

Exposome Profiling of Environmental Pollutants in Seminal Plasma and Novel Associations with Semen Parameters

Haotian Wu,^{*,#} Vrinda Kalia,[#] Katherine E. Manz, Lawrence Chillrud, Nathalie Hoffmann Dishon, Gabriela L. Jackson, Christian K. Dye, Raoul Orvieto, Adva Aizer, Hagai Levine, Marianthi-Anna Kioumourtzoglou, Kurt D. Pennell, Andrea A. Baccarelli, and Ronit Machtinger



Cite This: *Environ. Sci. Technol.* 2024, 58, 13594–13604



Read Online

ACCESS |



Metrics & More



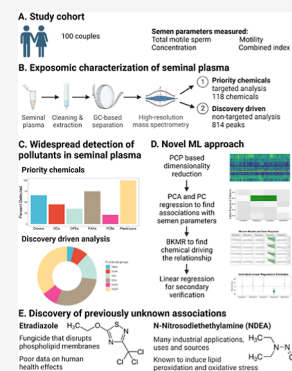
Article Recommendations



Supporting Information

ABSTRACT: Indicators of male fertility are in decline globally, but the underlying causes, including the role of environmental exposures, are unclear. This study aimed to examine organic chemical pollutants in seminal plasma, including both known priority environmental chemicals and less studied chemicals, to identify uncharacterized male reproductive environmental toxicants. Semen samples were collected from 100 individuals and assessed for sperm concentration, percent motility, and total motile sperm. Targeted and nontargeted organic pollutant exposures were measured from seminal plasma using gas chromatography, which showed widespread detection of organic pollutants in seminal plasma across all exposure classes. We used principal component pursuit (PCP) on our targeted panel and derived one component (driven by etridiazole) associated with total motile sperm ($p < 0.001$) and concentration ($p = 0.03$). This was confirmed by the exposome-wide association models using individual chemicals, where etridiazole was negatively associated with total motile sperm (FDR $q = 0.01$) and concentration ($q = 0.07$). Using PCP on 814 nontargeted spectral peaks identified a component that was associated with total motile sperm ($p = 0.001$). Bayesian kernel machine regression identified one principal driver of this association, which was analytically confirmed to be *N*-nitrosodiethylamine. These findings are promising and consistent with experimental evidence showing that etridiazole and *N*-nitrosodiethylamine may be reproductive toxicants.

KEYWORDS: *exposome, seminal plasma, fertility, etridiazole, NDEA, nitrosamine, sperm, pesticides*



INTRODUCTION

Infertility, defined as the inability to achieve a pregnancy after ≥ 12 months of trying, is a global public health concern that affects 15% of all reproductive age couples in the world.^{1,2} Infertility has significant negative physical, emotional, and social impacts on the affected individuals; addressing infertility will help realize the fundamental right for individuals and couples to have children and to improve the health of the affected individuals. Male infertility contributes to 40–50% of overall cases.³ While changes in gross rates of infertility have not been observed,² many alarming trends have been observed in male fertility, including notable declines in testosterone and sperm count, with concurrent rises in male genital anomalies and testicular cancer.⁴ Recent data show that sperm count declined by 51.6% between 1973 and 2018 and this decline has accelerated since 2000.⁵ Thus, if this trend persists, a greater proportion of the male population could drop below the threshold^{6,7} where lower sperm count could lead to a rise in male infertility, reflecting not only difficulty in conceiving but also a signal of poorer general health.

The cause of this decline is likely multifactorial, including both prenatal and adult factors,^{8–10} with environmental factors playing a major role.⁴ However, our current knowledge of the determinants of male fertility does not sufficiently explain this

constant and global decline. Over the past decades, several environmental pollutants have been found to be male reproductive toxicants, such as airborne pollutants, certain types of persistent organic pollutants, and chemicals commonly found in plastics. However, these known environmental reproductive toxicants might not be the primary drivers of this alarming trend because the overall trends of these specific pollutants are declining across the same time frame as the observed decline in sperm count and many observed effect sizes are relatively small. Thus, it is likely that there are other unidentified reproductive toxicants or uncharacterized interactions between chemicals that may impact male fertility.

The emerging field of exposomics offers a comprehensive approach for environmental health research. The concept of the “exposome” was first introduced by C.P. Wild in 2005 as a way to represent the totality of environmental (i.e., nongenetic) drivers

Received: December 7, 2023

Revised: July 11, 2024

Accepted: July 12, 2024

Published: July 25, 2024



of health and disease.¹¹ The exposome is a function of external forces and internal biological processes.¹² In practice, exposome studies attempt to capture a large set of environmental exposures simultaneously and have the potential to address limitations of existing studies by investigating the impact of “real-life” exposures and their combinations and interactions between exposures.

Historically, there have been several major challenges to identifying previously unknown reproductive toxicants. First, populations around the world are exposed to thousands of chemicals and other environmental exposures with large temporal and geospatial variations where most chemicals are introduced without rigorous reproductive toxicity data. This presents an obvious biomonitoring challenge as it is difficult to assess thousands of toxicant exposures and their metabolites simultaneously, particularly chemicals without reproductive toxicity data or those that have never been profiled in humans. Second, not every environmental pollutant has accessible assays, creating additional barriers to investigators and biomonitoring efforts. For pollutants that can be measured, the analytical platforms used are usually highly specific and lack the ability for high-throughput simultaneous assessment of many exposures from a reasonable amount of biospecimen. This restricts investigators to only a small subset of putative or known toxicants with accessible assays, impeding efforts for nontargeted discovery. This is compounded by a third challenge; human exposures are typically measured using available and minimally invasive biospecimen such as urine, blood, and hair. This can create exposure misclassification and spurious findings as male reproductive organs and processes are protected by the blood-testis-barrier, which can lead to different exposure profiles between systemic circulation and the male reproductive system.^{13,14} Fourth, the standard environment-wide (i.e., exposome) association studies (ExWAS) approach can be limited by multiple testing penalties, leading to high sample size requirements.¹⁵ While ExWAS remains useful, additional strategies are necessary to analyze high-dimensional exposure data. Together, these constraints collectively hinder investigators from more efficiently and comprehensively identifying environmental pollutant influences on male reproductive health.

The objectives of this study were to (1) profile a large set of organic pollutants relevant to male reproductive tissues using a targeted approach and a discovery-driven nontargeted analysis (NTA) and (2) assess associations between the detected pollutants and percent motility, concentration, and total motile sperm to discover previously uncharacterized male reproductive environmental toxicants. We leveraged state-of-the-art analytical advances in high-resolution mass spectrometry¹⁶ to simultaneously profile thousands of potential environmental chemicals in seminal plasma, which is more proximal and relevant for male reproductive health compared to measures of chemicals in systemic circulation.^{17,18} We then combined a novel machine learning pattern recognition approach, principal component pursuit (PCP),^{19,20} with modern statistical mixtures analyses²¹ to efficiently detect associations of environmental chemicals with male reproductive health. Typical studies model one feature (e.g., genetic polymorphism or environmental exposure) at a time,²² repeated through all features, which incurs severe multiple testing penalties on statistical power. In addition, environmental exposures do not occur in isolation as represented in these models. In real life, individuals are subjected to complex sets of environmental exposures that jointly act to induce endogenous responses. Our novel approach

helps address both limitations of traditional statistical approaches by removing noise, reducing the number of statistical comparisons, and allowing for discovery of interactions and exposure patterns.

METHODS

Cohort Recruitment and Clinical Data Collection. From 2020 to 2021, 158 couples undergoing a fresh IVF cycle at Sheba Medical Center, a tertiary medical center in Tel HaShomer, Israel, were approached, of which 100 consented. Couples were enrolled during ovarian stimulation and those with severe oligoasthenoteratozoospermia were excluded at the time of enrollment. Demographic, lifestyle, anthropometric, and medical history data were collected by trained clinic personnel. Standard semen parameters, including volume, percent motility, and total sperm count data were collected as part of routine IVF protocol following WHO protocol on semen sample treatment and semen parameter assessment.²³ A composite semen index was derived using principal component analysis (PCA).

Semen samples were collected in a sterile plastic specimen cup after a recommended 2–7 day abstinence period, per standard IVF protocol. Following liquefaction, sperm concentration and motility were evaluated by Makler counting cell (Sefi-Makler). The liquefied semen was gently added to the density gradient medium (80%/40% PureCeption, Quinn's, SAGE) and centrifuged at 600g for 20 min to separate the seminal plasma. Samples were stored at $-80\text{ }^{\circ}\text{C}$, defrosted in a $4\text{ }^{\circ}\text{C}$ refrigerator, and then brought to room temperature prior to extraction for NTA.

High Resolution Non-Targeted Exposome Measurement. Extensive details of the targeted and nontargeted high resolution platform are provided in the [Supporting Information](#), and are based on previously described methods.¹⁶ In brief, seminal plasma was extracted using a modified QuEChERS extraction. A 500 μL volume of seminal plasma was mixed with 500 μL hexane/acetone:dichloromethane, vortexed for 30 s, and transferred to a QuEChERS tube (150 mg MgSO_4 + 50 mg C18). The tubes were centrifuged and the supernatant was transferred to a glass vial for collection. This was repeated two more times with 250 μL hexane/acetone/dichloromethane. The final extract was reduced to 150 μL under nitrogen, transferred to an amber autosampler vial, spiked with 10 μL of an internal standard solution containing 62.5 $\mu\text{g/L}$ of phenanthrene D-10 and chrysene D-12 (AccuStandard) and with 10 μL of diluted retention time marker (AccuStandard DRH-TX-003-CNM), and sealed with a cap.

The extracts were analyzed using a high-resolution Thermo Q Exactive Orbitrap MS equipped with a Thermo Trace 1300 GC and a TriPlus RSH Autosampler. Limits of detection were determined by injecting a calibration standard seven times, and calculated as $[s \times t(df, 1 - \alpha = 0.99)]/m$ where s is the standard deviation, t is the student's t -value, df is the degrees freedom, α is the significance level ($n = 7$, $\alpha = 0.01$, $t = 3.14$), and m is the slope of the calibration curve. Full details of the analytical methods, including the QC, analytical sequence, chromatography and mass spectrometry settings, and data processing, are provided in the [Supporting Methods](#). [Table S1](#) shows the recovery spikes of the contaminant mixture in fetal bovine serum. [Table S2](#) shows the extraction recovery of the contaminants from NIST SRM.

For all statistical analyses, we restricted the analysis of targeted measurements to chemicals with $\geq 40\%$ detection rate, resulting in 43 chemicals. Any measurements below the limit of detection

were replaced with half the value of limit of quantitation or the minimum detected, whichever was smaller. The nontargeted chemical data analysis was restricted to peaks that were detected in all samples, with a minimum abundance >50,000 in order to ensure peak quality, resulting in 814 features for analysis. An overview of the statistical approach is shown in Figure S1.

Chemical Identification and Categorization. Chemicals measured via the targeted approach were categorized based on entries for each chemical on the EPA CompTox Chemicals Dashboard.²⁴ They were broadly categorized as PCB ($n = 46$), PAH ($n = 11$), pesticides were subclassified as chlorobenzene ($n = 7$), organochlorine ($n = 18$), triazine and triazole ($n = 4$), organophosphate ($n = 6$), and other pesticides ($n = 8$), and all other chemicals were categorized as “other” ($n = 18$). Further, we examined the correlation between detection and physicochemical properties obtained from the CompTox Dashboard, namely ExpoCast predicted median exposure, predicted octanol–water partition coefficient, predicted bioconcentration factor, and predicted biodegradation half-life. To characterize the nontargeted data, we used chemoinformatic tools to generate simplified molecular-input line-entry system notation by querying the chemical identifier resolver using chemical names, annotated by compound discoverer, using the webChem R package²⁵ (version 1.1.3). We determined the elemental composition and functional groups present using the rcdk package.²⁶

For the NTA, features were detected with 10 ppm mass tolerance, 10,000 total ion chromatogram threshold, signal-to-noise ratio of 3, and 99% allowable ion overlap. Each chromatogram was retention time-aligned using the carbon distribution marker (contains 9 alkanes; only compounds containing greater than 8 carbons were used since the compounds smaller than this eluted during the solvent delay) spiked into each sample and retention indices (RIs) were calculated for each peak detected.

For suspect screening analysis, the RI of each peak was used to limit suspects during identification; the allowed maximum RI difference was 300. Compounds were identified by searching their mass spectra in the NIST Mass Spectra Library (NIST/EPA/NIH EI and NIST Tandem Mass Spectral Library Version 2.3) and a high-resolution library developed in-house using certified standards containing 354 unique compounds. A minimum match factor and reverse match factor score of 500 was used for assigning library matches. Peaks with scores <500 were not assigned an identification. Chemicals that matched to our in-house library were assigned level 1 annotation if they were also detected in our standard mixture, based on previously described scoring scheme.²⁷

Covariate Selection. For all models, we decided to include age, body mass index (BMI), smoking and cannabis use, and infertility diagnoses (male, female, unexplained) as covariates. These were selected due to their known association with male fertility. Education and whether someone lived or worked near agriculture were not included in the model as it was not associated with any semen parameter outcomes in our population.

Individual Exposure Models. For the targeted organic pollutants, we applied an Exposome Wide Association Study (ExWAS) approach where we tested each chemical individually with the semen parameters as outcomes, analogous to GWAS studies.²⁸ Using linear regression models, chemicals with 40–70% detection rate were modeled as binary exposures (above and below detection) and those with >70% detection rate were

modeled as continuous variables. This was done to balance the statistical power gained by modeling variables as continuous when possible and the overall model fit as variables with high nondetects may be less appropriate for linear regression. Models were adjusted for age, BMI, smoking and cannabis use, and infertility diagnoses (male, female, unexplained). The model estimates were corrected for false discovery rate.²⁹ As a sensitivity analysis, we excluded 13 individuals who were diagnosed with male factor infertility and/or had total sperm count <39 million, which corresponds to fifth percentile of men whose partners became pregnant within 12 months.²³ Varying this cutoff made no material difference to our model estimates.

For the 814 highly abundant nontargeted features, we used linear regression as a verification of our PCP–PCA–BKMR results (see below) because potential bias introduced by coexposure adjustment may lead to spurious findings in statistical mixture models under specific scenarios.^{30,31} As all features were in arbitrary units that scale with concentration, we decided to model the feature variables after log 2-transformation to allow the resulting model estimates to be interpreted as the difference in outcome per doubling of exposure. This analysis was restricted to exposures identified by the PCP–PCA–BKMR pipeline. Models were adjusted for age, BMI, smoking and cannabis use, and infertility diagnoses.

Dimension Reduction via Unsupervised Machine Learning Pattern Recognition. PCP is a robust method for dimensionality reduction and pattern recognition.^{32,33} Its theory and application to environmental health research has been described previously.^{19,20} In brief, we can consider any set of exposure data to comprise two underlying matrices—the low-rank L-matrix and a sparse S-matrix. The L-matrix represents the latent underlying patterns that can be estimated but are not directly observed. The S-matrix represents the unusual or uniquely rare values that cannot be explained by the latent patterns observed by the L-matrix. By partitioning the observed data with PCP, we can effectively separate the underlying latent patterns from the unusual and rare events.

PCP offers two unique advantages. First, by removing the influence of unusual and rare events, we can remove outliers and statistical noise that are often products of random variability. Second, while the resulting L-matrix has the same dimensions as the original exposure matrix, it will have a low rank, effectively reducing the dimensionality of the data. Any subsequent matrix factorization techniques, such as PCA, will identify only a few factors/components that explain a large proportion of variability. This stands in contrast with PCA on typical observed—and noisy—exposure data matrices (i.e., raw exposure data) where a large number of factors/components are required to explain the majority of the variability in the data.

For our presented analysis, we used non-negative nonconvex square root PCP with a noise-independent universal choice of regularization parameters.^{19,34} This approach requires estimation of two parameters, λ and μ . We tested both default universal parameters ($\lambda = 1/\sqrt{n}$ and $\mu = p/\sqrt{2}$ where n is sample size and p is number of features)^{32,34} as well as parameter values obtained via a grid search. In each case, the resulting L matrices were very similar both in rank and subsequent regression results. Thus, we chose to present the results using the default universal parameters which have been shown to have theoretical guarantees.

PCP was applied separately for the targeted chemicals and on the nontargeted spectra. The input for the targeted chemicals included 43 pollutants with $\geq 40\%$ detection rate. The input for

the nontargeted chemicals included 814 high abundance features that were detected in all samples. All inputs were scaled to 0–1 for PCP.

Principal Component (PC) Regression. We applied PCA on the L-matrices and for downstream regression analyses, we chose the top 5 PCs from the priority common organic pollutants (99.1% cumulative variance) and the top 6 PCs from the nontargeted spectra (99.4% cumulative variance). We extracted the eigenvectors as well as the loadings of each chemical in the data set. Linear regression was then used to model each PC as the predictor and semen parameters (total motile sperm, concentration, and percent motility) as outcomes. All models were adjusted for age, BMI, smoking and cannabis use, and infertility diagnoses (male, female, unexplained). For PCs that were found to be associated with outcomes, we identified the top chemicals or features with the highest loadings as potential drivers and used this to inform our mixtures models.

Bayesian Kernel Machine Regression. BKMR²¹ is a flexible statistical mixtures method that can model multiple exposures to estimate the dose–response relationship of each mixture component while holding other components constant. Bayesian kernel machine regression (BKMR) makes no assumptions regarding the shape of the relationships and can be used to capture nonlinear curves and statistical interactions between the mixture components. For all BKMR models, we scaled observed exposures, outcomes, and continuous covariates, and specified 50,000 iterations with default priors. All models were adjusted for age, BMI, smoking and cannabis use, and infertility diagnoses.

From PC regression, we identified candidate chemicals or features that may be associated with semen parameters. We then modeled the observed continuous exposure values (i.e., observed concentrations) of these chemicals or features in BKMR models with our outcomes. We reasoned that real associations should be detectable using the observed values and that using the latent representations (i.e., L-matrix) may result in spurious findings induced by the data handling procedure. Furthermore, models using observed exposure values are more easily verified and replicated by others. Also, to ensure that our observations regarding individual chemicals were not products of arbitrary decisions for the number of mixture components, we tested BKMR models with top 5, 10, and 15 chemicals or features to ensure that the estimated dose–response relationships were a faithful reflection of the underlying data.

Software and Packages. All analyses were conducted in R v4.0.1. We used pcpr package (<https://github.com/Columbia-PRIME/pcpr>) for PCP and bkmr package for BKMR.

RESULTS

Population characteristics are shown in Table 1. Among the male partners of the 100 couples seeking assisted in vitro fertilization, five reported use of cannabis and 19 reported current smoking. Their mean age was 38.1 (SD = 5.7) years old and mean BMI was 26.3 (SD = 4.6). A substantial portion of the study population lived or worked near agricultural sites. The population semen parameter was similar to the WHO population parameters²³ and the majority of participants had normal volume and concentration according to WHO standards. Only 9 couples sought IVF treatment due to male-factor infertility. Our study design, methods, and key findings are shown in (Graphical Abstract).

Targeted Profiling of the Seminal Plasma Exposome and Exposure Patterns. Starting with 119 commonly

Table 1. Study Population Characteristics (N = 100)

demographics	mean (SD) or N (%)
age (years)	38.09 (5.65)
BMI (kg/m ²)	26.28 (4.59)
Cannabis use	5 (5.0%)
smoking	19 (19.0%)
live or work near agriculture ^a	
no	67 (68.4%)
yes	31 (31.6%)
education (years) ^b	14.6 (2.88)
clinical characteristics	
infertility diagnoses (couple)	mean (SD) or N (%)
endometriosis	12 (12.0)
fertility preservation	9 (9.0)
genetic	42 (42.0)
male factor	9 (9.0)
mechanic	5 (5.0)
ovarian	11 (11.0)
ovulatory	4 (4.0)
unexplained/idiopathic	8 (8.0)
volume (mL)	3.34 (1.70)
motility (%)	55.31 (21.96)
motile concentration (million/mL)	76.19 (53.63)
total motile sperm (million)	146.4 (142.2)

^aMissing data on 2 individuals. ^bRange is 8–21 years of formal education. Median is 15 years.

measured persistent or trace organic pollutants (all with level 1 identification confidence, i.e. confirmed chemical structures), we found that nearly all tested organic pollutant exposures were detectable in at least one sample, including all measured classes of pesticides, polyaromatic hydrocarbons (PAH), polychlorinated biphenyl (PCBs), and other common environmental contaminants such as dioxins and dioxin-like compounds, phthalates, and solvents (Figure 1). Cyanazine was the only chemical not detected in any of the samples. We observed predominantly positive correlations among the targeted chemicals, which is consistent with the fact that we expect chemical exposures to arise from complex chemical mixtures and not be mutually exclusive. Summary statistics for individual compounds can be found in Table S3.

As expected we found a positive correlation between detection rate and predicted median exposure and ($\rho = 0.57, p < 0.001$). We found negative correlations between detection rate and predicted octanol–water partition coefficient ($\rho = -0.56, p < 0.001$), predicted bioconcentration factor ($\rho = -0.52, p < 0.001$), and predicted biodegradation half-life ($\rho = -0.19, p < 0.05$).

Several specific exposure patterns can also be observed in our data (Figure 1). There were expected correlations between different isomers and related metabolites such as lindane (γ -hexachlorohexane [HCH]), α -lindane, δ -lindane, and β -HCH. There was also a cluster that comprised numerous pesticides, such as triadimefon, *N,N*-diethyl-3-methylbenzamide (DEET), lindane (and its related isomers and metabolites), 2,4,5-trichlorophenol, etridiazole, 1,4-dichlorobenzene, and chlopyrifos, and reagents that are commonly used for production of pesticides, including diethyl phthalate and isosafrole. PCB-49 is also part of this cluster of correlated chemicals, which may reflect the usage of PCBs in certain pesticides similar to diethyl phthalate and isosafrole. Other specific coexposure patterns and exposure profiles were observed among dibenzofuran, safrole,

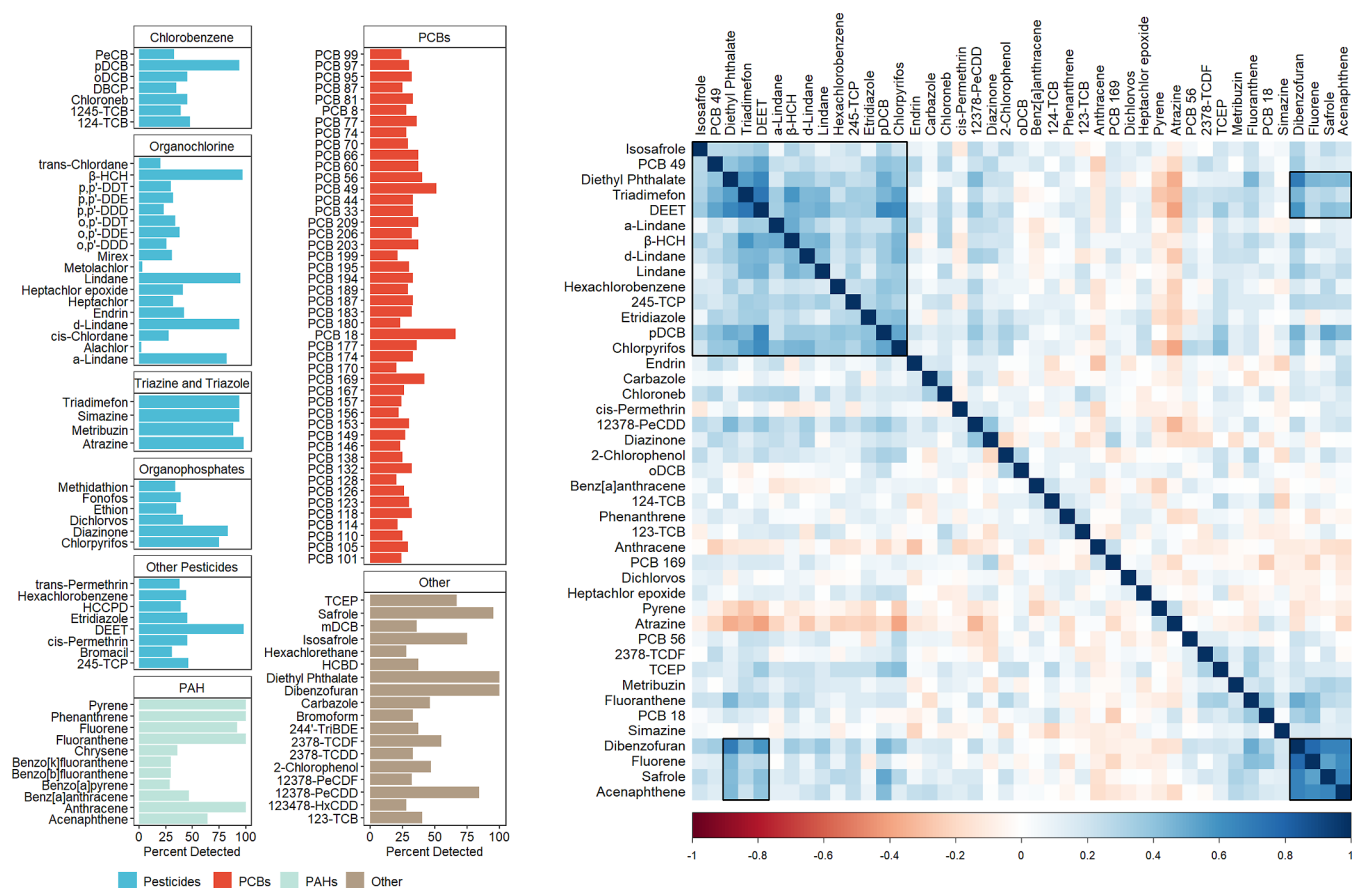


Figure 1. High-resolution gas chromatography mass spectrometry efficiently detects widespread detection of organic pollutants in seminal plasma. Left side—the detection rate of 118 targeted organic pollutants. Right—correlation plot of chemicals with $\geq 40\%$ detection rate.

and two PAHs, fluorene and acenaphthene. Although it is not clear which products or behaviors underlie this profile, these four chemicals also had positive correlations with many pesticides in our data.

Analysis of Targeted Organic Pollutants Shows Etridiazole was Associated with Poor Semen Parameters. We used two parallel approaches to quantify the associations between the targeted organic pollutants and semen parameters, including motility, concentration, total motile sperm, and a combined index that captures overall poor semen parameters (Figure S2).

First, we testing each exposure individually using an ExWAS approach, adjusting for age, BMI, smoking and cannabis use, and infertility diagnoses and correcting for multiple comparisons. We found that etridiazole was associated with lower total motile sperm (FDR $q = 0.01$), concentration (FDR $q = 0.07$), and overall index of semen parameters (FDR $q = 0.004$) (Figure 2A,B, Table S4). These results were robust when we excluded those with male factor infertility or very low total motile sperm count (Table S5).

We then sought to use our novel analysis approach, which combines a machine learning pattern recognition approach with modern mixture methods, to identify latent patterns in the exposure data and evaluate these exposure patterns, rather than individual exposures, with the semen parameter outcomes. In this approach, we first used PCP, a robust pattern recognition and data dimension reduction technique, to derive a low-rank exposure matrix that represents the latent underlying patterns without outliers and rare events (Figure S3). Despite the

apparent complexity in the raw observed data, ~ 5 PCs (PCP–PCs) explained $>99\%$ of the variability in the low-rank exposure matrix (Figure S4). Of these 5 PCP–PCs, we found that PCP–PC2 was associated with total motile sperm after adjusting for age, BMI, smoking and cannabis use, and infertility diagnoses and correcting for multiple comparisons (Figure 2C).

The loadings on PCP–PC2 showed high positive loadings for PAHs and triazine/triazole pesticides and highest negative loadings for etridiazole and 2,3,7,8-tetrachlorodibenzofuran (Figure S5). To formally investigate which exposures contributed to the association between PCP–PC2 and total motile sperm, the original observed exposure concentrations for the top 10 chemicals with highest absolute loadings were fit as a mixture in a BKMR model adjusting for age, BMI, smoking and cannabis use, and infertility diagnoses. The BKMR model showed that within this mixture, etridiazole had the highest posterior inclusion probability (Figure 2D) and was the only one that showed a negative association with any tested semen parameter (Figure 2E), which is evidence that etridiazole was the sole driver of this relationship. Indeed, we also saw that etridiazole was clearly associated with lower sperm motility in the BKMR model, which was also evidenced in the ExWAS models, but did not pass multiple testing corrections.

To ensure that our results were robust and not specific to model parameters, we conducted sensitivity analyses where we specified alternative model parameters for PCP and the number of mixture components for BKMR. There were no appreciable changes in the results as etridiazole was consistently associated with poor semen parameters (Figures S6–S7).

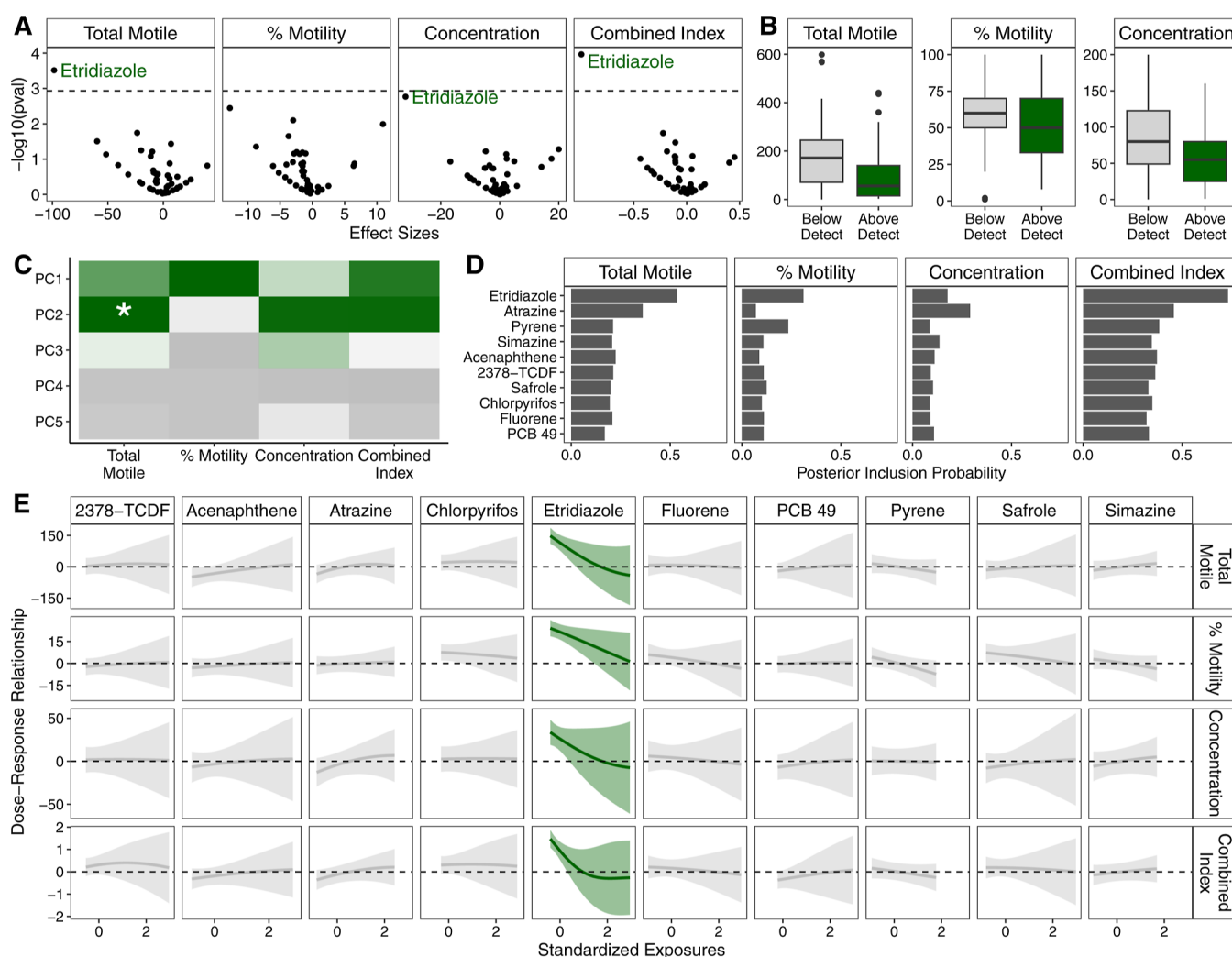


Figure 2. Exposome-wide association study using standard single-exposure regression models and machine learning pattern recognition based approach both identified etridiazole as negatively associated with worse semen parameters including lower total motile sperm, percent motility, and concentration. (A) Volcano plot of exposome-wide association study results, with the dashed line representing Bonferroni-corrected p -value cutoff threshold. (B) Boxplot of semen parameters among those with detectable levels of etridiazole compared to those with nondetectable levels of Etridiazole. (C) Heat plot of results PC regression. Green indicates negative association and the asterisk denotes statistical significance after Bonferroni correction. (D) Bar graph of posterior inclusion probability calculated from BKMR. (E) Estimated dose-response plots and associated 95% credible intervals from BKMR. Green line and credible interval indicates that the credible interval deviates from the null.

Non-Targeted Analysis Identified NDEA as Associated with Poor Semen Parameters. In addition to the set of targeted organic chemicals, we conducted a NTA of all high-quality spectra tentatively identified in seminal plasma with the goal of applying and leveraging the exposomic approach to identify potential male reproductive toxicants among less-investigated pollutants. We characterized the tentatively identified spectra for elemental composition and presence of functional groups (Figure S8). We found that the 39% of the tentatively identified features contained only hydrogen and carbon and $\sim 45\%$ did not carry any functional groups. Starting from 814 spectral peaks with abundances detected in all study samples, we again applied our PCP-based approach.

In these nontargeted data, PCP derived a low-rank exposure matrix with 6 components that explained $>99\%$ of the variance (Figure S9), one of which was associated with total motile sperm (Figure 3A). Using BKMR model with the top 10 loading peaks of this component, we found that one peak had the highest posterior inclusion probability for total motile sperm, concen-

tration, and combined index (Figure 3B). This peak had a retention time of 5.734 min and the mass-to-charge (m/z) ratio was 57.03352 and was the only one in this mixture that showed negative association with semen parameters (Figure 3C). These results did not differ when we considered alternative BKMR inputs (Figure S10).

To verify the observations from our PCP- and BKMR-based approach, we also modeled each of the top 10 peaks individually with linear regression, adjusting for age, BMI, smoking and cannabis use, and infertility diagnoses. These models also showed that the same peak (retention time = 5.734 min, m/z = 57.03352) was negatively associated with total motile sperm (p = 0.01) and combined index (p = 0.01) and marginally associated with lower sperm concentration (p = 0.07) (Figure 3D). These results persisted after excluding those with male factor infertility or very low total motile sperm count (p = 0.001 for total motile sperm, p = 0.02 for concentration, p = 0.003 for combined index).

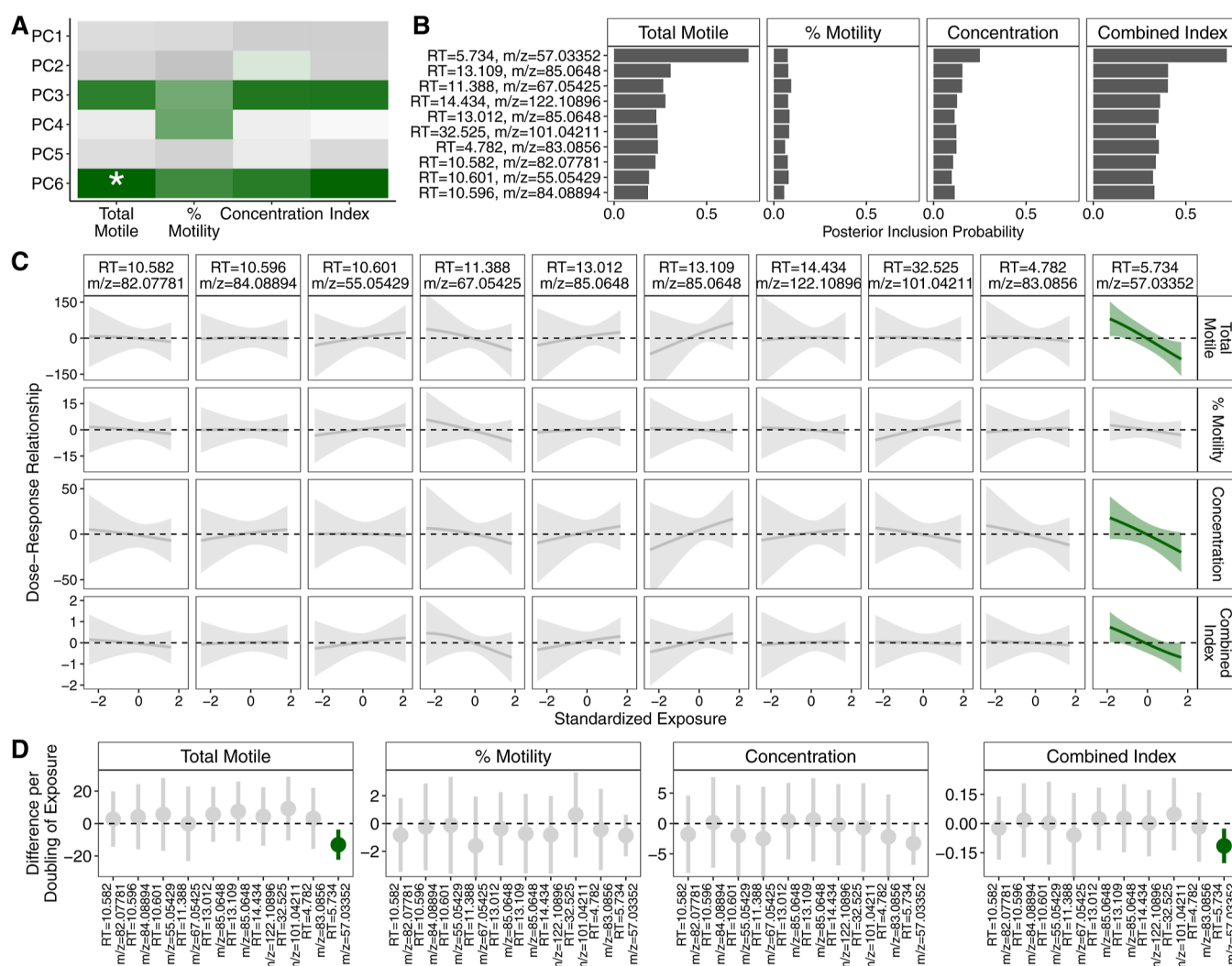


Figure 3. Machine learning pattern recognition based approach identified a spectral peak to be negatively associated with worse semen parameters. (A) Heat plot of results PC regression. Green indicates negative association and the asterisk denotes statistical significance after Bonferroni correction. (B) Bar graph of posterior inclusion probability calculated from BKMR. (C) Estimated dose-response plots and associated 95% credible intervals from BKMR. Green line and credible interval indicates that the credible interval deviates from the null. (D) Forest plot of estimates and associated 95% confidence intervals from standard single-exposure linear regression models.

In order to determine the chemical identity of the peak, we used a suspect screening analysis pipeline described previously¹⁶ and identified NDEA as a high quality match for this *m/z* feature with exact mass match, reverse match factor >600, and near exact retention time match (within 0.0008%). We compared this feature in our samples to the NDEA reference standard and found very similar chromatograph profile (Figure S11) and ion ratios (Table S6). Together, this provides level 1 identification confidence for this *m/z* feature as NDEA.

DISCUSSION

In this study, we show that numerous organic pollutants can be detected in seminal plasma and can be observed as several pollutants frequently co-occur with a range of other pollutants, forming distinct exposure profiles. Chemicals with low octanol-water partition coefficient, bioconcentration factor, and biodegradation half-life had higher detection rates among the participants. Additionally, we applied a novel pattern recognition approach to the exposure data to reduce dimensionality and subsequently identified etridiazole and NDEA as two potential male reproductive toxicants negatively associated with

semen parameters. These discoveries were supported by the classical ExWAS approach and linear regression analyses, demonstrating the effectiveness and validity of our novel data analysis approach, and persisted when men with male factor infertility and low sperm count were removed. On a broader level, our study shows that we can adapt high dimensional approaches to exposomic studies of environmental determinants of reproductive health, which may be necessary to better understand the environmental contributions of the global male fertility decline.

Pesticides have been previously link to semen parameters.^{35,36} To our knowledge, this is the first study to report the potential association of etridiazole, a pesticide commonly used to control rot due to fungi and oomycetes,³⁷ and male fertility parameters. Etridiazole is sold as the active ingredient in numerous commercial pesticides such as terrazole, Truban, and Banrot and has been used on golf courses, cotton ornaments, lawns, and agricultural seed products.³⁷ In addition to occupational exposures,^{38–40} etridiazole can be present in water and air near sites where it is frequently applied,^{37,41} which may lead to low levels of intermittent exposure in the general population.

Etridiazole is not typically found on food,^{37,42} which suggests that the observed exposure is unlikely to be a result of confounding via dietary patterns and associated habits. Ultimately, there is likely substantial variation in detection rate of etridiazole across different populations, depending on underlying exposure levels and the analytical methods. It is notable that a substantial portion of our population live or work near agricultural sites, which likely led to higher detection rates than other populations. Etridiazole is classified as a probable human carcinogen,^{43,44} in part due to evidence of testes tumor and testicular interstitial cell hyperplasia in rats,³⁷ which suggests potential reproductive toxicity. Mechanistically, etridiazole inhibits fungal growth by hydrolysis of phospholipid membrane of fungal mitochondria via increased phospholipase A activity.⁴⁵ The effect of etridiazole on mammalian cells is not well characterized, but there is evidence that it causes hemolysis and lipid peroxidation of human erythrocyte cell membranes by free radicals.⁴⁵ Thus, the biological mechanisms underlying the associations of etridiazole with semen parameters should be investigated in future studies.

NDEA is a nitrosamine and a probable human carcinogen^{43,44} with well characterized hepatotoxic, carcinogenic, and mutagenic properties.⁴⁶ NDEA is used in a wide range of industrial applications, including as a gasoline and lubricant additive, antioxidant, plastics stabilizer, and can be found in tobacco^{47,48} and food products.⁴⁹ It can also be found as a water treatment disinfection byproduct.^{50,51} A well-characterized mechanism of carcinogenesis by NDEA is through lipid peroxidation and generation of free radicals. However, lipid peroxidation and free radicals may result in a variety of other adverse health effects, including impaired spermatogenesis and poor semen parameters.^{52,53} Although there was no prior human evidence, there is compelling evidence demonstrating reproductive toxicity of NDEA in animal models. NDEA increased abnormal sperm and markedly decreased sperm count, sperm motility, and male reproductive organ weight in chronically exposed rats in two different experiments.^{54,55} These changes were complemented by depletion of antioxidants, increase in malondialdehyde and other markers of oxidative stress, elevated indicators of lipid peroxidation in the testes, pathological changes in seminiferous tubules, and changes in levels of key sex hormones.^{54,55} Similar induction of oxidative stress, changes in sex hormones, and pathological changes in seminiferous tubules were observed in rabbits.⁵⁶ Lycopene, an antioxidant, appeared to rescue the NDEA-induced effects, suggesting that oxidative stress is the underlying mechanism.⁵⁴ Thus, our study is consistent with the emergent experimental evidence and there is strong evidence that NDEA can lead to impaired spermatogenesis and reduced fertility in humans.

In addition to the discovery of two new potential human reproductive toxicants, our study also introduces a new workflow for the analysis of high dimensional exposome data. Our approach leverages a popular machine learning pattern recognition approach to remove outliers and rare events from noisy exposure data to uncover latent patterns. The resulting latent patterns can then use other machine learning and statistical techniques such as factor analysis (e.g., PCA) and mixtures analysis (e.g., BKMR) to screen hundreds of exposures with a smaller number of statistical tests while maintaining the ability to assess potential mixture effects and interactions. As proof of concept, we show that this workflow was able to recapitulate associations that were detectable in classical ExWAS analyses while identifying associations present in our data that

were not detectable via ExWAS analyses due to lack of power. In effect, our proposed workflow can maximize statistical efficiency and overcome a key challenge in exposome studies¹⁵ to efficiently screen hundreds of exposures using realistic study sample sizes in reproductive health studies to identify plausible reproductive toxicants.

There are additional considerations when applying this approach. First, we were able to identify specific patterns and chemicals related to our outcomes, but it is possible that PCs with a highly toxic chemical will not be associated with the outcome of interest if the rest of the chemicals loading highly on the pattern were null or protective, thereby diluting the effect of the highly toxic chemical. It may be advisable to conduct parallel ExWAS analyses, but whether ExWAS can identify such associations depend on the statistical power. Second, a collection of weakly toxic chemicals may have a significant joint effect without a single primary suspect. Although we did not observe this in our study, this is another strength of BKMR that should be kept in mind during the interpretation. Lastly, it is important to note that the individual chemical estimates are the product of a prior selection step, which means that the estimates do not reflect uncertainty in dimension reduction.

While our results are promising and consistent with experimental evidence, it raises some questions for future work. Given the absence of recent biomonitoring data for these chemicals, a comparison of exposure values in our study population to other populations cannot be conducted. For example, we detected relatively high levels of dioxin-like chemicals and etridiazole in our study compared to previous studies,^{57,58} but there are major differences between populations, biomatrices, and assay methods, as well as potential differences across time. Ultimately, this highlights the urgent need for better human exposure biomonitoring efforts for known environmental contaminants.⁵⁹ While there are several advantages to measuring parent compounds directly, future efforts may also consider measuring known human metabolites. For example, etridiazole⁶⁰ and NDEA⁶¹ have metabolites that may be detected at higher frequencies and quantities, making those metabolites potentially more attractive for targeted studies. Low detection may lead to nondifferential measurement error, which on average biases the results toward the null. Thus, while this likely did not result in spurious findings, in the case of etridiazole, additional efforts involving etridiazole metabolites and other populations would be helpful because the mechanism of action is unclear and complementary experimental evidence is necessary to understand how these putative male reproductive toxicants act mechanistically.

Our study also has some limitations. First, our study population was recruited from couples seeking IVF in Israel and the associations observed with etridiazole and NDEA need to be replicated in other populations. However, there is compelling support for the validity of the observed associations from prior literature and experimental evidence. Although recruitment of men from infertility center might not be representative of the general population, most of those included in our cohort did not have male factor infertility: 42% of the participants underwent IVF for preimplantation genetic diagnosis with normal sperm parameters, 9% underwent planned embryo cryopreservation and 32% due to female infertility. Only a few study participants were diagnosed with male factor infertility and had obvious poor semen parameters. Furthermore, our results were stronger when we excluded individuals with male factor infertility or very low total motile

sperm, which gives us confidence that these associations are likely also present in the general population. Additional limitation was that we assessed only three sperm parameters and not the full semen quality parameters according to the WHO. Upon the first visit in our unit, male partners were asked to provide a full semen analysis, including sperm morphology. However, these tests were performed in different laboratories and variable timing and only sperm volume, percent motility, and concentration were collected in our clinical IVF lab on the day of sample collection. Another limitation was that we only assessed semen parameters as the evidence suggested a global decline in sperm count. However, the relationship between semen parameters and fecundity is nonlinear^{7,62} and it is unknown whether etridiazole or NDEA are associated with measures of fecundity such as time-to-pregnancy. Lastly, we were unable to adjust for other potentially relevant factors such as diet, which could influence both our observed exposures and outcomes.

■ ASSOCIATED CONTENT

Data Availability Statement

Deidentified data from the current study are available from the corresponding author on reasonable request.

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.est.3c10314>.

Supporting Information includes additional methods used for the GC analysis of seminal plasma samples (Text S1, Tables S1–S2) and statistical analysis (Figures S1–S4). Summary statistics on exposure assessments can be found in Table S3. Additional details on the analysis of targeted exposome data can be found in Figures S5–S7 and Tables S4–S5. Additional details on the statistical analysis of nontargeted exposome data can be found in Figures S8–S10. Chemical identification information is for nontargeted data can be found in Figure S11 and Table S6 (PDF)

Recovery spikes of the contaminant mixture in fetal bovine serum; extraction recovery of the contaminants from NIST SRM (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Haotian Wu – Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York 10032, United States; orcid.org/0000-0002-6271-0249; Email: hw2694@cumc.columbia.edu

Authors

Vrinda Kalia – Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York 10032, United States; orcid.org/0000-0003-2244-6663

Katherine E. Manz – School of Engineering, Brown University, Providence, Rhode Island 02912, United States; orcid.org/0000-0002-8225-9168

Lawrence Chillrud – Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York 10032, United States

Nathalie Hoffmann Dishon – Infertility and IVF Unit, Department of Obstetrics and Gynecology, Chaim Sheba Medical Center (Tel Hashomer), Ramat Gan 5262000, Israel

Gabriela L. Jackson – Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York 10032, United States

Christian K. Dye – Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York 10032, United States

Raoul Orvieto – Infertility and IVF Unit, Department of Obstetrics and Gynecology, Chaim Sheba Medical Center (Tel Hashomer), Ramat Gan 5262000, Israel

Adva Aizer – Infertility and IVF Unit, Department of Obstetrics and Gynecology, Chaim Sheba Medical Center (Tel Hashomer), Ramat Gan 5262000, Israel

Hagai Levine – Braun School of Public Health, Hadassah Medical Organization, Faculty of Medicine, Hebrew University, Jerusalem 9112102, Israel

Marianthi-Anna Kioumourtzoglou – Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York 10032, United States

Kurt D. Pennell – School of Engineering, Brown University, Providence, Rhode Island 02912, United States; orcid.org/0000-0002-5788-6397

Andrea A. Baccarelli – Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York 10032, United States

Ronit Machtinger – Infertility and IVF Unit, Department of Obstetrics and Gynecology, Chaim Sheba Medical Center (Tel Hashomer), Ramat Gan 5262000, Israel; School of Medicine, Tel-Aviv University, Tel Aviv 6997801, Israel

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.est.3c10314>

Author Contributions

#H.W. and V.K. contributed equally to this paper.

Funding

This work was supported by the Environment and Health Fund (Israel, RPGA1901 to Levine), and National Institute of Environmental Health Sciences (USA, P30ES009089 to AAB, T32ES007272 to KEM, R01ES028805 to MAK),

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We would like to thank Lukas Scheidl for his assistance with sample extraction.

■ REFERENCES

- (1) World Health Organization. Infertility [Internet], 2023. <https://www.who.int/health-topics/infertility> (accessed Jan 23, 2023).
- (2) Mascarenhas, M. N.; Flaxman, S. R.; Boerma, T.; Vanderpoel, S.; Stevens, G. A. National, Regional, and Global Trends in Infertility Prevalence Since 1990: A Systematic Analysis of 277 Health Surveys. *PLoS Med.* **2012**, *9* (12), No. e1001356.
- (3) Agarwal, A.; Mulgund, A.; Hamada, A.; Chyatte, M. R. A unique view on male infertility around the globe. *Reprod. Biol. Endocrinol.* **2015**, *13*, 37.
- (4) Skakkebaek, N. E.; Lindahl-Jacobsen, R.; Levine, H.; Andersson, A. M.; Jørgensen, N.; Main, K. M.; Lidegaard, Ø.; Priskorn, L.; Holmboe, S. A.; Bräuner, E. V.; et al. Environmental factors in declining human fertility. *Nat. Rev. Endocrinol.* **2022**, *18* (3), 139–157.
- (5) Levine, H.; Jørgensen, N.; Martino-Andrade, A.; Mendiola, J.; Weksler-Derri, D.; Jolles, M.; et al. Temporal trends in sperm count: a systematic review and meta-regression analysis of samples collected

- globally in the 20th and 21st centuries. *Hum. Reprod. Update* **2022**, *29*, 157.
- (6) WHO. *WHO Laboratory Manual for the Examination and Processing of Human Semen* [Internet]; World Health Organization; Report No: 6th ed., 2021. <https://www.who.int/publications/i/item/9789240030787>.
- (7) Bonde, J. P.; Ernst, E.; Jensen, T. K.; Hjollund, N. H.; Kolstad, H.; Henriksen, T. B.; Giwercman, A.; Skakkebaek, N. E.; Henriksen, T. B.; Olsen, J. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *Lancet* **1998**, *352* (9135), 1172–1177.
- (8) Jørgensen, N.; Lamb, D. J.; Levine, H.; Pastuszak, A. W.; Sigalos, J. T.; Swan, S. H.; Eisenberg, M. L. Are worldwide sperm counts declining? *Fertil. Steril.* **2021**, *116* (6), 1457–1463.
- (9) Wu, H.; Hauser, R.; Krawetz, S. A.; Pilsner, J. R. Environmental Susceptibility of the Sperm Epigenome During Windows of Male Germ Cell Development. *Curr. Environ. Health Rep.* **2015**, *2* (4), 356–366.
- (10) Greeson, K. W.; Crow, K. M. S.; Edenfield, R. C.; Easley, C. A. Inheritance of paternal lifestyles and exposures through sperm DNA methylation. *Nat. Rev. Urol.* **2023**, *20* (6), 356–370.
- (11) Wild, C. P. Complementing the Genome with an “Exposome”: The Outstanding Challenge of Environmental Exposure Measurement in Molecular Epidemiology. *Cancer Epidemiol. Biomarkers Prev.* **2005**, *14* (8), 1847–1850.
- (12) Vermeulen, R.; Schymanski, E. L.; Barabási, A. L.; Miller, G. W. The exposome and health: Where chemistry meets biology. *Science* **2020**, *367* (6476), 392–396.
- (13) Cheng, C. Y.; Mruk, D. D. The Blood-Testis Barrier and Its Implications for Male Contraception. *Pharmacol. Rev.* **2012**, *64* (1), 16–64.
- (14) Su, L.; Mruk, D. D.; Cheng, C. Y. Drug transporters, the blood–testis barrier, and spermatogenesis. *J. Endocrinol.* **2011**, *208* (3), 207–223.
- (15) Chung, M. K.; Buck Louis, G. M.; Kannan, K.; Patel, C. J. Exposome-wide association study of semen quality: systematic discovery of endocrine disrupting chemical biomarkers in fertility require large sample sizes. *Environ. Int.* **2019**, *125*, 505–514.
- (16) Manz, K. E.; Yamada, K.; Scheidl, L.; La Merrill, M. A.; Lind, L.; Pennell, K. D. Targeted and Nontargeted Detection and Characterization of Trace Organic Chemicals in Human Serum and Plasma Using QuEChERS Extraction. *Toxicol. Sci.* **2021**, *185* (1), 77–88.
- (17) Galusha, A. L.; Farnsworth, A. C.; Bloom, M. S.; Kruger, P. C.; McGough, A.; Lenhart, N.; Wong, R.; Fujimoto, V. Y.; Mok-Lin, E.; Parsons, P. J. Trace element analysis of human seminal plasma: A cautionary tale of preanalytical variation and use of non-traditional matrices in human biomonitoring studies. *Int. J. Hyg Environ. Health* **2021**, *234*, 113751.
- (18) Drabovich, A. P.; Saraon, P.; Jarvi, K.; Diamandis, E. P. Seminal plasma as a diagnostic fluid for male reproductive system disorders. *Nat. Rev. Urol.* **2014**, *11* (5), 278–288.
- (19) Gibson, E. A.; Zhang, J.; Yan, J.; Chillrud, L.; Benavides, J.; Nunez, Y.; Herbstman, J. B.; Goldsmith, J.; Wright, J.; Kioumourtoglou, M. A. Principal Component Pursuit for Pattern Identification in Environmental Mixtures. *Environ. Health Perspect.* **2022**, *130* (11), 117008.
- (20) Tao, R. H.; Chillrud, L. G.; Nunez, Y.; Rowland, S. T.; Boehme, A. K.; Yan, J.; Goldsmith, J.; Wright, J.; Kioumourtoglou, M. A. Applying principal component pursuit to investigate the association between source-specific fine particulate matter and myocardial infarction hospitalizations in New York City. *Environ. Epidemiol.* **2023**, *7* (2), No. e243.
- (21) Bobb, J. F.; Valeri, L.; Claus Henn, B.; Christiani, D. C.; Wright, R. O.; Mazumdar, M.; Godleski, J. J.; Coull, B. A. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics.* **2015**, *16* (3), 493–508.
- (22) Maitre, L.; Bustamante, M.; Hernández-Ferrer, C.; Thiel, D.; Lau, C. H. E.; Siskos, A. P.; Vives-Usano, M.; Ruiz-Arenas, C.; Pelegrí-Sisó, D.; Robinson, O.; et al. Multi-omics signatures of the human early life exposome. *Nat. Commun.* **2022**, *13* (1), 7024.
- (23) Cooper, T. G.; Noonan, E.; von Eckardstein, S.; Auger, J.; Baker, H. W. G.; Behre, H. M.; Haugen, T. B.; Kruger, T.; Wang, C.; Mbizvo, M. T.; et al. World Health Organization reference values for human semen characteristics. *Hum. Reprod. Update* **2010**, *16* (3), 231–245.
- (24) Williams, A. J.; Grulke, C. M.; Edwards, J.; McEachran, A. D.; Mansouri, K.; Baker, N. C.; Patlewicz, G.; Shah, I.; Wambaugh, J. F.; Judson, R. S.; et al. The CompTox Chemistry Dashboard: a community data resource for environmental chemistry. *J. Cheminf.* **2017**, *9* (1), 61.
- (25) Szöcs, E.; Stirling, T.; Scott, E. R.; Scharmüller, A.; Schäfer, R. B. webchem: An R Package to Retrieve Chemical Information from the Web. *J. Stat. Software* **2020**, *93*, 1–17.
- (26) Guha, R. Chemical Informatics Functionality in R. *J. Stat. Software* **2007**, *18*, 1–16.
- (27) Koelmel, J. P.; Xie, H.; Price, E. J.; Lin, E. Z.; Manz, K. E.; Stelben, P.; Paige, M. K.; Papazian, S.; Okeme, J.; Jones, D. P.; et al. An actionable annotation scoring framework for gas chromatography-high-resolution mass spectrometry. *Exposome* **2022**, *2* (1), osac007.
- (28) Visscher, P. M.; Wray, N. R.; Zhang, Q.; Sklar, P.; McCarthy, M. L.; Brown, M. A.; Yang, J. 10 Years of GWAS Discovery: Biology, Function, and Translation. *Am. J. Hum. Genet.* **2017**, *101* (1), 5–22.
- (29) Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. Roy. Stat. Soc. B* **1995**, *57* (1), 289–300.
- (30) Weisskopf, M. G.; Seals, R. M.; Webster, T. F. Bias Amplification in Epidemiologic Analysis of Exposure to Mixtures. *Environ. Health Perspect.* **2018**, *126* (4), 047003.
- (31) Webster, T. F.; Weisskopf, M. G. Epidemiology of exposure to mixtures: we cant be casual about causality when using or testing methods. *arXiv* **2020**, arXiv 200701370. preprint
- (32) Candès, E. J.; Li, X.; Ma, Y.; Wright, J. Robust principal component analysis? *J. ACM* **2011**, *58* (3), 1–37.
- (33) Wright, J.; Ganesh, A.; Rao, S.; Peng, Y.; Ma, Y. Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices via Convex Optimization. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc., 2009. https://proceedings.neurips.cc/paper_files/paper/2009/hash/c45147dee729311ef5b5c3003946c48f-Abstract.html.
- (34) Zhang, J.; Yan, J.; Wright, J. Square Root Principal Component Pursuit: Tuning-Free Noisy Robust Matrix Recovery. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc., 2021; p 29464–29475. <https://proceedings.neurips.cc/paper/2021/hash/f65854da4622c1f1ad4feb361d7703c-Abstract.html>.
- (35) Chiu, Y. H.; Afeiche, M. C.; Gaskins, A. J.; Williams, P. L.; Petrozza, J. C.; Tanrikut, C.; Hauser, R.; Chavarro, J. Fruit and vegetable intake and their pesticide residues in relation to semen quality among men from a fertility clinic. *Hum. Reprod.* **2015**, *30* (6), 1342–1351.
- (36) Levine, H.; Swan, S. H. Is dietary pesticide exposure related to semen quality? Positive evidence from men attending a fertility clinic. *Hum. Reprod.* **2015**, *30* (6), 1287–1289.
- (37) USEPA. *Etridiazole (Terrazole)*. United States Environmental Protection Agency; (Reregistration Eligibility Decision (RED)); Report No.: EPA 738-R-00-019, 2000. <https://archive.epa.gov/pesticides/reregistration/web/pdf/0009red.pdf>.
- (38) Stamper, J. H.; Nigg, H. N.; Mahon, W. D.; Nielsen, A. P.; Royer, M. D. Pesticide exposure to greenhouse foggers. *Chemosphere* **1988**, *17* (5), 1007–1023.
- (39) Stamper, J. H.; Nigg, H. N.; Mahon, W. D.; Nielsen, A. P.; Royer, M. D. Pesticide exposure to a greenhouse drencher. *Bull. Environ. Contam. Toxicol.* **1989**, *42* (2), 209–217.
- (40) Stamper, J. H.; Nigg, H. N.; Mahon, W. D.; Nielsen, A. P.; Royer, M. D. Pesticide exposure to greenhouse handgunners. *Arch. Environ. Contam. Toxicol.* **1989**, *18* (4), 515–529.
- (41) Peck, A. M.; Hornbuckle, K. C. Gas-Phase Concentrations of Current-Use Pesticides in Iowa. *Environ. Sci. Technol.* **2005**, *39* (9), 2952–2959.
- (42) USDA. *Pesticide Data Program* [Internet]; United States Department of Agriculture, 2023. <https://www.ams.usda.gov/datasets/pdp>.

- (43) IARC. *IARC Publications—IARC Monographs*; International Agency for Research on Cancer: Lyon, 2010; Vol. 94. <https://www.iarc.who.int/news-events/iarc-publications-iarc-monographs-volume-94/>.
- (44) IARC. *Overall Evaluations of Carcinogenicity: An Updating of IARC Monographs Volumes 1–42* [Internet]; International Agency for Research on Cancer: Lyon, 1987. <https://publications.iarc.fr/Book-And-Report-Series/Iarc-Monographs-Supplements/Overall-Evaluations-Of-Carcinogenicity-An-Updating-Of-IARC-Monographs-Volumes-1%E2%80%9342-1987>.
- (45) Radzuhn, B.; Lyr, H. On the mode of action of the fungicide etridiazole. *Pestic. Biochem. Physiol.* **1984**, *22* (1), 14–23.
- (46) Verna, L.; Whysner, J.; Williams, G. M. N-Nitrosodiethylamine mechanistic data and risk assessment: Bioactivation, DNA-adduct formation, mutagenicity, and tumor initiation. *Pharmacol. Ther.* **1996**, *71* (1–2), 57–81.
- (47) Lv, F.; Guo, J.; Yu, F.; Zhang, T.; Zhang, S.; Cui, H.; Liu, X.; Chen, L.; Liu, L.; Liu, S.; et al. Determination of nine volatile N-nitrosamines in tobacco and smokeless tobacco products by dispersive solid-phase extraction with gas chromatography and tandem mass spectrometry. *J. Sep. Sci.* **2016**, *39* (11), 2123–2128.
- (48) Smith, C. J.; Perfetti, T. A.; Rumble, M. A.; Rodgman, A.; Doolittle, D. J. IARC group 2A Carcinogens” reported in cigarette mainstream smoke. *Food Chem. Toxicol.* **2000**, *38* (4), 371–383.
- (49) ParkSeo, J. J. e.; SeoLee, J. J. y.; LeeKwon, J. H.; Kwon, H. Distribution of Seven N-Nitrosamines in Food. *Toxicol. Res.* **2015**, *31* (3), 279–288.
- (50) Bond, T.; Huang, J.; Templeton, M. R.; Graham, N. Occurrence and control of nitrogenous disinfection by-products in drinking water – A review. *Water Res.* **2011**, *45* (15), 4341–4354.
- (51) NDMA and Other Nitrosamines - Drinking Water Issues | California State Water Resources Control Board [Internet]. 2023, https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/NDMA.html (accessed Jun 20, 2023).
- (52) Jones, R.; Mann, T. R. R. Lipid peroxidation in spermatozoa. *Proceedings of the Royal Society of London Series B Biological Sciences*; Royal Society, 1997; Vol. 184 (1074), pp 103–107.
- (53) Aitken, R. J. Free radicals, lipid peroxidation and sperm function. *Reprod. Fertil. Dev.* **1995**, *7* (4), 659–668.
- (54) Kaya, E.; Ozer Kaya, S.; Yilmaz, S.; Ceribasi, A. O.; Turk, G. Evaluation of ameliorating effect of lycopene against testicular toxicity due to diethylnitrosamine using biochemical, spermatological and histopathological data. *Andrologia.* **2019**, *51* (6), No. e13274.
- (55) Owumi, S. E.; Adedara, I. A.; Duro-Ladipo, A.; Farombi, E. O. Acute diethyl nitrosamine and cadmium co-exposure exacerbates deficits in endocrine balance, sperm characteristics and antioxidant defence mechanisms in testes of pubertal rats. *Andrologia.* **2019**, *51* (4), No. e13230.
- (56) Sheweita, S. A.; El Banna, Y. Y.; Balbaa, M.; Abdullah, I. A.; Hassan, H. E. N-nitrosamines induced infertility and hepatotoxicity in male rabbits. *Environ. Toxicol.* **2017**, *32* (9), 2212–2220.
- (57) Torres-Moreno, A. C.; Mejia-Grau, K.; Puente-DelaCruz, L.; Codling, G.; Villa, A. L.; Ríos-Marquez, O.; Patequiva-Chauta, L.; Cobo, M.; Johnson-Restrepo, B. Polybrominated diphenyl ethers (PBDEs), polychlorinated biphenyls (PCBs), organochlorine pesticides (OCPs) in human breast milk from Colombia: A probabilistic risk assessment approach. *Chemosphere* **2023**, *339*, 139597.
- (58) Mayani, A.; Barel, S.; Soback, S.; Almagor, M. Dioxin concentrations in women with endometriosis. *Hum. Reprod.* **1997**, *12* (2), 373–375. Feb
- (59) Yudkovski, Y.; Herut, B.; Shefer, E.; Tom, M. Dioxin-like biological activity of organic extracts from sediments and fish livers sampled along the Israeli Mediterranean and Red Sea coasts. *Mar. Pollut. Bull.* **2015**, *98* (1–2), 295–300.
- (60) van Welie, R. T.; Mensert, R.; Van Duyn, P.; Vermeulen, N. P. Identification and quantitative determination of a carboxylic and a mercapturic acid metabolite of etridiazole in urine of rat and man. Potential tools for biological monitoring. *Arch. Toxicol.* **1991**, *65* (8), 625–632.
- (61) Sierra, L. M.; Tosal, L.; Nivard, M. J. M.; Comendador, M. A.; Vogel, E. W. The importance of distinct metabolites of N-nitrosodiethylamine for its in vivo mutagenic specificity. *Mutat. Res. Fund Mol. Mech. Mutagen* **2001**, *483* (1–2), 95–104.
- (62) Slama, R.; Eustache, F.; Ducot, B.; Jensen, T. K.; Jørgensen, N.; Horte, A.; et al. Time to pregnancy and semen parameters: a cross-sectional study among fertile couples from four European cities. *Hum. Reprod.* **2002**, *17* (2), 503–515.