome vector are simultaneously decomposed into latent variables and regressed in a way that maximises the covariance between X and Y.[18 19] These latent variables, called components, are orthogonal, linear combinations of the full set of input variables. Chun and Keleş[10] proposed sPLS, embedding variable selection into the PLS algorithm by eliminating 'uninformative' X-variables via penalisation. A penalty term (L_1) is applied during dimension reduction, and regression coefficients are obtained via PLS for the reduced set of X-variables (see online supplementary methods). Model complexity is a function of the number of components used to construct the model (K) and the degree of sparsity (η). To determine the optimal model (avoid overfitting), we performed Monte Carlo cross-validation,[20] and added a null (intercept only) model in the cross-validation loop to assess significance. We tested models with K values 1–5 and η values between 0.01 and 0.99 (more sparse) in steps of 0.01, using a fivefold cross-validation, repeated (with different random partitions of the data) 500 times. We selected the most parsimonious model within one SE of the overall minimum mean squared error of prediction (MSEP).[19]

We took several data pretreatment and exploration steps before undertaking sPLS analyses. We imputed exposure data below the limit of detection (LOD) (0–18%) and values missing-at-random in the matrix of exposures (2–16%). Values were imputed from a log-normal probability distribution via single conditional imputation, dependent on the population, and detected values for the other exposures. In the primary analyses, HCB, PCB-153 and *p,p*′-DDE were lipid adjusted (see online supplementary methods), and four contaminants for which <70%[21] of the samples had concentrations above the LOD were excluded: MEOHP, MOiNP, PFUnDA and PFDoDA. To approximate normality, we used natural logarithm (ln)-transformed values for all exposure variables and for 14 of the 22 outcome variables.

To address potential confounding, we employed a two-stage regression approach[22]: first, each outcome and each exposure were separately regressed on potential confounders, and second, sPLS regression models were fit inputting the residuals. We a priori selected the set of potential confounders,[9] including study population and serum cotinine level, and variably age, body mass index (BMI), abstinence period and time of blood sampling, depending on the outcome (as specified in table 1). We imputed missing covariate data (0.8–16%; as elaborated in the online supplementary methods). To assess the potential for

Table 1 Summary statistics for biomarkers of reproductive function and potential confounders among 602 male partners of pregnant women

Outcomes and covariates	n*	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†
Male reproductive hormones in serum‡§¶					
Follicle-stimulating hormone (IU/L)	456	4.38 (1.70, 9.20)	3.46 (1.40, 7.40)	3.67 (1.60, 9.40)	0.001
Luteinising hormone (IU/L)	456	4.05 (1.80, 8.50)	3.81 (1.90, 7.10)	3.79 (1.90, 7.60)	0.34
Inhibin B (ng/L)**	456	184.21 (83.00, 314.00)	157.94 (71.00, 275.00)	194.74 (98.00, 321.00)	<0.001
Sex hormone-binding globulin (nmol/L)	455	26.90 (15.60, 46.70)	21.61 (9.45, 40.90)	26.19 (15.50, 43.20)	<0.001
Total testosterone (nmol/L)**	456	14.81 (7.27, 24.41)	13.08 (6.83, 20.76)	18.01 (10.26, 26.89)	<0.001
Free testosterone (nmol/L)**	455	0.31 (0.18, 0.48)	0.29 (0.16, 0.42)	0.39 (0.22, 0.58)	<0.001
Estradiol (pmol/L)	454	63.44 (40.50, 98.70)	71.50 (45.80, 112.60)	80.73 (52.50, 137.50)	<0.001
Conventional semen characteristics					
Semen volume (mL)††,‡‡	535	3.21 (1.20, 7.20)	3.37 (1.25, 7.19)	3.29 (1.59, 7.57)	0.68
Sperm concentration (10^6/mL)††	600	51.12 (10.75, 199.00)	55.95 (7.15, 257.50)	53.07 (10.15, 185.80)	0.65
Total sperm count (10^6/ejaculate)††,‡‡	533	160.15 (31.50, 692.30)	186.68 (18.82, 1061.96)	179.46 (34.20, 804.80)	0.39
Morphologically normal (%)**	598	6.97 (2.00, 13.00)	6.66 (2.00, 14.00)	7.37 (1.00, 15.00)	0.20
Progressive motility (% A+B)**,††,§§	565	55.29 (22.00, 81.00)	60.98 (19.00, 87.00)	55.08 (14.50, 87.00)	0.007
Sperm chromatin integrity§,††					
SCSA DNA fragmentation index (%)	546	7.66 (3.32, 20.28)	10.07 (4.16, 31.50)	10.83 (3.54, 31.56)	<0.001
High DNA stainability (%)	546	11.37 (4.37, 29.88)	8.66 (3.60, 25.14)	10.29 (4.35, 26.55)	<0.001
TUNEL DNA fragmentation index (%)	462	2.95 (0.79, 13.37)	11.64 (2.57, 38.28)	6.46 (1.90, 30.18)	<0.001
Apoptotic markers§,††					
Fas positivity (%)**	454	25.00 (2.22, 70.60)	48.57 (3.13, 96.42)	28.24 (0.00, 90.60)	<0.001
Bcl-xL positivity (%)**	286	26.54 (0.44, 90.65)	18.05 (0.81, 69.64)	66.01 (9.54, 99.17)	<0.001
Epididymal and accessory sex gland function§,††,‡‡					
Neutral α-glucosidase (mU/ejaculate)	448	16.32 (4.80, 52.00)	21.81 (6.55, 71.52)	17.51 (4.86, 52.56)	0.001
Prostate-specific antigen (μg/ejaculate)	507	3187.74 (1078.00, 8330.85)	3708.59 (1312.37, 11127.89)	2497.94 (784.08, 7707.70)	<0.001
Zinc (μmol/ejaculate)	497	4.68 (1.06, 12.63)	6.31 (1.50, 25.23)	4.22 (1.11, 16.56)	<0.001
Fructose (μmol/ejaculate)	506	47.00 (12.67, 153.85)	42.24 (6.10, 161.65)	39.36 (10.69, 117.35)	0.18
Y chromosome sperm cells (%)**	411	51.26 (48.52, 55.49)	50.30 (48.79, 52.11)	50.76 (48.72, 53.61)	<0.001
Age (years)**	597	31.34 (20.83, 43.15)	30.34 (25.36, 36.80)	27.88 (20.67, 38.75)	<0.001
BMI (kg/m^2)**	592	25.99 (20.56, 31.77)	25.81 (20.43, 31.93)	24.20 (19.71, 29.41)	<0.001
Abstinence period (days)	557	2.86 (0.50, 7.10)	4.27 (1.00, 30.00)	3.41 (1.50, 7.00)	<0.001
Current smoker, n (%)	597	144 (72.7%)	53 (27.2%)	136 (66.7%)	<0.001
Current smokers: cotinine (ng/mL)	589	113.73 (0.96, 465.23)	23.95 (<LOD, 358.42)	33.59 (<LOD, 447.36)	<0.001
Time of blood sampling, <12:00, n (%)	504	17 (16.8%)	187 (94.9%)	183 (88.8%)	<0.001

*Number available out of 602.
†Test for difference in levels between the three study populations: ANOVA for means and χ^2 test for proportions.
‡In adjusted analyses, adjusted for time of blood sampling.
§In adjusted analyses, adjusted for age.
¶In adjusted analyses, adjusted for BMI.
**Outcome was not ln-transformed; arithmetic mean presented.
††In adjusted analyses, adjusted for ln-abstinence period.
‡‡Samples with spillage were excluded (n=67 for semen volume and sperm count; n=41–52 for the 4 epididymal and accessory sex gland function markers).
§§Samples with a delay of >1 h from collection to semen analysis were excluded (n=28).
BMI, body mass index; GM, geometric mean; LOD, limit of detection; P, percentile, SCSA, sperm chromatin structure assay; TUNEL, terminal deoxynucleotidyl transferase dUTP nick end-labelling.

overadjustment, in a secondary analysis we evaluated sPLS models with exposure and outcome variables only prestandardised (centred) for study population. This was considered the minimal adjustment, necessary due to the large differences in central tendencies between the populations for the exposures (figure 1), and to a lesser degree but still conspicuous, for the outcomes. Mean centring also removed baseline offsets, while scaling by log-transformation reduced the variance differences between X-variables, and thus the relative importance during dimension reduction.[23]

As an additional screening approach, we tested 330 single-pollutant, covariate-adjusted OLS linear regression models for associations between the 15 exposures (with >70% detection frequency) and 22 outcomes. To adjust for multiple comparisons, we computed the false discovery rate (FDR),[24] and set the significance threshold at an FDR <10%. In sensitivity analyses, we tested the four contaminant exposures with <70% detection frequency dichotomised as detect versus non-detect variables, and tested models excluding participants with imputed confounder and (missing-at-random) exposure data. We also assessed models with lipid-unadjusted organochlorines, including total lipids as a covariate. Finally, we tested for heterogeneity in effect estimates between study populations with an interaction term, and by assessing plots of the population-stratified regressions.

sPLS and OLS regression coefficients are presented per ln-unit increase in exposure, and for interpretability, converted into the per cent change in outcome per IQR increase in exposure; for ln-transformed outcomes, the proportional change, and for untransformed outcomes, the absolute change relative to the

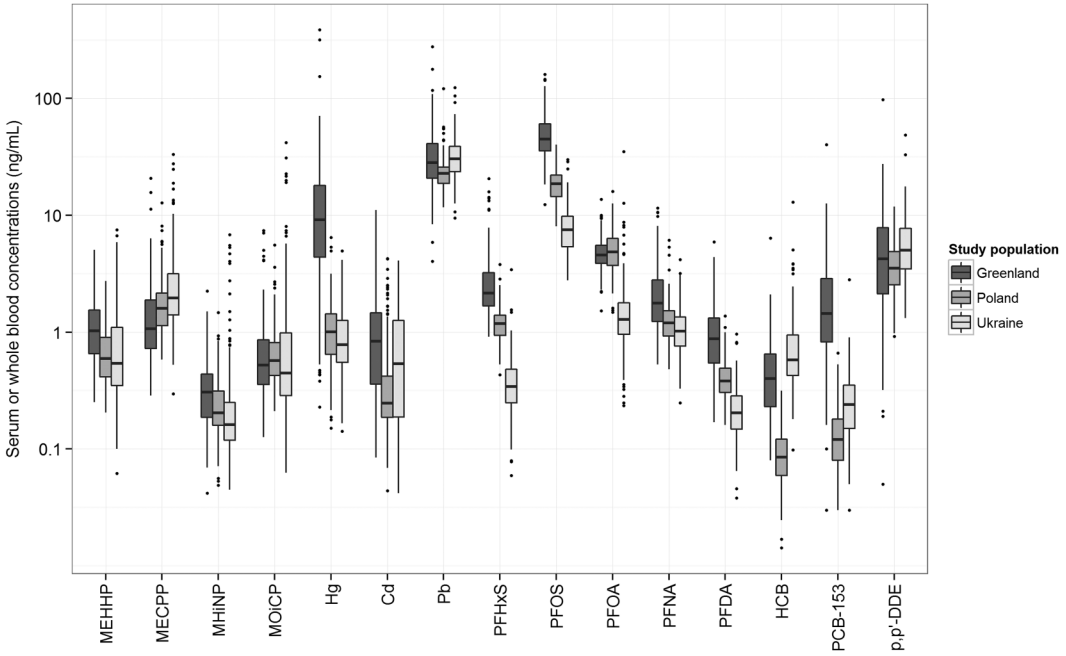

Figure 1 Box plots of the distributions of exposure biomarker concentrations per study population. MEHHP, mono(2-ethyl-5-hydroxyhexyl) phthalate; MECPP, mono(2-ethyl-5-carboxypentyl) phthalate; MHiNP, mono(4-methyl-7-hydroxyoctyl) phthalate; MOiCP, mono(4-methyl-7-carboxyheptyl) phthalate; Hg, mercury; PFHxS, perfluorhexane sulfonic acid; PFOS, perfluorooctane sulfonic acid; PFOA, perfluorooctanoic acid; PFNA, perfluorononanoic acid; PFDA, perfluorodecanoic acid; HCB, hexachlorobenzene; PCB, polychlorinated biphenyl; DDE, dichlorodiphenyldichloroethylene.

mean outcome level. Statistical analyses were performed using R V.3.0.1 (R Foundation for Statistical Computing, Vienna, Austria), and specifically the *spls* package for sPLS modelling.[25]

RESULTS

Descriptive characteristics of the study populations are shown in table 1. Participants from Kharkiv were slightly younger than their counterparts from Greenland and Warsaw, and fewer participants from Warsaw smoked. In exploring the structure of the exposure data, we observed large variations within and between the three study populations (figure 1). Levels differed across study populations for all contaminant exposures (analysis of variance p value <0.05, online supplementary table S2), except for MOiCP. Levels were higher in samples from Greenland compared to Warsaw and Kharkiv for many exposures, especially for Hg and PCB-153. Correlations within the X-matrix ranged from r_p=−0.48 to 0.91 (figure 2). We also observed distributional differences for many outcomes (table 1), although distributions for the three populations were generally more similar for the outcomes than for the exposures.

Of the 22 sPLS regression models tested in the primary analysis (n=286–600 complete cases), the X-matrix significantly improved the prediction error beyond the null model for 8 outcomes: luteinising hormone (LH), inhibin B, total testosterone, free testosterone, semen volume, progressive sperm motility, terminal deoxynucleotidyl transferase dUTP nick end-labelling assay (TUNEL) DNA fragmentation index (DFI), and neutral α-glucosidase (NAG). Within each of these significant sPLS regression models, 1 to 3 exposures within the X-matrix contributed to a total of 10 significant exposure–outcome associations (table 2). The optimal models were all K=1 component models and quite sparse (η between 0.64 and 0.99).

In the secondary analysis with exposure and outcome variables only prestandardised for study population but not for other potential confounders, sex hormone-binding globulin (SHBG) was additionally selected (the inclusion of the X-matrix improved the MSEP beyond the null model), and the LH model was no longer selected (see online supplementary table S3). Extra exposures were also selected, for a total of 24 associations.

In the covariate-adjusted OLS regression analyses, 8 overlapping exposure–outcome associations, comprising the same 8 outcomes but fewer exposures (8 of 10) compared to the primary sPLS analyses, were identified as significant (FDR <10%). Two associations, between MHiNP and total testosterone and TUNEL DFI, were significant at an FDR <5% significance threshold (see online supplementary figure S2). An additional association between MEHHP and TUNEL DFI, not selected in the sPLS model, was significant in OLS analyses (see online supplementary table S3). Consistent with the sPLS modelling results, a greater number of associations was significant in the secondary analysis with OLS models only adjusted for study population compared to the primary analysis with further adjusted models (15 compared to 9; 14 of which overlapped with those selected in sPLS secondary analysis only prestandardised for study population). No additional associations were detected in the sensitivity analysis testing contaminants with <70% detection frequency inputted as binary detect versus non-detect variables. In a reanalysis using only the complete (non-imputed) exposure and covariate data, 2 associations—*p,p′*-DDE and LH, and MHiNP and total testosterone—were selected in both the sPLS and OLS modelling (with n=321–480 and 328–586 complete cases, respectively, except for Bcl-xL positivity models which had 192–252 complete cases). This analysis had reduced power, and while CIs were wider, coefficients for the complete case analysis were remarkably similar to the primary analysis with imputed data (data not shown). Coefficients for lipid unadjusted organochlorines, adjusted for total lipids, were only slightly different in magnitude (see footnotes of table 2).

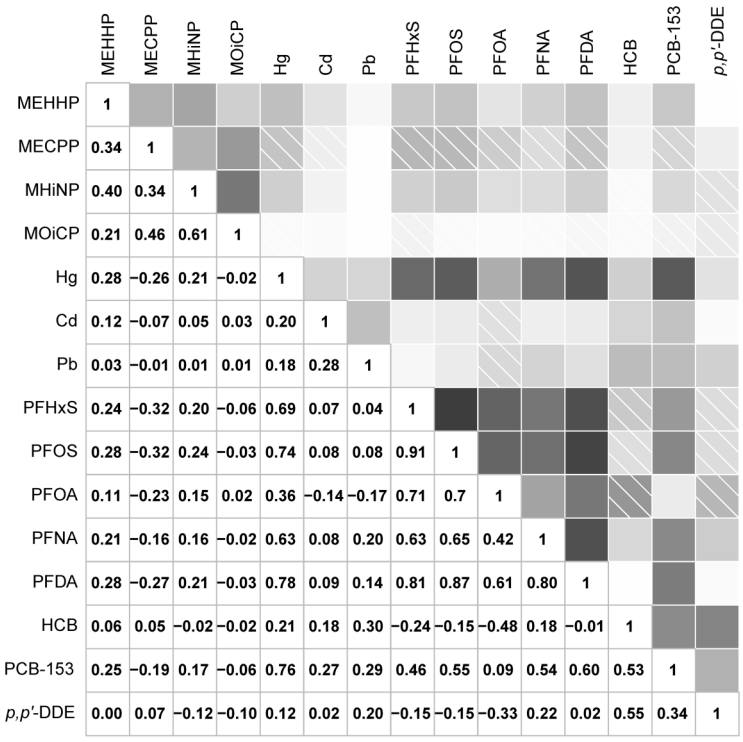

Figure 2 Pairwise Pearson correlation coefficients, also represented as a heat map, of the exposure biomarkers. MEHHP, mono (2-ethyl-5-hydroxyhexyl) phthalate; MECPP, mono(2-ethyl-5-carboxypentyl) phthalate; MHiNP, mono(4-methyl-7-hydroxyoctyl) phthalate; MOiCP, mono (4-methyl-7-carboxyheptyl) phthalate; Hg, mercury; PFHxS, perfluorohexane sulfonic acid; PFOS, perfluorooctane sulfonic acid; PFOA, perfluorooctanoic acid; PFNA, perfluorononanoic acid; PFDA, perfluorodecanoic acid; HCB, hexachlorobenzene; PCB, polychlorinated biphenyl; DDE, dichlorodiphenyldichloroethylene.

For 5 of the 10 sPLS-selected associations, there was a significant interaction between exposure and study population in the OLS sensitivity analysis (table 2). It was apparent that 8 associations were consistent in slope direction across study populations in regression models stratified by population (see online supplementary figure S3). Poland exhibited a slope in the opposite direction for LH and *p,p′*-DDE, and for the oxidative phthalates and TUNEL DFI. Associations with consistent slope directions, and their β_{sPLS} corresponding overall per cent change per IQR increase in exposure, included a 11% increase in inhibin B with Hg; a 6% decline in total testosterone with MHiNP; a 3% and 2% decline in free testosterone with DiNP metabolites, and a 5% increase with Cd; a 10% decline in semen volume with MEHHP; a 11% decline in the fraction of progressively motile sperm with PCB-153; and a 16% decline in NAG with MEHHP.

DISCUSSION

Using a multipollutant modelling approach, we identified 10 associations between biomarkers of environmental contaminant

Table 2 Exposure–outcome associations and regression coefficients selected from multipollutant sPLS regression modelling and corresponding single-pollutant OLS regression coefficients

		sPLS		OLS		
Outcome	**Exposure***	**K, η†**	**β_{sPLS}**	**β_{OLS}**	**95% CI**	**p Value**
LH* (IU/L)	*p,p′*-DDE (ng/g)	1, 0.99	0.083	0.083‡,§	(0.031 to 0.135)	0.002¶
Inhibin B (ng/L)	Hg (ng/mL)	1, 0.99	10.79	10.82	(3.90 to 17.73)	0.002¶
Total testosterone (nmol/L)	MHiNP (ng/mL)	1, 0.90	−1.14	−1.15‡	(−1.74 to −0.57)	0.0001¶
Free testosterone (nmol/L)	MHiNP (ng/mL)	1, 0.64	−0.0113	−0.019‡	(−0.032 to −0.007)	0.002¶
	MOiCP (ng/mL)		−0.0091	−0.010‡	(−0.020 to 0.000)	0.043
	Cd (ng/mL)		0.0091	0.012	(0.001 to 0.023)	0.029
Semen volume* (mL)	MEHHP (ng/mL)	1, 0.99	−0.106	−0.106	(−0.167 to −0.045)	0.0008¶
Progressive sperm (%)	PCB-153 (ng/g)	1, 0.99	−3.37	−3.37§	(−5.48 to −1.25)	0.002¶
TUNEL DFI* (%)	MHiNP (ng/mL)	1, 0.99	−0.218	−0.218‡	(−0.332 to −0.104)	0.0002¶
NAG* (mU/ejaculate)	MEHHP (ng/mL)	1, 0.77	−0.163	−0.164	(−0.255 to −0.073)	0.0005¶

Regression coefficients (β) are expressed per ln-unit change in exposure. Refer to online supplementary table S3 for per cent changes and coefficients from models only adjusted for study population. sPLS and OLS models were adjusted for study population, cotinine and for additional potential confounders as specified in table 1; sPLS models by inputting exposures and outcomes 'prestandardised' for covariates (residuals).
*Variable ln-transformed in statistical analyses.
†sPLS tuning parameters: K, the number of components used and η, the degree of sparsity.
‡Interaction p value <0.10 for the cross-product term between exposure and study population (see online supplementary figure S3 for population-stratified regression plots).
§Sensitivity analysis: 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); progressive sperm and PCB-153, −3.592 (−5.871 to −1.313).
¶Significant after adjustment for multiple comparisons in single-pollutant OLS regression modelling (FDR <10%).
DDE, dichlorodiphenyldichloroethylene; DFI, DNA fragmentation index; FDR, false discovery rate; Hg, mercury; LH, luteinising hormone; MHiNP, mono(4-methyl-7-hydroxyoctyl) phthalate; MOiCP, mono(4-methyl-7-carboxyheptyl) phthalate; MEHHP, mono(2-ethyl-5-hydroxyhexyl) phthalate; NAG, neutral α-glucosidase; OLS, ordinary least squares; PCB, polychlorinated biphenyl; sPLS, sparse partial least squares; TUNEL, terminal deoxynucleotidyl transferase dUTP nick end-labelling.

exposure and biomarkers of male reproductive function from several hundred associations tested in (sub)fertile men from Greenland, Warsaw and Kharkiv. This is the first time several of the relationships have been assessed in an epidemiological study, or simultaneously.

One of the most statistically robust and consistent associations was between DiNP metabolites and decreasing free (bioactive) testosterone and total testosterone. DEHP metabolites did not exceed the significance threshold in confounder-adjusted analyses but exhibited inverse relationships. DiNP exposure is outpacing DEHP exposure in Europe, where DiNP use is less restricted than that of DEHP. These inverse associations are consistent with an antiandrogenic effect, which has more often been observed compared to pro-androgenic effects in experimental studies.[26] Phthalates are generally considered to act by inhibiting Leydig cell synthesis of testosterone. While most experimental studies have focused on in utero exposure, DEHP and its primary metabolite MEHP were recently shown to suppress steroidogenesis of adult human testis explants in vitro, without affecting production of the hormone insulin-like factor 3 (INSL3).[27] Several epidemiological studies have found inverse associations between (urinary) MEHP and testosterone in men; in 74 occupationally exposed workers and 63 controls in China,[28] and in a pooled analysis of 425 men of infertile couples and 425 men recruited through a US infertility clinic.[29] No association was found in 234 young Swedish men sampled during routine medical conscript examinations.[30] Joensen *et al*[31] considered another exposure metric, the proportion of the summed urinary metabolites excreted as the primary metabolite, as a marker of DEHP and DiNP metabolism; %MEHP and %MiNP were associated (significant at $p<0.05$ or near-significant) with total testosterone and free androgen index in 881 healthy young men from Denmark around 19 years of age. However, no significant associations were reported for the sum of primary and secondary metabolites. Primary monoester phthalate metabolites were not measured in the current study—only secondary oxidative metabolites—due to the lipase activity of serum and risk of contamination from the diesters from the sampling devices.

We found a clear dose-dependent decrease in serum testosterone levels with increasing levels of MHiNP. We recently discovered, however, that the measurement of MHiNP was possibly confounded by an unknown co-eluting agent. Nevertheless, in an additional analysis using the molar sum of the other two DiNP metabolites versus all three oxidative metabolites tested in relation to total testosterone, the magnitude of the regression coefficients was similar: β_{OLS} −0.713 (p value=0.0037) for $\Sigma DiNP_{MOiNP+MOiCP}$ versus −0.929 (0.0007) for ΣDiNPom. Moreover, that other phthalate metabolites were significantly inversely associated with testosterone lends weight to the evidence that one or more phthalates is indeed related to decreasing testosterone levels.

We observed a positive association between cadmium and free testosterone in the sPLS analyses, although this association was not significant in the OLS analyses adjusted for potential confounders, notably cotinine. Disentangling the effect of smoking and cadmium is difficult as smoking is an important source of cadmium exposure. A cadmium–testosterone association has previously been reported in 219 men recruited from two US infertility clinics,[32] and in a representative sample of 1262 US men,[33] although this latter association did not sustain statistical significance after adjustment for smoking. An experimental study of chronic low-dose oral exposure in rats reported cadmium-related elevated testosterone[34]; however, the underlying mechanism remains unknown.

We observed decreasing NAG levels with increasing MEHHP. Sperm undergo maturation in the epididymis, and NAG is a marker of epididymal function. A smaller study (n=234) of urinary phthalates and NAG in Swedish conscripts was null.[30] In a study of 40 rats administered di-*n*-butyl phthalate, epididymal weight and NAG activity decreased dose dependently, in line with our finding.[35]

We found decreasing sperm DNA damage with increasing levels of the secondary hydroxylated metabolite of DiNP, counter to the hypothesis of a deleterious effect of exposure. However, the exposure–outcome relationship was inconsistent in direction, exhibiting a positive slope in the Polish study population. We detected associations for DNA fragmentation assessed with TUNEL, but not the sperm chromatin structure assay (SCSA). TUNEL represents DNA strand breaks measured directly in the ejaculated spermatozoa, whereas SCSA is a measure of susceptibility to acid denaturation.[36] Sperm DNA damage might be mediated by testicular oxidative stress. In rats, DEHP decreased antioxidant enzymes and concomitantly increased reactive oxygen species.[37] In 379 US men seeking infertility treatment, urinary monoethyl phthalate and MEHP were positively associated with neutral Comet assay parameters.[38] MEHP was more strongly associated when adjusted for oxidative metabolites (MEHHP and MEOHP), and similar to our finding for MEHHP, the metabolites were inversely associated with DNA damage. Hauser *et al*[38] suggested that an increased ratio of oxidative metabolites to MEHP could represent interindividual differences in rates of DEHP metabolism to 'less toxic' metabolites, or it may reflect inhibition of phase 1 enzymes involved in xenobiotic metabolism.

Some of the exposure–outcome associations tested in the current analysis have previously been tested in the same or overlapping study populations in analyses considering a single-pollutant or a single class of chemicals (refer to refs. [39–41]; previous investigations of PCB-153 and *p,p′*-DDE were reviewed by Bonde *et al*[5]). In addition to identifying several novel associations, we confirmed, using a more conservative significance threshold and joint modelling approach, several previously reported associations, including an unfavourable inverse association between PCB-153 and the proportion of progressively motile sperm.[5 17] Other smaller studies have reported comparable findings, for instance, for PCB-153 in 305 Swedish men,[42] and for PCB-138 but not PCB-153 and the odds of <50% progressive motility in 212 US men recruited at an infertility clinic (uncorrected for multiple comparisons).[43]

We also confirm a previously reported strong positive association between mercury and inhibin B.[40] As inhibin B, predominantly excreted by Sertoli cells, is correlated with sperm concentration and counts and serves as a marker of spermatogenesis,[44] this positive association is contrary to a deleterious effect of exposure. The authors[40] suggest that this association may be driven by a higher consumption of seafood—particularly relevant for Greenlandic Inuit—and the resultant concomitant exposure to mercury and ω-3 polyunsaturated fatty acids. There are indications that the latter may beneficially impact certain semen quality parameters.[45] This mercury–inhibin B association was not observed in 219 US men recruited from infertility clinics.[32]

We observed an association between MEHHP and a decline in semen volume, as previously reported for this study population.[39] A null finding for MEHP was reported for a study of

234 young Swedish men,[30] and a statistically significant increase in semen volume for the highest versus lowest %MiNP quartile in 881 Danish men was interpreted by the authors as a chance finding.[31]

We also observed a positive, yet inconsistent-across-populations, association between *p,p′*-DDE and increasing LH, as previously reported.[17] As LH regulates testosterone production by Leydig cells, this is in line with an antiandrogenic effect of *p,p′*-DDE. A null finding was reported for 341 men from a US infertility clinic with lower *p,p′*-DDE levels (geometric mean 236 ng/g lipids).[46]

We found indications of associations between phthalates, metals and organochlorines and SHBG, which did not sustain significance on adjustment for confounders, specifically BMI. These non-significant perturbations in SHBG, together with the other associations with hormones, might be indicative of effects of exposure profiles on hypothalamo–pituitary–gonadal axis regulation. Similar to the potential (over)adjustment for cotinine, adjustment for BMI may introduce overadjustment bias rather than (only) prevent confounding bias, given that some exposures may causally influence BMI.

A limitation of this assessment is the cross-sectional nature of this study, although, besides phthalates, all measured contaminants are relatively stable over time and as such were most likely reasonable proxies of recent past exposure as related to the measured outcomes. However, while blood and semen samples were collected on the same day for nearly all of the participants from Poland and Ukraine, for around 60% of the participants from Greenland, samples were collected more than 3 days apart (elaborated in the online supplementary methods), contributing to misclassification of exposure, especially for the phthalates. Analyses excluding these participants yielded generally consistent results.

Recent advances in molecular epidemiology (high-throughput screening techniques) and the introduction of the exposome concept[47] imply that epidemiological researchers will be confronted more and more with the challenging task of analysing data with high numbers of measured analytes with often complex correlation structures. Single-pollutant modelling may result in a high potential for false positives (type I errors), although multiple testing corrections and the recently proposed environment-wide association study[48] design addresses this. Using multipollutant OLS regression modelling may lead to inflated variances and flipped signs of parameter estimates, thus making it difficult to disentangle the independent effects of these variables. Methods which project data onto a lower dimensional space, like PLS regression, can better cope with collinear variables and, as a corollary, reduce the number of models tested.[10 18] While PLS-based methods are well established in chemometrics and bioinformatics, they are only recently gaining traction in genomics, metabolomics and epidemiology.[49]

In the current analysis, the associations selected using sPLS rather closely matched those we would have selected using single-pollutant OLS models. However, if we had screened for associations using multipollutant OLS models, accounting for correlations between exposures, these estimates would have suffered from multicollinearity and been unreliable (variance inflation factors exceeded 3 for 8 to 9 of the 15 exposure terms across multipollutant OLS models).

As sPLS captures the complementary contribution of X-variables, it has the power to detect multiple pollutants which might not be detected in single-pollutant modelling. This can be explained by additivity, and/or consistency at large, in that noise from measurement error and biological variation can be averaged out when a larger number of X-variables inform the latent variable. This was evident for the free testosterone model, and for several models in the sensitivity analysis with inputs only prestandardised for study population, in which up to 8 exposures were selected. Further, in constructing latent variables, sPLS regression delineates groupings in the X-matrix of exposures, which is a step forward in tackling the long-standing challenge of assessing mixture effects of contaminants.

Some aspects of (s)PLS modelling implementation are underdeveloped, such as quantifying uncertainty (stability of the selection), and explicitly correcting for multiple testing, beyond the implicit sparseness step. We estimated an empirical null distribution using a Monte Carlo approach to establish a significance threshold. A caveat of sPLS is that penalisation methods optimise prediction and will tend to select one of two highly correlated exposures with nearly equal coefficient sizes.[19] Competing multipollutant modelling approaches exist[22] (eg, tree-based models, elastic net penalised regression); however, simulation and validation assessments, specifically for data structures relevant for environmental epidemiology (often low dimensional), are lacking.

There is no straightforward way to adjust for potential confounders in (s)PLS analyses. Adjusting for confounders in sPLS models by inputting residuals in a two-step approach should, as with OLS modelling,[50] yield unbiased regression coefficients, although with a slight loss of efficiency under certain (sample size, correlation) scenarios. We observed large contrasts in exposures across the three study populations, and some differences in outcomes such as TUNEL DFI. To reduce the risk of ecological bias due to potential between-population confounding, in all analyses we prestandardised (mean centred) the exposure and outcome data by population (sPLS models) or adjusted by population (OLS models). This may have artificially reduced true contrast in exposure and outcome data between study populations, and negatively affected our power to detect some associations.

While (s)PLS modelling has some power to detect non-linear relationships (weighted over components), the PLS and OLS regression coefficients reflect linear relationships. In an exploratory analysis with generalised additive modelling, this assumption of linearity was met for most exposure–outcome relationships (<2% deviated from linearity, with an estimated degree of freedom >1.0). Further research is needed to compare the efficacy of sPLS regression with other multipollutant modelling and variable selection approaches, including non-parametric techniques.

CONCLUSIONS

We used a systematic, multipollutant modelling approach to perform a global assessment of biomarkers of contaminant exposure and male reproduction function in one of the largest studies to date. We identified several environmentally perturbed outcomes, including a robust inverse association between DiNP phthalates and circulating testosterone, and between PCB-153 and sperm motility.

Author affiliations
[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, Skåne 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]Department of 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]Department of Occupational Medicine, Danish Ramazzini Center, Aarhus University Hospital, Aarhus, Denmark
[10]Department of Occupational and Environmental Medicine, Bispebjerg Hospital, University of Copenhagen, Copenhagen, Denmark

Contributors VL, LP, DH and RV conceived and designed the current data analyses. JPB, GT, AG, LR, BAGJ, MS, HSP, JKL and LC designed and collected data for the cohort study. BAGJ and CHL performed chemical analyses. VL analysed the data and drafted the manuscript, with supervision from LP and RV. All authors, particularly LP, RV, DH, LAMS, AHP, JPB, GT, AG and MS, participated in the interpretation of results and revision of the manuscript. All authors approved the final version of the manuscript.

Funding This work was supported by the European Commission's 7th and 5th Framework Programmes (grants FP7-ENV-2008-1-226217 and QLK4-CT-2001-00202).

Competing interests None.

Patient consent Obtained.

Ethics approval Polish Bioethical Committee, Ethical Committee for Human Research in Greenland, Commission on Ethics and Bioethics of Kharkiv National Medical University.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1 Jørgensen N, Asklund C, Carlsen E, *et al*. Coordinated European investigations of semen quality: results from studies of Scandinavian young men is a matter of concern. *Int J Androl* 2006;29:54–61.
2 Travison TG, Araujo AB, O'Donnell AB, *et al*. A population-level decline in serum testosterone levels in American men. *J Clin Endocrinol Metab* 2007;92:196–202.
3 Jørgensen N, Joensen UN, Jensen TK, *et al*. Human semen quality in the new millennium: a prospective cross-sectional population-based study of 4867 men. *BMJ Open* 2012;2.
4 Carlsen E, Giwercman A, Keiding N, *et al*. Evidence for decreasing quality of semen during past 50 years. *BMJ* 1992;305:609–13.
5 Bonde JP, Toft G, Rylander L, *et al*. Fertility and markers of male reproductive function in Inuit and European populations spanning large contrasts in blood levels of persistent organochlorines. *Environ Health Perspect* 2008;116:269–77.
6 Toft G, Axmon A, Giwercman A, *et al*. Fertility in four regions spanning large contrasts in serum levels of widespread persistent organochlorines: a cross-sectional study. *Environ Health* 2005;4:26.
7 Groshart C, Okkerman PC. *Towards the establishment of a priority list of substances for further evaluation of their role in endocrine disruption–preparation of a candidate list of substances as a basis for priority setting*. European Commission DG ENV, 2000. http://ec.europa.eu/environment/archives/docum/pdf/bkh_main.pdf (accessed Apr 2014).
8 CDC (Centers for Disease Control and Prevention). Fourth National Report on Human Exposure to Environmental Chemicals. Atlanta, GA. 2009. http://www.cdc.gov/exposurereport/pdf/fourthreport.pdf (accessed Apr 2014).
9 Lindh CH, Rylander L, Toft G, *et al*. Blood serum concentrations of perfluorinated compounds in men from Greenlandic Inuit and European populations. *Chemosphere* 2012;88:1269–75.
10 Chun H, Keleş S. Sparse partial least squares regression for simultaneous dimension reduction and variable selection. *J R Stat Soc Series B Stat Methodol* 2010;72:3–25.
11 Jönsson BAG, Rylander L, Lindh C, *et al*. 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* 2005;4:27.
12 Otero R, Santiago-Silva M, Grimalt JO. Hexachlorocyclohexanes in human blood serum. *J Chromatogr A* 1997;778:87–94.
13 Elzanaty S, Rignell-Hydbom A, Jönsson BAG, *et al*. Association between exposure to persistent organohalogen pollutants and epididymal and accessory sex gland function: multicentre study in Inuit and European populations. *Reprod Toxicol* 2006;22:765–73.
14 Stronati A, Manicardi GC, Cecati M, *et al*. Relationships between sperm DNA fragmentation, sperm apoptotic markers and serum levels of CB-153 and p,p′-DDE in European and Inuit populations. *Reproduction* 2006;132:949–58.
15 Spanò M, Toft G, Hagmar L, *et al*. Exposure to PCB and p,p′-DDE in European and Inuit populations: impact on human sperm chromatin integrity. *Hum Reprod* 2005;20:3488–99.
16 Toft G, Rignell-Hydbom A, Tyrkiel E, *et al*. Semen quality and exposure to persistent organochlorine pollutants. *Epidemiology* 2006;17:450–8.
17 Giwercman AH, Rignell-Hydbom A, Toft G, *et al*. Reproductive hormone levels in men exposed to persistent organohalogen pollutants: a study of inuit and three European cohorts. *Environ Health Perspect* 2006;114:1348–53.
18 Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab* 2001;58:109–30.
19 Hastie TJ, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction*. 2nd ed. New York, NY: Springer, 2009.
20 Xu Q, Liang Y, Du Y. Monte Carlo cross-validation for selecting a model and estimating the prediction error in multivariate calibration. *J Chemom* 2004;18:112–20.
21 Lubin JH, Colt JS, Camann D, *et al*. Epidemiologic evaluation of measurement data in the presence of detection limits. *Environ Health Perspect* 2004;112:1691–6.
22 Chadeau-Hyam M, Campanella G, Jombart T, *et al*. Deciphering the complex: methodological overview of statistical models to derive OMICS-based biomarkers. *Environ Mol Mutagen* 2013;54:542–57.
23 Van den Berg RA, Hoefsloot HCJ, Westerhuis JA, *et al*. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* 2006;7:142.
24 Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B* 1995;57:289–300.
25 Chung D, Chun H, Keles S. spls: Sparse Partial Least Squares (SPLS) Regression and Classification. R Packag version 22-1. 2013. http://cran.r-project.org/package=spls (accessed Jan 2014).
26 Albert O, Jégou B. A critical assessment of the endocrine susceptibility of the human testis to phthalates from fetal life to adulthood. *Hum Reprod Update* 2014;20:231–49.
27 Desdoits-Lethimonier C, Albert O, Le Bizec B, *et al*. Human testis steroidogenesis is inhibited by phthalates. *Hum Reprod* 2012;27:1451–9.
28 Pan G, Hanaoka T, Yoshimura M, *et al*. Decreased serum free testosterone in workers exposed to high levels of di-n-butyl phthalate (DBP) and di-2-ethylhexyl phthalate (DEHP): a cross-sectional study in China. *Environ Health Perspect* 2006;114:1643–8.
29 Mendiola J, Meeker JD, Jørgensen N, *et al*. Urinary concentrations of di (2-ethylhexyl) phthalate metabolites and serum reproductive hormones: pooled analysis of fertile and infertile men. *J Androl* 2012;33:488–98.
30 Jönsson BAG, Richthoff J, Rylander L, *et al*. Urinary phthalate metabolites and biomarkers of reproductive function in young men. *Epidemiology* 2005;16:487–93.
31 Joensen UN, Frederiksen H, Jensen MB, *et al*. Phthalate excretion pattern and testicular function: a study of 881 healthy Danish men. *Environ Health Perspect* 2012;120:1397–403.
32 Meeker JD, Rossano MG, Protas B, *et al*. Environmental exposure to metals and male reproductive hormones: circulating testosterone is inversely associated with blood molybdenum. *Fertil Steril* 2010;93:130–40.
33 Menke A, Guallar E, Shiels MS, *et al*. The association of urinary cadmium with sex steroid hormone concentrations in a general population sample of US adult men. *BMC Public Health* 2008;8:72.
34 Zeng X, Jin T, Zhou Y, *et al*. Changes of serum sex hormone levels and MT mRNA expression in rats orally exposed to cadmium. *Toxicology* 2003;186:109–18.
35 Zhou D, Wang H, Zhang J. Di-n-butyl phthalate (DBP) exposure induces oxidative stress in epididymis of adult rats. *Toxicol Ind Health* 2011;27:65–71.
36 Evenson DP, Larson KL, Jost LK. Sperm chromatin structure assay: its clinical use for detecting sperm DNA fragmentation in male infertility and comparisons with other techniques. *J Androl* 2002;23:25–43.
37 Kasahara E, Sato EF, Miyoshi M, *et al*. Role of oxidative stress in germ cell apoptosis induced by di(2-ethylhexyl)phthalate. *Biochem J* 2002;365:849–56.
38 Hauser R, Meeker JD, Singh NP, *et al*. DNA damage in human sperm is related to urinary levels of phthalate monoester and oxidative metabolites. *Hum Reprod* 2007;22:688–95.
39 Specht IO, Toft G, Hougaard KS, *et al*. Associations between serum phthalates and biomarkers of reproductive function in 589 adult men. *Environ Int* 2014;66:146–56.
40 Mocevic E, Specht IO, Marott JL, *et al*. Environmental mercury exposure, semen quality and reproductive hormones in Greenlandic Inuit and European men: a cross-sectional study. *Asian J Androl* 2013;15:97–104.
41 Toft G, Jönsson BAG, Lindh CH, *et al*. Exposure to perfluorinated compounds and human semen quality in arctic and European populations. *Hum Reprod* 2012;27:2532–40.
42 Richthoff J, Rylander L, Jönsson BAG, *et al*. Serum levels of 2,2′,4,4′,5,5′-hexachlorobiphenyl (CB-153) in relation to markers of reproductive function in young males from the general Swedish population. *Environ Health Perspect* 2003;111:409–13.
43 Hauser R, Chen Z, Pothier L, *et al*. The relationship between human semen parameters and environmental exposure to polychlorinated biphenyls and p,p′-DDE. *Environ Health Perspect* 2003;111:1505–11.

44 Jørgensen N, Liu F, Andersson A-M, *et al*. Serum inhibin-b in fertile men is strongly correlated with low but not high sperm counts: a coordinated study of 1,797 European and US men. *Fertil Steril* 2010;94:2128–34.
45 Safarinejad MR. Effect of omega-3 polyunsaturated fatty acid supplementation on semen profile and enzymatic anti-oxidant capacity of seminal plasma in infertile men with idiopathic oligoasthenoteratospermia: a double-blind, placebo-controlled, randomised study. *Andrologia* 2011;43:38–47.
46 Ferguson KK, Hauser R, Altshul L, *et al*. Serum concentrations of p, p′-DDE, HCB, PCBs and reproductive hormones among men of reproductive age. *Reprod Toxicol* 2012;34:429–35.
47 Wild CP. Complementing the genome with an "exposome": the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005;14:1847–50.
48 Patel CJ, Bhattacharya J, Butte AJ. An Environment-Wide Association Study (EWAS) on type 2 diabetes mellitus. *PLoS ONE* 2010;5:e10746.
49 Holmes E, Loo RL, Stamler J, *et al*. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 2008;453:396–400.
50 Xing G, Lin C-Y, Xing C. A comparison of approaches to control for confounding factors by regression models. *Hum Hered* 2011;72:194–205.

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.

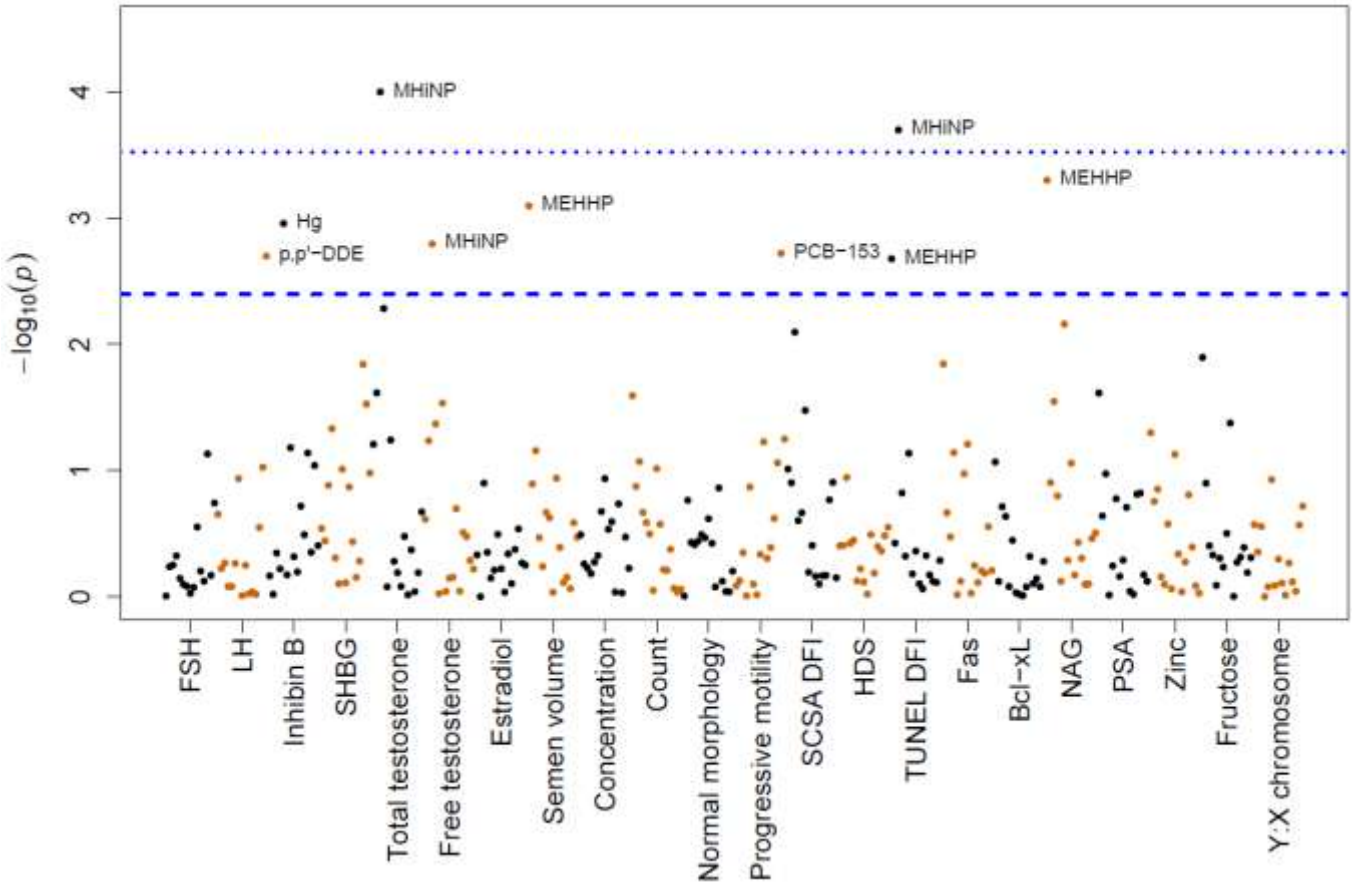

Figure S2. The p-values ($-\log_{10}$ scale) from single exposure-outcome OLS regression analyses, plotted per outcome.

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.

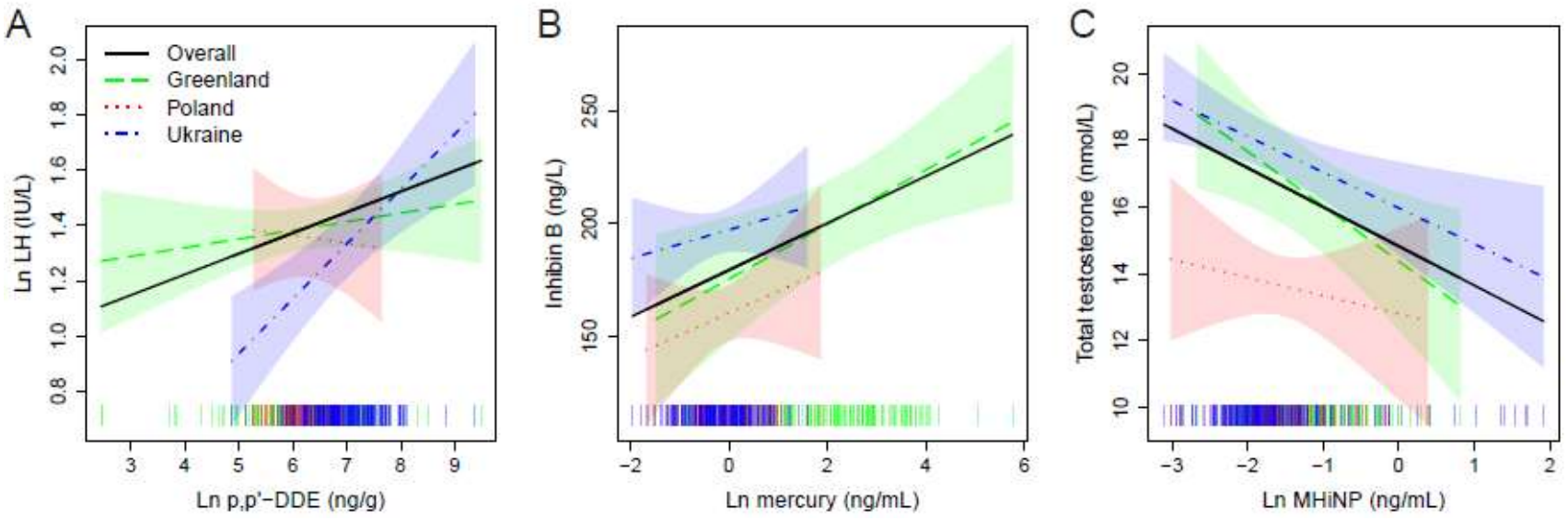

Figure S3. Exposure-outcome associations, plotted as linear regressions across all and stratified by study population (A-L).

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.

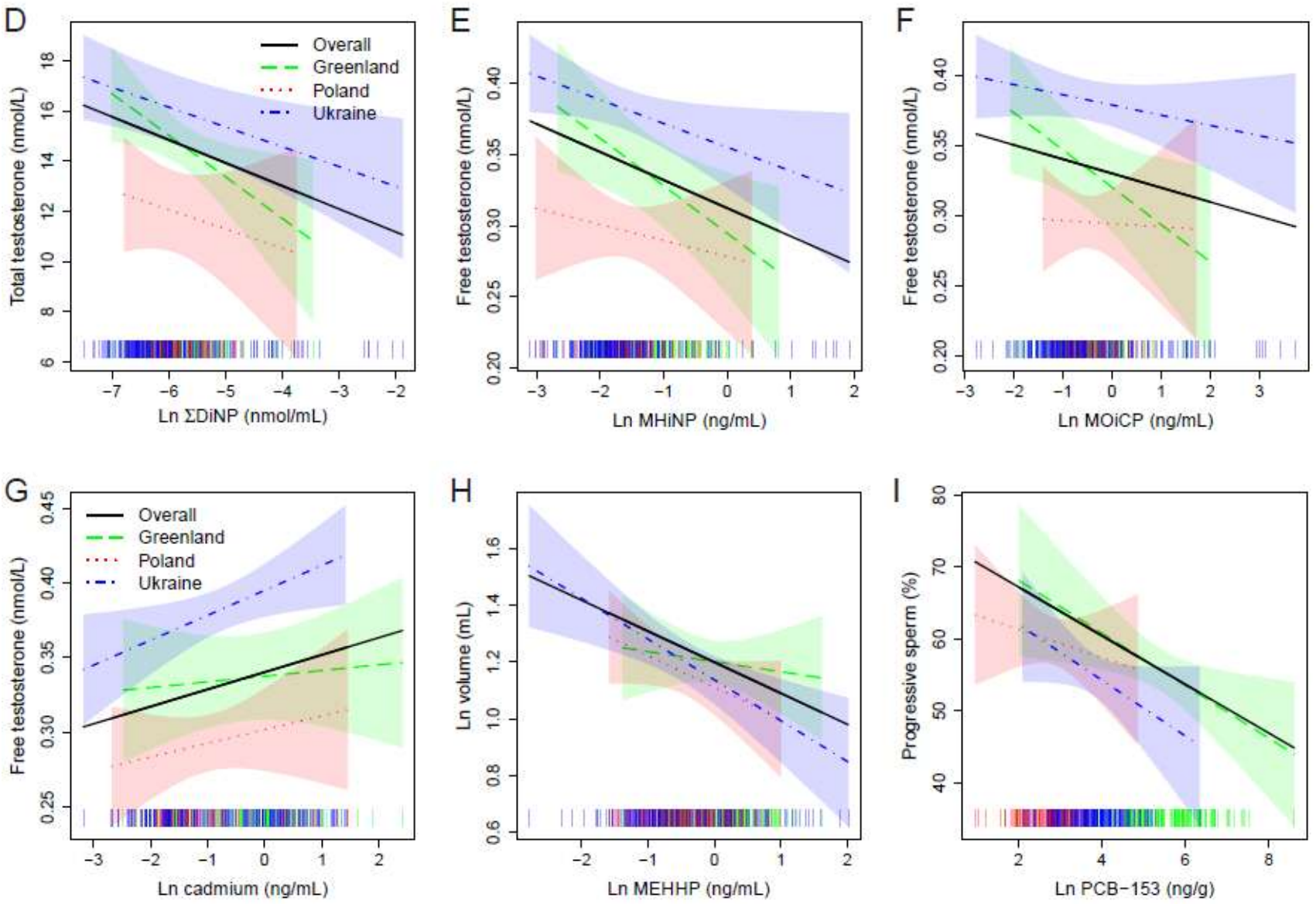

Figure S3. *Continued*

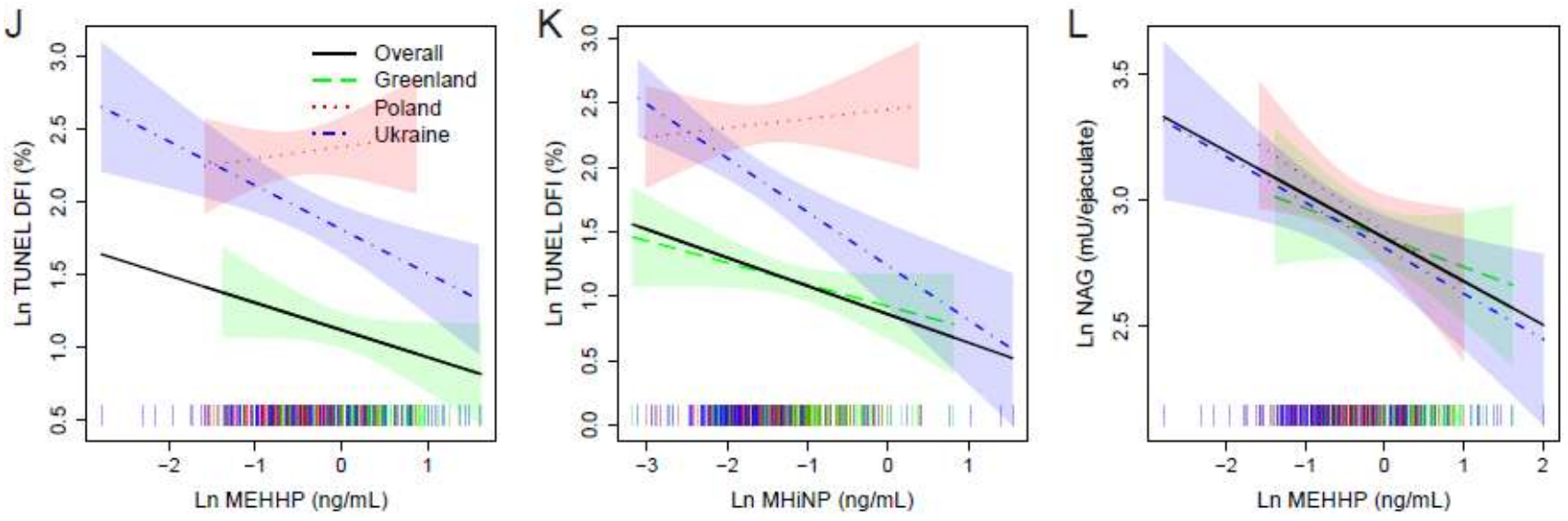

Figure S3. *Continued*

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.