A Section 508–conformant HTML version of this article is available at https://doi.org/10.1289/EHP12678.

Adult Organophosphate and Carbamate Insecticide Exposure and Sperm Concentration: A Systematic Review and Meta-Analysis of the Epidemiological Evidence

Lauren B. Ellis,[1,2] Karen Molina,[2] C. Rebecca Robbins,[2] Marlaina Freisthler,[2] Daria Sgargi,[3] Daniele Mandrioli,[3,4] and Melissa J. Perry[5]*

[1]Bouvé College of Health Sciences, Northeastern University, Boston, Massachusetts, USA
[2]Department of Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Washington, District of Columbia, USA
[3]Cesare Maltoni Cancer Research Center, Ramazzini Institute, Bologna, Italy
[4]Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy
[5]College of Public Health, George Mason University, Fairfax, Virginia, USA

Background: Evidence of the negative impacts of contemporary use insecticides on sperm concentration has increased over the last few decades; however, meta-analyses on this topic are rare.

Objectives: This investigation assessed the qualitative and quantitative strength of epidemiological evidence regarding adult exposure to two classes of contemporary use insecticides—organophosphates (OPs) and *N*-methyl carbamates (NMCs)—and sperm concentration using robust and reproducible systematic review and meta-analysis methods.

Methods: Three scientific databases (PubMed, Scopus, and Web of Science), two U.S. government databases (NIOSHTIC-2 and Science.gov), and five nongovernmental organization websites were searched for relevant primary epidemiological studies published in any language through 11 August 2022. Risk of bias and strength of evidence were evaluated according to Navigation Guide systematic review methodology. Bias-adjusted standardized mean difference effect sizes were calculated and pooled using a three-level, multivariate random-effect meta-analysis model with cluster-robust variance estimation.

Results: Across 20 studies, 21 study populations, 42 effect sizes, and 1,774 adult men, the pooled bias-adjusted standardized mean difference in sperm concentration between adult men more- and less-exposed to OP and NMC insecticides was −0.30 (95% CI: −0.49, −0.10; $P_{Satt} < 0.01$). Sensitivity and subgroup analyses explored statistical heterogeneity and validated the model robustness. Although the pooled effect estimate was modified by risk of bias, insecticide class, exposure setting, and recruitment setting, it remained negative in direction across all meta-analyses. The body of evidence was rated to be of moderate quality, with sufficient evidence of an association between higher adult OP and NMC insecticide exposure and lower sperm concentration.

Discussion: This comprehensive investigation found sufficient evidence of an association between higher OP and NMC insecticide exposure and lower sperm concentration in adults. Although additional cohort studies can be beneficial to fill data gaps, the strength of evidence warrants reducing exposure to OP and NMC insecticides now to prevent continued male reproductive harm. https://doi.org/10.1289/EHP12678

Introduction

Studies conducted across a wide range of populations have found significant decreases in average sperm concentrations over the last century,[1–6] and the prospect of further declines threatens population-wide indicators of male fecundity.[7,8] Oligospermia (low sperm concentration), in isolation or in concert with other abnormal semen parameters, is commonly used to diagnose idiopathic male factor infertility.[9,10] Oligospermia is diagnosed by semen analysis using the World Health Organization (WHO) lower reference limit of <16 million spermatozoa/mL of semen, which in 2021 represented the fifth percentile of sperm concentration among fertile men.[11,12] Low sperm concentration is also associated with general indicators of reduced male health,[13,14] including higher rates of cancer,[15,16] chronic comorbidities,[17,18] and all-cause mortality.[19,20]

Causes of low sperm concentration are not fully known, but important risk factors include age,[21] nutrition and lifestyle factors,[22] and exposure to reproductive toxicants in the environment, particularly endocrine disrupting chemicals (EDCs).[23] Reproductive toxicants are ubiquitous in the environment, and usually go unnoticed until unintended adverse effects are observed. Pesticides are a prime example in that pesticides have known endocrine disrupting and reproductive effects[24,25] but continue to be manufactured and widely applied, resulting in occupational and environmental exposures.[26,27] Although the intention of pesticide application is to target pests, pesticides can also induce acute and chronic toxicity in nontarget species, including reproductive harm in humans.[28]

In the latter half of the twentieth century, growing concerns about the persistence and bioaccumulation of organochlorine insecticides led to the adoption of less-persistent compounds, including organophosphate (OP) and *N*-methyl carbamate (NMC) insecticides.[29] OP and NMC insecticides are characterized as cholinesterase inhibitors because of their shared principal mechanism of toxicity—the inhibition of cholinesterase enzymes that hydrolyze or degrade important neurotransmitters, including acetylcholine.[30] As such, cholinesterase activity is a well-accepted biomarker of OP and NMC insecticide exposure and effect in humans.[30–32]

Prior qualitative reviews have reported deleterious associations between environmental and occupational pesticide exposure,

Address correspondence to Melissa J. Perry, College of Public Health, George Mason University, 4400 University Dr., Fairfax, VA 22030 USA. Telephone: (703) 993-1918. Email: mperry27@gmu.edu

*L.B.E. is currently at Northeastern University and was previously at the George Washington University at the time she conducted the bulk of the research.

Supplemental Material is available online (https://doi.org/10.1289/EHP12678).

This research did not receive funding from agencies in the public, commercial, or not-for-profit sectors. M.J.P. provided expert consultation in legal cases pertinent to occupational COVID-19 and pesticide health effects in 2021. The other authors of this paper declare they have no actual or potential competing financial interest and that their freedom to design, conduct, interpret, and publish research is not compromised by any controlling sponsor as a condition of review or publication.

Received 31 December 2022; Revised 27 September 2023; Accepted 5 October 2023; Published 15 November 2023.

Note to readers with disabilities: *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact ehpsubmissions@niehs.nih.gov. Our staff will work with you to assess and meet your accessibility needs within 3 working days.

including OP and NMC compounds, and sperm concentration[33–36]; however, meta-analyses on this topic are rare. Based on decades of studies of the negative impacts of contemporary use insecticides on sperm concentration and the lack of related meta-analyses, the objective of this investigation was to assess the qualitative and quantitative strength of epidemiological evidence regarding adult exposure to two classes of contemporary use insecticides—OPs and NMCs—and sperm concentration using robust and reproducible systematic review and meta-analysis methods.

Methods

Systematic Review Methodology

This systematic review was performed according to the Navigation Guide, a validated systematic review methodology designed to evaluate the quality and strength of evidence in the context of environmental and reproductive health.[37–39] Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 reporting guidelines were also followed (Table S1).[40]

A review team with expertise in environmental and reproductive epidemiology, pesticide exposure assessment, semen analysis, and systematic review was assembled in October 2020. Prior to initiating the literature search in late November 2020, the review team developed a protocol to prespecify key elements of the project, including search methods, study selection criteria, study screening criteria, data extraction forms, risk of bias domains, statistical analysis plans, quality of evidence factors, and strength of evidence factors.

The protocol was adapted from the protocol for Navigation Guide Case Study #5 on polybrominated diphenyl ethers (PBDEs) and human neurodevelopment.[41] Although the protocol was not formally published, the original protocol with amendments made during the review is available on GitHub (https://github.com/Lauren-Ellis/Ellis-et-al.-2023-OP-NMC-Insecticides-and-Sperm-Concentration).

As discussed in the protocol, the research team narrowed the review to OP and NMC insecticides given *a*) both are contemporary use insecticides, *b*) evidence of shared exposure measurement through cholinesterase monitoring, and *c*) the wealth of studies available to assess.

Study Question

This sytematic review and meta-analysis sought to answer the research question "What is the association between adult exposure to OP and NMC insecticides and sperm concentration?" The review team developed a Population, Exposure, Comparator, Outcome (PECO) statement[42] to guide the study search and selection process:

- ***Population:*** adult human males (≥18 years of age)
- ***Exposure:*** OP and NMC insecticides; occupation-based, self-report, proxy or biomonitoring assessment methods; documented, measured, or estimated exposure directly in study subjects; occupational or environmental exposure settings; non-acute exposure contexts
- ***Comparator:*** non- or less-exposed adult human males (≥18 years of age)
- ***Outcome:*** sperm concentration (operationally defined as millions of sperm per milliliter of semen) measured on a continuous or dichotomous scale (i.e., clinical diagnosis of low sperm concentration, or oligospermia)

Evidence Selection

Literature search strategy. Three scientific databases (PubMed, Scopus, and Web of Science), two U.S. government databases (NIOSHTIC-2 and Science.gov), and five nongovernmental organization (NGO) websites [Beyond Pesticides (https://www.beyondpesticides.org/home), National Pesticide Information Center (http://npic.orst.edu/), Pesticide Action Network North America (https://www.panna.org/), Collaborative for Health & Environment (https://www.healthandenvironment.org/home), and Environmental Working Group (https://www.ewg.org/)] were searched for relevant primary epidemiological studies published in any language through 11 August 2022. Scientific databases were searched five times throughout the duration of the review, whereas government databases were searched four times throughout the duration of the review. See Table S2 for more detail.

Search terms based on the Medical Subject Headings database from Knapke et al.[36] were used to search scientific databases. These search terms were simplified to search government databases and NGO websites. Database search terms are presented in Table S2 and NGO website search strategies are presented in Table S3.

Scientific database search results were imported into the systematic review management software Covidence (https://www.covidence.org), where duplicate records (titles and abstracts) were automatically and manually removed prior to screening. The lead review author (L.B.E.) also hand-searched government database and NGO website search results to capture studies that may have been missed. The reference lists of all reports (full-text journal articles, government reports, or other documents supplying relevant information about a particular study or studies) assessed for eligibility were also hand-searched. Titles deemed potentially relevant during hand-searching were recorded, checked against the existing Covidence database, and, if unique, imported into Covidence for screening.

Study selection. Primary epidemiological studies that examined the association between adult OP, NMC, or combined OP and NMC insecticide exposure and sperm concentration were considered eligible for inclusion in this review. Studies published in any language other than English were accepted and translated to English using Google Translate prior to screening. A diverse range of study designs, geographies, and exposure scenarios were accepted. Studies that examined exposures that occurred before adulthood or in someone other than the study participant, acute exposure events (e.g., chemical warfare or industrial accidents), other chemical exposures, and other sperm-related outcomes were excluded from consideration.

Study selection took place in three sequential screening stages according to predefined inclusion and exclusion criteria: *a*) title and abstract screening, *b*) full-text screening, and *c*) OP and NMC insecticide screening. The title and abstract and full-text screening stages aimed to identify studies that examined the association between any type of pesticide exposure and sperm concentration (Table S4), whereas the OP and NMC insecticide screening stage aimed to identify studies that examined the association between OP, NMC, or combined OP and NMC insecticide exposure and sperm concentration (Table S5). All reports assessed for eligibility during full-text screening were categorized in Covidence according to pesticide class, which facilitated the OP and NMC insecticide screening stage. Only studies that met the full-text screening criteria and examined an OP, NMC, or combined OP and NMC insecticide exposure were included in this review (Table S5).

Two review authors (L.B.E. and K.M., screeners) independently screened all retrieved records and reports according to predetermined criteria. Screeners resolved discrepancies via discussion and consulted other review authors (C.R.R., M.F., M.J.P.) as needed. If a record was inaccessible, screeners completed a brief online search before screening according to the title, erring on the side of overinclusion at the title and abstract

screening stage. If a report was inaccessible, the lead review author (L.B.E.) contacted study authors via email to request access to the report. Study screening results are presented in a PRISMA flow diagram (Figure 1).

Data extraction. Two review authors (L.B.E. and K.M., data extractors) independently extracted study characteristics (e.g., study design, exposure and recruitment setting, exposure assessment method) and meta-analysis eligibility factors (e.g., statistical transformations) from all studies included in the systematic review. Quantitative study results (e.g., measures of association or "effect sizes") were extracted from studies included in the meta-analysis. If multiple reports on the same study population reported on identical analyses, they were considered to be separate reports on the same study and were consolidated as a single study prior to data extraction. If multiple reports on the same study population reported on different analyses, they were considered to be separate studies on the same study population and were not consolidated prior to data extraction. The complete data extraction form, including amendments and clarifications, is presented in Table S6. Study characteristics, meta-analysis eligibility factors, and study results are presented in tabular format. Standardized results from studies included in the meta-analysis are shown in a forest plot, whereas results from studies excluded from the meta-analysis are summarized in the text.

Results from all relevant exposure definitions, doses, and sampling time periods were included. If a study reported multiple effect size indices (i.e., mean difference, odds ratio, correlation) on the same exposure and outcome relationship, the most adjusted effect size index was extracted. For studies that reported more than one most adjusted effect size index, the most commonly reported effect size index across the body of evidence was extracted in an effort to reduce the number of statistical effect size transformations required. Detailed instructions on methods for choosing which effect size(s) to extract from each study are presented in the data extraction form (Table S6C).

The data extraction form was piloted on three randomly selected studies to increase extractor consistency and address nuances in the data extraction form. Discrepancies in the extracted data were discussed and resolved via data extractor discussion; other review authors (C.R.R., M.F., M.J.P.) were consulted as needed. The data extractors then independently performed a quality control (QC) check by reviewing the extracted data (including resolved discrepancies) for accuracy. Any adjustments considered during the individual QC effort were recorded, discussed, and finalized. The lead review author (L.B.E.) contacted corresponding study authors via email to retrieve data needed for meta-analysis, including data needed to back-transform log-scale effect sizes and missing results data.

Risk of Bias Assessment

Consistent with the Navigation Guide, the following eight risk of bias domains were assessed for each study included in the systematic review: recruitment strategy, blinding, exposure assessment, outcome assessment, confounding, incomplete outcome data, selective outcome reporting, and conflict of interest.[38,39,43] These risk of bias domains are based on those from the Cochrane Collaboration[44] and the Agency for Healthcare Research and Quality,[45] as well as suggestions to consider conflict of interest in the Cochrane risk of bias tool.[46] Risk of bias ratings and rationales are reported in tabular format in the supplemental materials. This data was used to create study-specific and summary-level data visualizations on Microsoft Excel (version 16.73), which are included as figures in the main text. The summary-level figure was generated by calculating the frequency of each risk of bias rating for each domain and visualizing the frequency data in a horizontal bar chart.

The review team adapted the risk of bias assessment criteria used in Lam et al.[41] to the research question at hand (Table S7). Possible ratings were low, probably low, probably high, or high risk of bias for the following domains: recruitment strategy, blinding, exposure assessment, outcome assessment, and incomplete outcome data. The review team determined that the rating option probably high was not applicable for the selective outcome reporting and conflict of interest domains (Table S7, in the section "Clarifications made during pilot"); thus, the three rating options for these two domains were low, probably low, and high. Finally, risk of confounding bias was rated on a dichotomous scale as either low or high based on whether the study accounted for age and smoking. The review team identified only age and smoking as strict confounders based on consistent evidence of their association with both the exposure and outcome of interest.

Two of seven possible risk of bias assessors [four review authors (L.B.E., K.M., C.R.R., M.F.) and three review contributors (D.P.M., S.A., E.G.)] independently assessed each included study for risk of bias according to prespecified risk of bias criteria (Table S7). Discrepancies in risk of bias domain ratings were resolved by a third independent assessor (one of seven possible assessors listed above) and discussed as a group as needed. A training session was held before initiating the risk of bias assessments to review and clarify nuances in the risk of bias criteria. Following the training, all risk of bias assessors completed three pilot studies before independently assessing studies for risk of bias. The prespecified risk of bias criteria was available as a reference to assessors during risk of bias assessments in an effort to improve inter-assessor consistency. The overall risk of bias across the body of evidence was determined according to quality of evidence criteria adapted from the Navigation Guide (Table S9, in the section "Risk of bias across studies"). Study authors were not contacted for additional information to inform the risk of bias assessments.

Statistical Methods

Effect size index chosen for synthesis. Bias-adjusted standardized mean difference (Hedges' g) was used as the effect size index for synthesis. Hedges' g represents the mean difference in sperm concentration between adult men more- and less-exposed to OP and NMC insecticides on a standardized scale in units of pooled standard deviations, corrected for small-sample bias.[47] The term effect size is used generally to characterize the outcome measure used for the meta-analysis, consistent with the meta-analysis literature,[47] and does not necessarily indicate a causal effect.

Hedges' g was chosen as the effect size index for synthesis because *a*) the measure required the least amount of statistical effect size transformations, given that mean difference in sperm concentration between more- and less-exposed men was the most commonly reported effect size index among included studies; and *b*) the measure enabled the pooling of dichotomous and continuous effect sizes (e.g., mean difference and odds ratio) in the same meta-analysis.[47,48] For studies that measured the mean difference between more- and less-exposed men, the review team relied on the primary study authors' exposure group designations, irrespective of the statistical significance of examined differences in exposure levels between exposure and comparator groups, if applicable. Study-specific Hedges' g effect sizes are presented in a forest plot, whereas pooled Hedges' g (G_{Pooled}) estimates are presented in tabular format. To generate the forest plot, study-specific Hedges' g effect sizes were exported from R and imported into Tableau Public (version 2022.4.0) for visualization.

Guidance from Bakker et al.[49] was used to interpret the magnitude of the G_{Pooled} estimate in the context of the study question

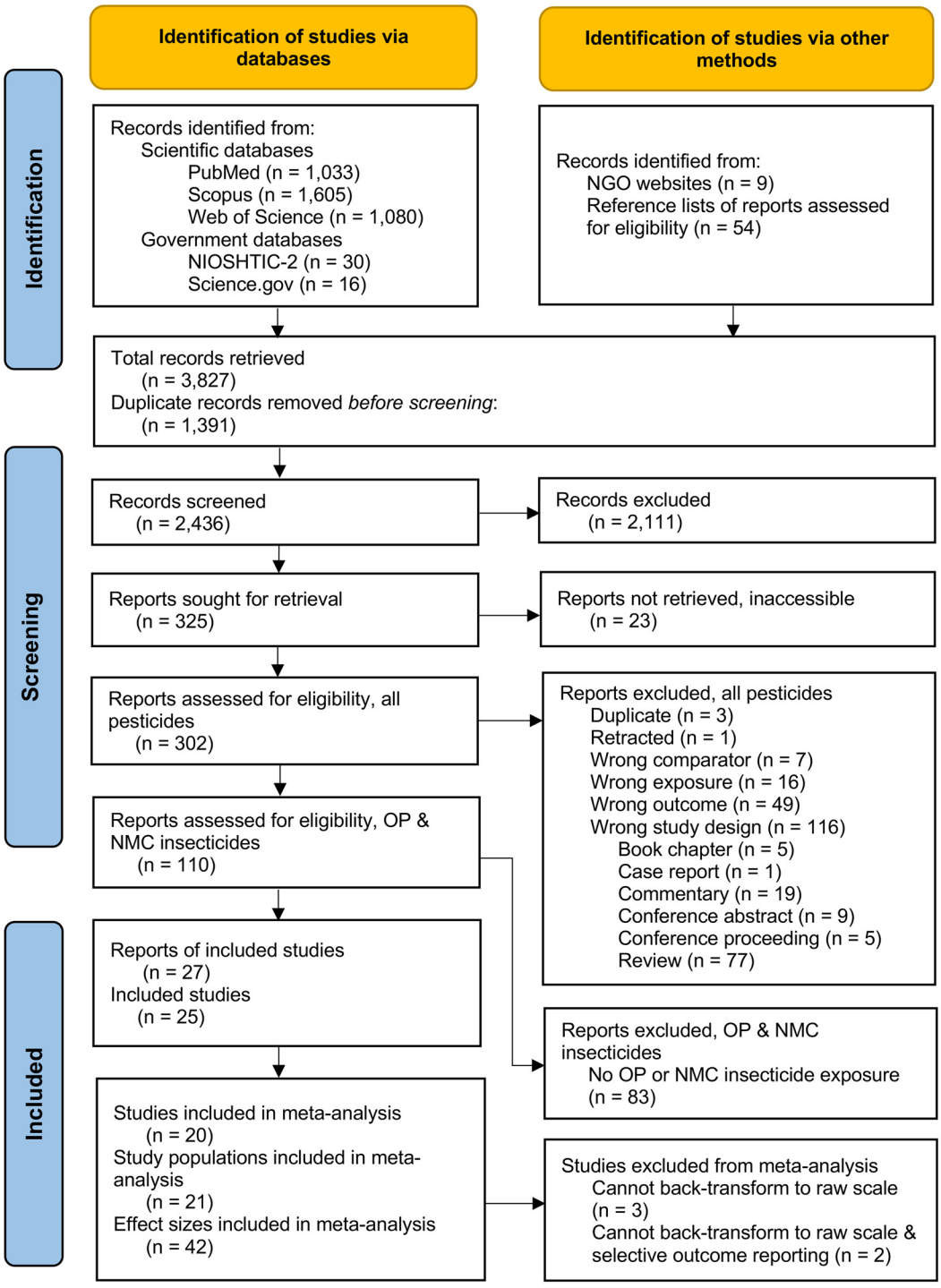


Figure 1. PRISMA flow diagram of the identification, screening, and selection of studies on adult organophosphate (OP) and *N*-methyl carbamate (NMC) insecticide exposure and sperm concentration. "Record" refers to a title and abstract. "Report" refers to a full-text journal article, government report, or other document supplying relevant information about a particular study or set of studies. See Table S2 for literature search terms and search dates, Table S3 for NGO search strategies, Table S4 for full-text (all pesticides) screening criteria, and Table S5 for OP and NMC insecticide screening criteria. Regarding "Reports sought for retrieval," see Excel Table S1 for a complete list of records sought for retrieval and their final inclusion decision and rationale. Note: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

rather than relying solely on commonly used benchmark values (0.2, small; 0.5, medium; 0.8, large).[50] In the context of environmental and reproductive health, the review team considered a G_{Pooled} value >0.5 or a difference in sperm concentration equal to one-half of the sample-specific pooled standard deviation to be large enough to warrant an upgrade in the quality of evidence. Despite this, magnitudes of effect below this cutoff may still be significant from a population health perspective as a result of widespread global use of and exposure to OP and NMC insecticides.[27]

Meta-analysis eligibility. Study-reported results were eligible for inclusion in the meta-analysis if a Hedges' *g* effect size could be calculated on the raw scale. Studies were excluded from the meta-analysis if *a*) results were not reported quantitatively as a result of

selective outcome reporting, *b*) Hedges' *g* could not be calculated, or *c*) results were reported on a statistically transformed scale that could not be back-transformed to the raw scale (Table S8).

Data preparation. **Mean difference data.** The standardized mean difference (Cohens' *d*) was calculated from study-reported mean difference (*MD*) data according to Equations 1 and 2,[47,51] where *MD* is the difference in the unadjusted or adjusted mean sperm concentration between two independent groups, S_{Pooled} is the within-group standard deviation pooled across groups, n_1 and n_2 are the sample sizes in the two groups, and S_1 and S_2 are the unadjusted or adjusted standard deviations in the two groups (with 1 and 2 representing the more- and less-exposed groups, respectively):

$$d = \frac{MD}{S_{Pooled}}, \tag{1}$$

$$S_{Pooled} = \sqrt{\frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{n_1+n_2-2}}. \tag{2}$$

In Equation 1, *MD* is either the difference in the unadjusted or adjusted sample means of independent two groups ($MD = Mean_1 - Mean_2$) or a beta coefficient from a linear regression model with a binary independent variable that represents exposure group membership (binary beta coefficient).

The variance of *d* (V_d) was calculated according to Equation 3,[47,51] and the standard error of *d* (SE_d) was calculated by taking the square root of V_d, so that $SE_d = \sqrt{V_d}$[47]:

$$V_d = \frac{n_1+n_2}{n_1 n_2} + \frac{d^2}{2(n_1+n_2)}. \tag{3}$$

Generally, the standard error (*SE*) of a given effect estimate was calculated according Equation 4,[52] where $Estimate_{Upper}$ is the upper bound and $Estimate_{Lower}$ is the lower bound of the 95% confidence interval (CI) around the estimate and *Z* is the *z*-value for a 95% CI, equal to 1.96:

$$SE_i = \frac{Estimate_{i_{Upper}} - Estimate_{i_{Lower}}}{2Z}. \tag{4}$$

In Equation 2, unadjusted or adjusted sample standard deviations were either directly reported or calculated by multiplying the sample *SE* by the square root of the sample size of the group, so that $S_i = SE_i\sqrt{n_i}$.[52] When dealing with adjusted mean difference data, unadjusted standard deviations were used to calculate S_{Pooled} where possible, given that this is recommended for calculating partial standardized mean differences.[47,51] Unadjusted standard deviations were sought from authors but were not required for inclusion in the meta-analysis. Of the six studies (with one study having three reports) included in the meta-analysis that reported adjusted mean difference data,[53–60] unadjusted standard deviations were available for two studies.[54,57,58,60]

If sample standard deviations or standard errors were not reported or calculable, Equation 5[53,51] was used to calculate *d*, where *t* is the *t*-statistic associated with a *t*-test comparison of group means or a binary beta coefficient:

$$d = \left(t\sqrt{\frac{n_1+n_2}{n_1 n_2}}\right). \tag{5}$$

t was calculated as either the absolute value of the mean difference (*MD*) divided by its standard error (SE_{MD}), so that $t_{MD} = MD/SE_{MD}$, or from the *t* distribution table,[52] so that $t = tinv(P, df)$, where *P* is the exact *p*-value of the effect estimate, *df* is the degrees of freedom $(n-p-1)$,[51] *p* is the number of predictors, and *n* is the total sample size. After calculating *d* using Equation 5, V_d was derived according to Equation 3.

When dealing with adjusted mean difference data using the *t*-statistic approach in Equation 5, it is recommended to multiple *d* by $\sqrt{1-R_{YZ}^2}$, as well as to integrate R_{YZ}^2 into the calculation of V_d to attempt to restore the original unadjusted standard deviation, which is reduced by the inclusion of one or more covariates in the regression model.[47,51] The term R_{YZ}^2 represents the proportion of explained variance in the regression model that includes one or more covariates, but excludes the grouping variable.[51] R_{YZ}^2 values were sought from authors but were not required for inclusion in the meta-analysis. R_{YZ}^2 values were not available in any the studies included in the meta-analysis.

If mean difference data was provided on the logarithmic (log) scale, Equations 6 and 7 were used to convert a difference in means on log scale (MD_{Log}) to an approximate difference on the raw scale (MD_{Raw}),[61] where $\bar{x}_{\text{geom}}$ is the geometric mean of the geometric means across groups (equivalent to the exponential of the arithmetic mean of the means of log-transformed values) and MD_{Log} and $SE_{MD_{Log}}$ are the mean difference and its standard error from log-transformed values.[61] This was done for two studies included in the meta-analysis[62,63]:

$$MD_{Raw} = MD_{Log} \times \bar{x}_{\text{geom}}, \tag{6}$$

$$SE_{MD_{Raw}} = SE_{MD_{\text{Log}}} \times \bar{x}_{\text{geom}}. \tag{7}$$

$SE_{MD_{Log}}$ was calculated according to Equation 8, where $MD_{Log_{Upper}}$ is the upper bound and $MD_{Log_{Lower}}$ is the lower bound of the 95% CI around the log scale estimate and *Z* is the *z*-value for a 95% CI, equal to 1.96:

$$SE_{MD_{Log}} = \frac{MD_{Log_{Upper}} - MD_{Log_{Lower}}}{2Z}. \tag{8}$$

If neither group-level standard deviations nor the *t*-statistic were reported or calculable, missing group-level standard deviations were imputed with the average of the available group-level standard deviations in each of the exposure groups. This imputation was done for one study included in the meta-analysis.[63]

Odds ratio data. Odds ratios (*OR*s), either unadjusted or adjusted (i.e., a beta coefficient from a multiple logistic regression model, or logistic beta coefficient), and their log variances [$V_{ln(OR)}$] were transformed to standardized mean differences (*d*) and variances of *d* (V_d) according to Equations 9 and 10,[47] where ln is the natural log (base *e*) and $V_{ln(OR)}$ is the natural log variance of the *OR*:

$$d = \frac{ln(OR)\sqrt{3}}{\pi}, \tag{9}$$

$$V_d = \frac{3V_{ln(OR)}}{\pi^2}. \tag{10}$$

Where appropriate, Equation 9 was multiplied by –1 to ensure *d* reflects the correct direction of effect,[47] so that *d* is negative where the *OR* input is >1, representing higher odds of having low sperm concentration, and *d* is positive where the *OR* input is <1, representing lower odds of having low sperm concentration.

If not directly reported, *OR*s were calculated from exposure and outcome prevalence data provided in the study using formulas presented in Borenstein and Hedges.[47] Specifically, *OR* was calculated as $OR = AD/BC$, where *A* is the number of exposed (or more-exposed) men with low sperm concentration, *B* is the number of exposed men without low sperm concentration, *C* is

the number of unexposed (or less-exposed) men with low sperm concentration, and D is the number of unexposed men without low sperm concentration. $V_{ln(OR)}$ was then calculated according to Equation 11[47]:

$$V_{ln(OR)} = \frac{1}{A} + \frac{1}{B} + \frac{1}{C} + \frac{1}{D}. \quad (11)$$

For studies that directly reported *OR*s, including in the form of a logistic beta coefficient, $V_{ln(OR)}$ was calculated according to Equation 12,[47] where OR_{Upper} is the upper bound and OR_{Lower} is the lower bound of the 95% CI around the estimate and Z is the z-value for a 95% CI, equal to 1.96:

$$V_{ln(OR)} = \left(\frac{OR_{Upper} - OR_{Lower}}{2Z}\right)^2. \quad (12)$$

Correlation data. Correlations (r) and the variances of r (V_r) were transformed to standardized mean differences (d) and the variances of d (V_d) according to Equations 13 and 14[47]:

$$d = \frac{2r}{\sqrt{1-r^2}}, \quad (13)$$

$$V_d = \frac{4V_r}{(1-r^2)^3}. \quad (14)$$

V_r was approximated according to Equation 15,[47] where r is the correlation and n is the total sample size:

$$V_r = \frac{(1-r^2)^2}{n-1}. \quad (15)$$

For studies that reported beta coefficients from a multiple linear regression model that included continuous exposure and outcome variables (continuous beta coefficient), partial correlations (r_p) and the variances of r_p (V_{r_p}) were calculated according to Equations 16 and [17,64] where df is the degrees of freedom equal to $n-p-1$. Partial correlations describe the magnitude of an effect after controlling for the influence of other variables included in a model[64]:

$$r_p = \frac{t}{\sqrt{t^2 + df}}, \quad (16)$$

$$V_{r_p} = \frac{(1-r_p{}^2)^2}{df}. \quad (17)$$

In Equation 16, t is the t-statistic of the continuous beta coefficient $\beta_{Exposure}$ (i.e., the test of null hypothesis that $\beta_{Exposure} = 0$), df is the degrees of freedom, calculated as $df = n-p-1$, where p is the number of predictors included in the model. t was calculated as either the absolute value of the continuous beta coefficient ($\beta_{Exposure}$) divided by its standard error ($SE_{\beta_{Exposure}}$), so that $t_{\beta_{Exposure}} = \beta_{Exposure}/SE_{\beta_{Exposure}}$, or from the t distribution table, so that $t = tinv(P, df)$, where P is the exact p-value of the effect estimate and df is the degrees of freedom ($n-p-1$).[64]

Studies reported continuous beta coefficients on both the raw and log scales. Because it is not recommended to pool log and raw scale data in the same meta-analysis,[61] methods presented by Rodríguez-Barranca et al.[65] were applied to homogenized log scale linear beta coefficients prior to calculating r_p and V_{r_p}. Continuous beta coefficients were homogenized to represent an absolute (raw scale) change in sperm concentration for every 1-unit increase in exposure as measured by the study using a downloadable Excel spreadsheet embedded with various formulas provided by the study authors.[65] This was done for two studies included in the meta-analysis.[66,67] Arithmetic mean exposure and outcome levels for the total study sample were needed to homogenize beta coefficients, but these values were not directly reported; missing arithmetic mean values were imputed with either the average of the arithmetic mean values across groups[66] or geometric mean values.[67] After homogenization, r_p and V_{r_p} were calculated according to Equations 16 and 17 based on the t-statistic associated with the homogenized continuous beta coefficient.

Hedges' *g* bias correction. The bias-adjusted standardized mean difference (Hedges' g) was calculated by multiplying the standardized mean difference (d), which has a slight bias, by a correction factor [$J(df)$] calculated according to Equation 18,[47,51] so that $g = d \times J(df)$. The correction factor is a function of the degrees of freedom (df), where $df = n-p-1$,[51] n is the total sample size, and p is the number of predictors included in the statistical model:

$$J(df) = 1 - \left(\frac{3}{4df - 1}\right). \quad (18)$$

The variance of g (V_g) was calculated by multiplying the variance of d by the square of the correction factor, so that $V_g = V_d \times [J(df)]^2$,[47,51] and the standard error of g (SE_g) was calculated by taking the square root of V_g so that $SE_g = \sqrt{V_g}$.[47]

It should be noted that unadjusted and adjusted values were used to calculate Hedges' g in the above equations, leading to a combination of bivariate and partial effect sizes included in the meta-analysis. Although not all possible covariates were assessed, the impact of controlling for two key confounders, age and smoking, was explored in a sensitivity analysis. Moreover, in some cases, the formulas[47] used to transform *OR*s and correlations to standardized mean differences were applied to partial effect sizes, such as adjusted *OR*s or partial correlation coefficients derived from regression models. To our knowledge, the use of these formulas on partial effect sizes has not yet been validated in the literature. Transformed effect sizes were removed in a sensitivity analysis to examine the potential impact of these exploratory statistical transformations.

Study results data, including log to raw scale back-transformations and beta coefficient homogenization (sheet 2), as well as the R code used to calculate Hedges' g effect sizes are available on GitHub (https://github.com/Lauren-Ellis/Ellis-et-al.-2023-OP-NMC-Insecticides-and-Sperm-Concentration). The R code can also be found at the end of the Supplemental Material file, "R Code."

Statistical analysis. **Meta-analytical model.** Meta-analyses were performed using R statistical software (version 4.2.2; R Development Core Team) packages *metafor* (version 3.8.1),[68] *clubSandwich* (version 0.5.8),[69] and *dmetar* (version 0.0.9000).[70] For all meta-analyses, a three-level, multivariate random-effect model[71] with cluster-robust variance estimation was employed to account for both hierarchical and correlational dependencies in the effect size data.[72–74] Weights were calculated using the generic inverse variance method.[75] The meta-analysis R code is available in the Supplemental Material file ("R Code") as well as on GitHub (https://github.com/Lauren-Ellis/Ellis-et-al.-2023-OP-NMC-Insecticides-and-Sperm-Concentration).

Effect sizes were nested under their respective study populations for meta-analysis. The review team assumed a strong correlation ($r = 0.8$) between dependent effect sizes originating from, and nested under, fully or partially overlapping study participants (the study population cluster). A cluster-robust variance estimator (robust to the assumed correlation between dependent effect sizes within the same cluster) with bias-reduced linearization small-sample adjustment based on Satterthwaite approximated degrees

of freedom was employed to reduce the chance of a Type I error in hypothesis testing.[76–79] A Satterthwaite-adjusted p-value (P_{Satt}) of ≤ 0.05 was considered statistically significant.

Heterogeneity point estimates (τ^2) were estimated using restricted maximum likelihood estimation (REML)[75,80] for each of the three levels of random-effects in the model: *a*) sampling error, *b*) within-cluster heterogeneity, and *c*) between-cluster heterogeneity.[81] Statistical significance of heterogeneity beyond sampling error was tested using the Q-test statistic based on a chi-square distribution.[82,83] Corresponding I^2 statistics were calculated using the var.comp function available in *dmetar*,[70] representing the proportion of total heterogeneity that can be attributed to each level of random effect.[84]

REML heterogeneity point estimates (τ^2) are known to be imprecise when the number of studies included in the meta-analysis is small (<10 studies), or when the sample sizes of the included studies are small.[80] Therefore, sensitivity analyses fixing heterogeneity values to the upper- and lower-bounds of their Q-profile CI[85] were performed to assess the impact of the heterogeneity estimate value on the pooled effect estimate.

Primary meta-analysis. A primary meta-analysis was performed across all eligible studies and effect sizes. Effect sizes were considered outliers if they fell outside of Tukey's fences[86] (defined as 1.5 times the lower and upper bounds of the interquartile range) and influential if their Cook's distance was at least three times the mean Cook's distance across included studies.[87] Cook's distance estimates represent the scaled change in fitted values resulting from the removal of each study (or other unit of analysis) from the model fitting.[87] Further, a prediction interval around the primary pooled effect estimate was calculated to examine where the true effects would be expected for 95% of similar studies that may be performed in the future.[75,88,89]

Secondary meta-analyses. A series of secondary meta-analyses were performed to assess the robustness of the meta-analysis findings. Specifically, data sensitivity analyses were performed to assess the impact of outlier and influential effect sizes, meta-analytic model parameters, statistical imputations, and statistical transformations. Moderator sensitivity analyses were also performed to assess the impact of risk of bias and potential effect modifiers, including coexposures to other chemicals, medical risk factors, and abstinence time. We considered these factors to be potential effect modifiers rather than confounders based on a lack of evidence that these factors lead to higher insecticide exposure. Finally, moderator subgroup analyses were performed to explore potential sources of heterogeneity and to test the hypothesis that the pooled effect estimate significantly differs by the following subgroups: insecticide class (OP, NMC, mixture), exposure setting (occupational, environmental), recruitment setting (general population, infertility clinic), and continent (Asia, North America, South America, Europe). Subgroup analysis was conducted rather than meta-regression based on the assumption that heterogeneity varies between subgroups.[90]

A given meta-analysis was considered to have sufficient statistical power if there was at least 4 degrees of freedom after Satterthwaite-adjustment ($df_{Satt} > 4$), as the probability of a Type I error can be substantially larger than the significance level ($\alpha = 0.05$) below this cutoff.[79] For subgroup analyses in which both subgroups were sufficiently powered and statistically independent, the review team planned to compare the pooled effect estimates of each subgroup using a fixed-effect meta-regression model and an F-test with separate heterogeneity estimates.[90,91]

Publication bias. A modified Egger's meta-regression test[92] using cluster-robust variance estimation was used to examine the impact of study precision on study-reported effect sizes (small-study effects), a known source of publication bias.[93] A sample size-based precision estimate ($1/\sqrt{n}$) was assessed as the predictor rather than the standard error to avoid known distortions in funnel plot assessments of standardized mean differences and standard errors.[94] Statistical adjustments for publication bias were not performed because methods such as trim and fill are not yet available for multivariate meta-analyses.[95] It should be noted that meta-regression tests for small-study effects have limited statistical power and account for only one potential source of publication bias, requiring cautious interpretation.[93]

Rating the Quality and Strength of Evidence across Studies

The quality of the overall body of evidence was rated as high, moderate, or low. An initial quality rating of moderate was assumed, consistent with the Navigation Guide's approach to observational human evidence.[38] The review team considered whether to downgrade the quality of evidence one or two levels according to five factors: *a*) risk of bias across studies, *b*) indirectness, *c*) inconsistency, *d*) imprecision, and *e*) publication bias.[39,43] The review team considered whether to upgrade the quality of evidence one or two levels according to three factors: *a*) large magnitude of effect, *b*) dose–response, and *c*) confounding minimizes effect.[39,43] Table S9 presents specific considerations for each of these quality factors.

Strength of evidence of an association between adult OP and NMC insecticide exposure and sperm concentration was determined according to four factors: *a*) quality of the body of evidence, *b*) direction of the pooled effect estimate, *c*) confidence in the pooled effect estimate, and *d*) other compelling attributes of the data that may influence certainty. Table S10 presents specific considerations for each of these strength factors. The Navigation Guide strength of evidence definitions[38,39,43] (Table S11) are based on categories used by the International Agency for Research on Cancer,[96] the U.S. Preventive Services Task Force,[97] and U.S. Environmental Protection Agency[98,99] and reflect the level of certainty in the overall toxicity of adult OP and NMC insecticide exposure on sperm concentration.

Two review authors (L.B.E. and K.M.) independently rated the quality and strength of evidence and resolved discrepancies through discussion. The full review team then met to review, discuss, and finalize the quality and strength of evidence ratings and rationales. Individual and collective rating rationales were recorded throughout the deliberation process. Final ratings and rationales are presented in tabular format.

Results

Study Selection

Of the 3,827 records retrieved in the literature search, 325 reports were sought for retrieval (Figure 1). All reports sought for retrieval and their corresponding screening results are provided in Excel Table S1. 23 reports were inaccessible, largely owing to a lack of digitization of studies conducted prior to the year 2000. Of the 302 accessible reports assessed for eligibility, 27 reports were included in the systematic review,[53–60,62,63,66,67,100–114] representing 25 unique studies.

Padungtod et al.[57,58] and Lin et al.[54] were treated as three reports on the same study based on identical analyses and were combined as a single study prior to data extraction. In contrast, Perry et al.[62,109] were treated as two separate studies on the same study population based on differences in the exposure assessment. Whorton et al.[112] and Wyrobek et al.[59] were also treated as two separate studies on the same study population based on differences in the control groups. Separate studies on the same study population were not combined prior to data extraction but, rather,

nested under their respective study populations for meta-analysis. Padungtod et al.[107] reported separate results for three study populations based on genotype and Swan et al.[66] reported separate results for two study populations based on geography.

Study Characteristics

Summary statistics for selected characteristics of included studies are reported in Table 1. More detailed study characteristics are presented in Table 2 and the complete set of extracted study characteristics with notes is available in Excel Table S2. Of the 25 studies included in this review, 21 studies were cross-sectional,[53–60,62,63,67,100–108,112–114] 2 were case-control,[66,109] and 2 were cohort by design.[110,111] Included studies were carried out all over the globe, with about one-third of them performed in Asia (China,[54,57,58,62,63,107,109,113] India,[104] Iran,[101] Malaysia[102]). The remaining studies were conducted in North America (Mexico,[108,110,111] United States[59,66,105,112]), South America (Brazil,[100] Guadeloupe,[56] Peru,[60,114] Venezuela[55,106]), and Europe (Denmark,[103] Poland,[53] Spain[67]).

Over half (60%) of the studies included in this review examined exposure to OP insecticides only.[54,57,58,60,62,66,67,100–102,104,107–111,114] Four studies examined exposure to NMC insecticides only, specifically carbaryl.[59,63,112,113] In addition, three studies examined exposure to a mixture of OP and NMC insecticides, defined by exposure to parent insecticides in both classes[56] or cholinesterase monitoring.[55,106] The remaining three studies examined exposures to both OP and NMC insecticides, performing separate analyses to assess each class in isolation.[53,103,105]

More studies assessed occupational exposures (72%)[54–60,63,100–102,104,106–108,110–114] than environmental exposures (28%).[53,62,66,67,103,105,109] Biomonitoring of insecticide metabolites (44%)[53,54,57,58,62,66,67,104,105,109–111,114] and proxy methods (i.e., job exposure matrix,[108] food frequency questionnaire,[103] environmental monitoring including dermal residue measurements,[59,63,107,112,113] and cholinesterase monitoring[55,100,101,106]) (44%) were the most commonly applied exposure assessment methods across the included studies; self-report exposure assessment methods were the least common (12%).[56,60,102] More studies recruited participants from population-based settings (84%)[54–60,62,63,66,100–103,106–114] than infertility clinic-based settings (16%).[53,67,104,105] See Table 2 for more detailed recruitment and exposure characteristics of included studies.

Less than half (44%) of the included studies accounted for abstinence time,[53,54,56–58,66,67,103,105,108–111] whereas the majority (76%) of studies accounted for medical risk factors for low sperm concentration[53–59,66,67,100–107,109–111,113] regardless of recruitment setting. Most studies (68%) accounted for known or measured coexposures to other reproductive toxicants.[53–60,67,101,102,104,105,107–109,112–114]

Table 1. Summary statistics of included studies on adult organophosphate (OP) and *N*-methyl carbamate (NMC) insecticide exposure and sperm concentration.

Study characteristic	Studies [*n* (%)]
Total	25 (100)
Study design	
Cross-sectional	21 (84)
Case-control	2 (8)
Cohort	2 (8)
Continent[a]	
Asia	9 (36)
North America	7 (28)
South America	6 (24)
Europe	3 (12)
Insecticide class	
OP	15 (60)
NMC	4 (16)
OP, NMC[b]	3 (12)
OP and NMC[c]	3 (12)
Exposure setting	
Occupational	18 (72)
Environmental	7 (28)
Exposure assessment method	
Biomonitoring	11 (44)
Proxy	11 (44)
Self-report	3 (12)
Recruitment setting	
Population	21 (84)
Infertility clinic	4 (16)
Accounted for abstinence time	
Yes	11 (44)
No	14 (56)
Accounted for medical risk factors	
Yes	19 (76)
No	6 (24)
Accounted for known or measured coexposures	
Yes	17 (68)
No	8 (32)

Note: See Excel Table S2 for the complete set of extracted data and related notes for all studies included in this review.
[a]Countries and regions represented on each continent: Asia (China, India, Iran, Malaysia), North America (Mexico, United States), South America (Brazil, Guadeloupe, Peru, Venezuela), and Europe (Denmark, Poland, Spain).
[b]"OP, NMC" refers to studies that assessed and reported results on OPs and NMCs separately and in isolation from one another; studies with this tag contributed to the separate OP and NMC subgroup meta-analyses.
[c]"OP and NMC" refers to studies that assessed and reported results on exposure to a mixture of OPs and NMCs or cholinesterase activity.

Risk of Bias

Risk of bias assessment ratings and rationales across all domains for each study are presented in Excel Table S3. A significant degree of risk of bias was present in the recruitment strategy, blinding, and exposure assessment domains (Figure 2A). These trends remained consistent after stratifying studies by meta-analysis inclusion status (Figure 2B).

Four studies were considered to have a high risk of recruitment strategy bias[60,102,103,114] based on likely or known differences in exposure or outcome between participants and nonparticipants. Eleven studies were considered to have a probably high risk of blinding bias[55,60,63,67,100,101,104,106,108,111,112] based on lack of reporting on whether investigators were blind to exposure and outcome status. Three studies were considered to have high risk of exposure assessment bias[56,60,102] based on their reliance on self-report exposure assessment methods, whereas 11 studies were considered to have probably high risk of exposure assessment bias[55,59,63,100,101,103,106–108,112,113] owing to reliance on proxy exposure assessment methods. The remaining 11 biomonitoring studies were determined to have either probably low[53,54,57,58,62,66,67,104,105,109,114] or low[110,111] risk of exposure assessment bias, depending on whether temporal variability could be established.

The review team considered Whorton et al.[112] and Wyrobek et al.[59] to have a high risk of outcome assessment bias. In both studies, which analyzed the same group of exposed men, semen samples were collected off-site (at home) and not necessarily delivered to the lab within 2 h, in contradiction to the methods outlined in the WHO semen analysis manual.[11] It is worth noting that these studies were conducted prior to the large-scale adoption of the WHO semen manual, originally published in 1980.

Whorton et al.[112] and Wyrobek et al.[59] were also considered to have a high risk of conflict of interest based on affiliations with and contributions from employees of Union Carbide Corporation, who manufactured carbaryl and other carbamate pesticides at the time the studies

Table 2. Selected characteristics of included studies on adult organophosphate (OP) and *N*-methyl carbamate (NMC) insecticide exposure and sperm concentration.

Study	Study design	Study participants	Place of recruitment	Insecticide class[a]	Exposure matrix (ces)[b]	Relevant exposure definitions (LOD; LOQ; DF)[c]
Occupational						
Self-report						
Hossain et al.[102]	CS	152 male farmers (62 exposed, 90 unexposed)	3 rural communities (Kungasang, Telpuid, and Papar districts), Sabah, Malaysia	OP	Questionnaire	Malathion-exposed farmers
Yucra et al.[60]	CS	111 male workers (31 exposed OP pesticide applicators, 80 unexposed men)	Agricultural workplaces and municipality (unexposed friends, city council, medical center, or hotel), Majes Valley, southern Peru	OP	Questionnaire	OP pesticide applicator, years of OP pesticide exposure, number of OP pesticide applications in the preceding month
Multigner et al.[56]	CS	87 male workers (42 exposed banana plantation workers, 45 unexposed manual and office workers)	Guadeloupe, an island in the French West Indies	OP and NMC	Questionnaire	Banana plantation worker exposed to one or more of the following pesticides: cadusaphos, ethoprophos, isazophos, pyrimiphos-ethyl, terbufos, aldicarb
Proxy						
Pérez-Herrera et al.[108]	CS	54 healthy, chronically exposed male agricultural workers	Agricultural community, Yucatan State, southeastern region of Mexico	OP	Job exposure matrix	OP exposure index
Cremonese et al.[100]	CS	135 randomly selected young males (99 rural men, 36 urban men)	Municipality of Farroupilha, Rio Grande do Sul, the southernmost state of Brazil	OP	Blood	AChE, BChE
Ghafouri-Khosrowshahi et al.[101]	CS	60 randomly selected males (30 OP-exposed rural farmers, 30 unexposed urban men)	Rural and urban communities, Hamadan Province (Karafs region), Iran	OP	Serum	BChE
Miranda-Contreras et al.[106]	CS	99 male workers (64 agricultural workers, 35 unexposed healthy men)	Rivas Dávila Municipality and the Mérida City, State of Mérida, Venezuela	OP and NMC	Plasma	AChE, BChE
Mármol-Maneiro et al.[55]	CS	59 male workers (29 exposed health department employees, 30 unexposed police department employees)	Department of Malariology of the Ministry of Health and Social Development of the Maracaibo Municipality and Zulia State Police Department, Zulia State, Venezuela	OP and NMC	Blood	ChE
Padungtod et al.[107]	CS	42 male factory workers (20 exposed pesticide workers, 22 unexposed textile workers)	OP pesticide manufacturing plant and nearby textile factory, Anqing, China		Plasma, personal air	AChE, ethylparathion, methamidophos
Whorton et al.[112]	CS	137 male workers (47 exposed, 90 unexposed)	Carbaryl production plant and unspecified Environmental Health Associates (EHA) study locations, USA	NMC	Indoor air, personal air	Carbaryl
Wyrobek et al.[59]	CS	84 male workers (50 exposed, 34 unexposed)	Carbaryl production plant, USA	NMC	Indoor air, personal air	Carbaryl
Tan et al.[63]	CS	99 male workers (31 exposed, 46 internal controls, 22 external controls)	Carbaryl production plant and Center for Disease Control and Prevention, Changzhou, China	NMC	Indoor air, personal air, dermal	Carbaryl
Xia et al.[113]	CS	46 healthy male workers (16 exposed, 12 internal controls, 18 external controls)	Carbaryl production plant and other professional settings, Changzhou, China	NMC	Indoor air, personal air, dermal	Carbaryl
Biomonitoring						
Padungtod et al.[57,58]; Lin et al.[54]	CS	43 male factory workers (20 exposed pesticide workers, 23 unexposed textile workers)	OP pesticide manufacturing plant and nearby textile factory, Anqing, China	OP	Urine	PNP (LOD/DF: NR)
Yucra et al.[114]	CS	62 male workers (31 exposed OP pesticide applicators, 31 unexposed men)	Agricultural workplaces and municipality (unexposed friends, medical center, or hotel), Majes Valley, southern Peru	OP	Urine	DMP (LOD: 5 μg/L; DF: 79%), DMTP (LOD: 5 μg/L; DF: 69%), DMDTP (LOD: 10 μg/L; DF: 10%), DEP (LOD: 5 μg/L; DF: 21%), DETP (LOD: 5 μg/L; DF: 19%), DEDTP (LOD: 10 μg/L; DF: 44%)

Table 2. (*Continued.*)

Study	Study design	Study participants	Place of recruitment	Insecticide class[a]	Exposure matrix (ces)[b]	Relevant exposure definitions (LOD; LOQ; DF)[c]
Manikandan et al.[104]	CS	100 male partners attending an infertility clinic for inability to conceive within 1 y	Jawaharlal Institute of Postgraduate Medical Education and Research hospital, Puducherry, India	OP	Urine	DMP (LOD: 10; LOQ: 33.4; DF: NR), DETP (LOD: 0.14; LOQ: 0.457; DF: NR), DEDTP (LOD: 3.06; LOQ: 10.2; DF: NR), ΣDAP (DF: 59%) (units NR)
Sánchez-Peña et al.[111]	C	33 randomly selected male agricultural workers ($n = 66$ semen samples)	Agricultural community of Villa Juarez, Durango, Mexico	OP	Urine	DMP (DF: 51%), DMTP (DF: 52%), DMDTP (DF: 52%), DEP (DF: 68%), DETP (DF: 74%), DEDTP (DF: 40%), ΣDAP [DF: 87% (at least 1 DAP)] (LOD: 2–3 μg/L)
Recio-Vega et al.[110]	C	52 long-term male residents of an agricultural community (19 pesticide sprayers, 16 agricultural workers, 17 nonoccupationally exposed men) ($n = 139$ semen samples)	Agricultural community of Villa Juarez, Durango, Mexico	OP	Urine	DMP (DF: NR), DMTP (DF: 26%), DMDTP (DF: 30%), DEP (DF: 48%), DETP (DF: 55%), DEDTP (DF: 11%), ΣDAP [DF: 83% (at least 1 DAP)] (LOD: 30–85 ng/mL)
Environmental Proxy						
Juhler et al.[103]	NCS	251 male farmers (166 with 0% organic intake, 39 with 1%–49% organic intake, 46 with 50%–100% organic intake)	Registers in the Danish Ministry of Agriculture	OP, NMC	Food frequency questionnaire	Acephate, azinphos-methyl, chlorpyrifos, diazinon, dimethoate, ethion, fenitrothion, mecarbam, methamidophos, methidathion, parathion, parathion-methyl, phenthoate, phosmet, pirimiphos-methyl, quinalphos, pyrazophos, phosalone, carbaryl
Biomonitoring						
Meeker et al.[105]	CS	330 male partners in subfertile couples seeking infertility diagnosis	Vincent Burnham Andrology Laboratory, Massachusetts General Hospital, Boston, Massachusetts, USA	OP, NMC	Urine	TCPY (LOD: 0.25 μg/L; DF: 94%), 1N (LOD: 0.40 μg/L; DF: 100%)
Dziewirska et al.[53]	NCS	315 males with normal sperm concentration (at least 15 million/mL) enrolled in the Environmental Factors and Male Fertility study	Infertility clinic, Lodz, Poland	OP, NMC	Urine	TCPY (DF: 100%), 1N (DF: 95%) (LOD: 0.5 μg/L)
Melgarejo et al.[67]	CS	116 male partners of infertile couples seeking infertility evaluation	Virgen de la Arrixaca University Hospital, Murcia region, southern Spain	OP	Urine	DMP (LOD: 0.1 μg/L; DF: 87%), DMTP (LOD: 0.1 μg/L; DF: 85%), DMDTP (LOD: 0.1 μg/L; DF: 70%), DEP (LOD: 0.1 μg/L; DF: 97%), DETP (LOD: 0.1 μg/L; DF: 98%); DEDTP (LOD: 0.01 μg/L; DF: 62%); ΣDAP
Perry et al.[62]	NCS	18 young males randomly selected from a larger cohort of newly married male partners of couples attending clinics for prenatal care	Agricultural regions and prenatal care clinics, Anhui Province, China	OP	Urine	DMP (LOD: 0.5 μg/L, DF: 28%), DMTP (LOD: 0.5 μg/L, DF: 89%), DMDTP (LOD: 0.25 μg/L, DF: 17%), DEP (LOD: 0.25 μg/L, DF: 17%), DETP (LOD: 0.25 μg/L, DF: 100%), DEDTP (LOD: 0.25 μg/L, DF: 11%), PNP (LOD: 0.14 μg/L, DF: 100%), MDA (LOD: 0.29 μg/L, DF: 6%), CMHC (LOD: 0.18 μg/L, DF: 17%), TCPY (LOD: 0.26 μg/L, DF: 0%), IMPY (LOD: 0.69 μg/L, DF: 6%), DEAMPY (LOD: 0.22 μg/L, DF: 0%), CIT (LOD: 1.5 μg/L, DF: 0%)

Table 2. (*Continued*.)

Study	Study design	Study participants	Place of recruitment	Insecticide class[a]	Exposure matrix (ces)[b]	Relevant exposure definitions (LOD; LOQ; DF)[c]
Perry et al.[109]	NCC	189 young males from a larger cohort of newly married male partners of couples attending clinics for prenatal care (94 cases, 95 controls)	Agricultural regions and prenatal care clinics, Anhui Province, China	OP	Urine	DMP (LOD: 0.25 μg/L; DF: 77%), DMTP (LOD: 0.25 μg/L; DF: 94%), DMDTP (LOD: 0.125 μg/L; DF: 18%); DEP (LOD: 0.125 μg/L; DF: 71%), DETP (LOD: 0.125 μg/L; DF: 96%), DEDTP (LOD: 0.125 μg/L; DF: 11%)
Swan et al.[66]	NCC	86 male partners of pregnant women enrolled in the Study for Future Families in Missouri (MO) and Minnesota (MN) (MO: 25 cases, 25 controls; MN: 9 cases, 27 controls)	Prenatal clinics affiliated with university hospitals, Minneapolis, MN and Columbia, MO, USA	OP	Urine	IMPY (DF: 73%), TCPY (DF: 72%), 1N (DF: 54%), PNP (DF: 11%), MDA (DF: 27%), carbofuranphenol (DF: 0%) (LOD: 0.1 μg/g creatinine)

Note: See Excel Table S2 for the complete set of extracted data and related notes for all studies included in this review. 1N, 1-naphthol; AChE, acetylcholinesterase; BChE, pseudocholinesterase; C, cohort; ChE, cholinesterase; CIT, 5-chloro-1,2-dihydro-1-isopropyl-[3H]-1,2,4-triazol-3-one; CMHC, 3-chloro-4-methyl-7-hydroxycoumarin; CS, cross-sectional; DAP, dialkylphosphate; DEAMPY, 2-diethylamino-6-methyl pyrimidin-4-ol; DEDTP, diethyldithiophosphate; DEP, diethylphosphate; DETP, diethylthiophosphate; DF, detection frequency; DMDTP, dimethyldithiophosphate; DMP, dimethylphosphate; DMTP, dimethylthiophosphate; IMPY, 2-isopropoxy-4-methyl-pyrimidinol; LOD, limit of detection; LOQ, limit of quantification; MDA, malathion dicarboxylic acid; NCC, nested case-control; NCS, nested cross-sectional; PNP, *para*-nitrophenol; ΣDAP, sum of dialkylphosphates; TCPY, 3,5,6-trichloropyridinol; USA, United States of America.

[a]"OP, NMC" refers to studies that assessed and reported results on OPs and NMCs separately and in isolation from one another; studies with this tag contributed to the separate OP and NMC subgroup meta-analyses. "OP and NMC" refers to studies that assessed and reported results on exposure to a mixture of OPs and NMCs or cholinesterase activity.

[b]Exposure assessment matrix(ces) correspond(s) to the most advanced exposure assessment method used in the study, regardless of the exposure variable included in the analysis of the relationship between exposure and outcome.

[c]Exposure definitions represent exposures examined in the exposure assessment, not necessarily analyzed in relation to sperm concentration; LOD, LOQ, and DF were extracted from biomonitoring studies only; LOD was listed at the end if applicable to all metabolites; LOQ was listed after LOD if available.

were conducted.[115] In addition, Multigner et al.[56] received a high risk of conflict of interest rating because it was funded in part by an unrestricted grant from the French Crop Protection Association, an agricultural lobbying group (http://www.agropages.com/CompanyDirectory/Detail-9232.htm).

Meta-Analysis Screening

Five studies were excluded from the meta-analysis (Figure 1; Excel Table S2). Three studies were excluded because they did not report effect sizes that could be back-transformed to the raw scale.[100,110,114] Two studies were excluded owing to the inability to back-transform results in addition to selective outcome reporting.[103,111]

Of the 20 studies included in the primary meta-analysis[53–60,62,63,66,67,101,102,104–109,112,113] (Figure 1; Excel Table S2), 13 studies reported results in the mean difference index[53–60,62,63,101,102,106,107,113] (Excel Table S2). Hedges' g effect sizes were directly calculated from mean difference data. The remaining study-reported results were transformed from either an unadjusted[112] or adjusted odds ratio[105,109] or a bivariate[104] or partial correlation coefficient[66,67,108] to standardized mean difference before converting to the bias-adjusted Hedges' g. Of the 12 authors or groups of authors contacted to retrieve data needed for meta-analysis, 1 author provided a previously inaccessible full-text article[63] and 3 authors responded that they no longer had access to the study data.[56,67,110] The remaining authors either could not be contacted owing to a lack of current contact information, did not respond to the inquiry, or were otherwise unable to provide the data requested.

Summary of Meta-Analysis Results

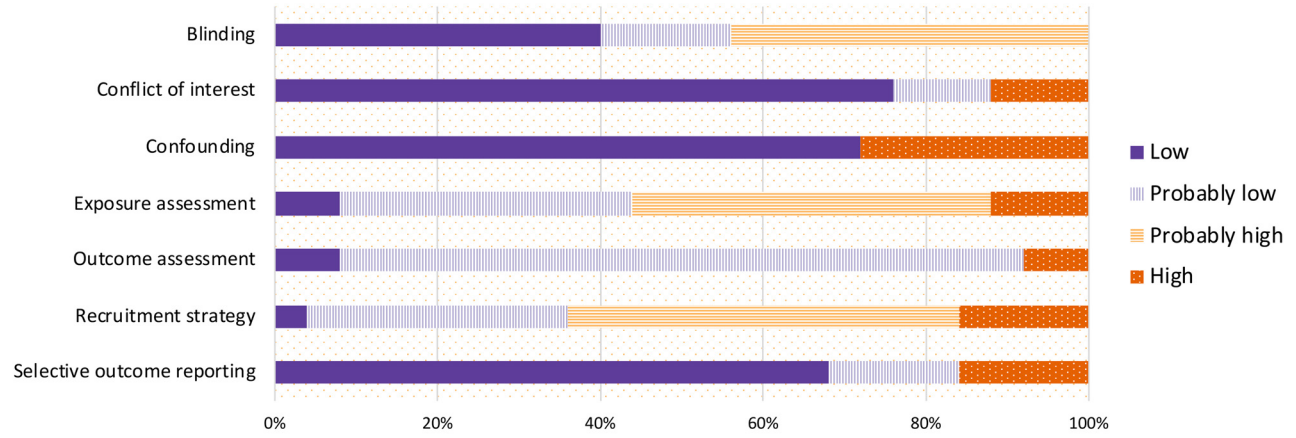
Primary meta-analysis results. Across 20 studies, 21 study populations, 42 effect sizes, and 1,774 adult men, the pooled bias-corrected standardized mean difference (G_{Pooled}) in sperm concentration between men more- and less-exposed to OP and NMC insecticides was −0.30 (95% CI: −0.49, −0.10; $P_{Satt} < 0.01$) (Table 3). A forest plot of study-specific Hedges' g effect sizes included in the primary meta-analysis, as well as their respective weights, is presented in Figure 3; an interactive version is available on Tableau Public (https://public.tableau.com/app/profile/lauren.ellis3833/viz/Ellisetal_2023ForestPlot_16952586826050/ForestPlot).

The prediction interval around the primary pooled effect estimate ranged from −0.95 to 0.36. This represents the expected range of results from future studies similar to those included in this review.

Sensitivity meta-analyses results. Removal of outlier and influential effect sizes from the primary meta-analysis did not impact the statistical significance of the observed association between higher OP and NMC insecticide exposure and lower sperm concertation (Excel Table S4). The primary pooled effect estimate was also robust to changing model parameters (the assumed correlation between dependent effect sizes and the heterogeneity estimate) as well as data manipulations (imputations of missing data, log to raw scale back-transformations, and effect size transformations) (Excel Table S4). However, fixing heterogeneity estimates at the lower-bounds of their Q-profile confidence intervals rendered the pooled effect estimate statistically insignificant ($P_{Satt} = 0.07$).

The removal of studies with high or probably high risk of bias in three domains (recruitment strategy, exposure assessment, and confounding) reduced the magnitude of the negative pooled effect estimate, whereas the removal of studies with high or probably high risk of bias in the other domains (blinding,

A



B

Study	Blinding	Conflict of interest	Confounding	Exposure assessment	Outcome assessment	Recruitment strategy	Selective outcome reporting
Included in meta-analysis							
Dziewirska (2019)[53]	Probably low	Low	Low	Probably low	Probably low	Probably high	Low
Ghafouri-Khosrowshahi (2019)[101]	Probably high	Low	High	Probably high	Probably low	Probably high	Low
Hossain (2010)[102]	Low	Probably low	Low	High	Probably low	High	Low
Manikandan (2021)[104]	Probably high	Low	High	Probably low	Probably low	Probably high	Low
Mármol-Maneiro (2003)[55]	Probably high	Probably low	Low	Probably high	Probably low	Probably high	Low
Meeker (2004)[105]	Low	Low	Low	Probably low	Probably low	Probably low	Probably low
Melgarejo (2015)[67]	Probably high	Low	Low	Probably low	Probably low	Probably low	Low
Miranda-Contreras (2013)[106]	Probably high	Low	High	Probably high	Probably low	Probably high	Low
Multigner (2008)[56]	Low	High	Low	High	Probably low	Low	Low
Padungtod (1999a[57]; 2000[58]); Lin (2000)[54]	Low	Probably low	Low	Probably low	Probably low	Probably low	Low
Padungtod (1999b)[107]	Low	Low	Low	Probably high	Probably low	Probably low	Low
Pérez-Herrera (2008)[108]	Probably high	Low	Low	Probably high	Probably low	Probably high	Probably low
Perry (2007)[62]	Low	Low	High	Probably low	Probably low	Probably low	Low
Perry (2011)[109]	Low	Low	Low	Probably low	Probably low	Probably low	Low
Swan (2003)[66]	Low	Low	Low	Probably low	Low	Probably low	High
Tan (2005)[63]	Probably high	Low	Low	Probably high	Probably low	Probably high	Low
Whorton (1979)[112]	Probably high	High	High	Probably high	High	Probably high	Low
Wyrobek (1981)[59]	Probably low	High	Low	Probably high	High	Probably high	Low
Xia (2005)[113]	Low	Low	Low	Probably high	Probably low	Probably high	Low
Yucra (2006)[60]	Probably high	Low	Low	High	Probably low	High	High
Excluded from meta-analysis							
Cremonese (2017)[100]	Probably high	Low	Low	Probably high	Probably low	Probably low	Probably low
Juhler (1999)[103]	Low	Low	High	Probably high	Probably low	High	High
Recio-Vega (2008)[110]	Probably low	Low	Low	Low	Low	Probably high	Probably low
Sánchez-Peña (2004)[111]	Probably high	Low	Low	Low	Probably low	Probably high	High
Yucra (2008)[114]	Probably low	Low	High	Probably low	Probably low	High	Low

Figure 2. Summary of risk of bias assessment ratings (A) across all included studies on adult organophosphate (OP) and *N*-methyl carbamate (NMC) insecticide exposure and sperm concentration and (B) for each study, stratified by meta-analysis inclusion status. The summary-level risk of bias horizontal bar chart (A) shows the cumulative frequency of each risk of bias rating option for each domain. Study is defined by the first author's last name and publication year followed by the numbered citation in superscript. See Excel Table S3 for the study-specific risk of bias assessment ratings and rationales used to generate this figure.

outcome assessment, selective outcome reporting, and conflict of interest) increased the magnitude of the negative pooled effect estimate (Table 3). The negative pooled effect estimate remained statistically significant across most of the risk of bias sensitivity analyses, except for the recruitment strategy and exposure assessment domains. Notably, the pooled effect estimate across studies that controlled for two key confounders (age and smoking)[53–60,63,66,67,100,102,105,107–111,113] ($G_{Pooled} = -0.27$; 95% CI: −0.48, −0.06; $P_{Satt} = 0.01$) proved similar to the primary meta-analysis result. The impact of risk of bias from incomplete outcome data was not statistically analyzed because the domain was only applicable to three studies.[66,110,111]

Regarding the effect modifier sensitivity meta-analyses, accounting for coexposures to other reproductive toxicants and medical risk factors for low sperm concentration increased the magnitude of the negative pooled effect estimate, which remained statistically significant (Table 3). In contrast, accounting for abstinence time reduced the magnitude of the negative pooled effect estimate, which was no longer statistically significant (Table 3). The lack of a statistically significant finding across studies that accounted for abstinence time ($P_{Satt} = 0.18$) should be interpreted with caution, given that the subgroup meta-analysis had limited statistical power ($df_{Satt} < 4$).

The pooled effect estimate remained negative in direction across the six studies that controlled for confounders (age and

Table 3. Results from primary and moderator (sensitivity and subgroup) meta-analyses examining the relationship between adult organophosphate (OP) and *N*-methyl carbamate (NMC) insecticide exposure and sperm concentration.

Category	Study populations [*n* (effect sizes)]	G_{Pooled} (cluster-robust 95% CI)	df_{Satt}	P_{Satt}	Between-cluster [τ^2 (I^2, %)]	Within-cluster [τ^2 (I^2, %)]	Q_H (*p*-value)
Total	21 (42)	−0.30 (−0.49, −0.10)	16.94	<0.01	0.08 (63)	0.02 (19)	<0.01
Risk of bias[a]							
Blinding	13 (26)	−0.31 (−0.60, −0.01)	10.19	0.04	0.11 (61)	0.05 (29)	<0.01
Recruitment strategy	10 (26)	−0.22 (−0.51, 0.07)	6.37	0.11	0.05 (59)	0.02 (21)	<0.01
Exposure assessment	8 (25)	−0.11 (−0.37, 0.15)	3.90	0.31	0.02 (32)	0.02 (39)	<0.01
Confounding[b]	18 (35)	−0.27 (−0.48, −0.06)	14.92	0.01	0.09 (77)	0.01 (6)	<0.01
Outcome assessment	20 (40)	−0.31 (−0.51, −0.10)	16.35	0.01	0.09 (70)	0.02 (13)	<0.01
Selective outcome reporting	18 (39)	−0.35 (−0.58, −0.13)	14.66	<0.01	0.11 (69)	0.02 (16)	<0.01
Conflict of interest	19 (39)	−0.31 (−0.53, −0.09)	15.63	0.01	0.11 (73)	0.02 (12)	<0.01
Effect modifiers accounted for[c]							
Abstinence time	9 (24)	−0.09 (−0.27, 0.10)	1.95	0.18	0.00 (0)	0.01 (31)	<0.01
Medical risk factors	18 (33)	−0.37 (−0.59, −0.15)	14.46	<0.01	0.10 (79)	0.01 (6)	<0.01
Coexposures	17 (34)	−0.37 (−0.60, −0.13)	13.85	0.01	0.12 (80)	0.01 (7)	<0.01
Insecticide class							
OP[d]	15 (30)	−0.32 (−0.59, −0.06)	11.97	0.02	0.12 (80)	0.01 (8)	<0.01
NMC[e]	5 (9)	−0.16 (−0.60, 0.28)	3.63	0.36	0.04 (30)	0.02 (17)	0.04
OP and NMC[f]	3 (3)	−0.50 (−1.34, 0.35)	1.97	0.13	0.03 (28)	0.03 (28)	0.11
Exposure setting							
Occupational	15 (19)	−0.43 (−0.71, −0.16)	12.60	0.01	0.12 (51)	0.02 (11)	<0.01
Environmental	6 (23)	−0.03 (−0.31, 0.25)	2.43	0.75	0.00 (12)	0.02 (50)	<0.01
Recruitment setting							
Population	17 (29)	−0.30 (−0.54, −0.07)	13.81	0.02	0.08 (49)	0.06 (38)	<0.01
Infertility clinic	4 (13)	−0.14 (−0.62, 0.33)	1.57	0.27	0.00 (0)	0.01 (18)	<0.01
Continent[g]							
Asia	10 (20)	−0.48 (−0.95, 0.00)	8.01	0.05	0.24 (77)	0.06 (18)	<0.01
North America	5 (10)	−0.20 (−0.56, 0.16)	3.78	0.20	0.00 (0)	0.02 (21)	0.06
South America	4 (4)	−0.37 (−0.96, 0.22)	2.98	0.14	0.04 (32)	0.04 (32)	0.05
Europe	2 (8)	−0.07 (−0.14, 0.01)	1.00	0.05	0.00 (0)	0.01 (23)	0.01

Note: The pooled Hedges' *g* (G_{Pooled}) represents the bias-adjusted standardized mean difference in sperm concentration between adult men more- and less-exposed to OP and NMC insecticides pooled across studies included in a given meta-analysis. The review team assumed a strong correlation ($r = 0.8$) between dependent effect sizes originating from, or nested under, fully or partially overlapping study participants (the study population "cluster"). A cluster-robust variance estimator (robust to the assumed correlation between dependent effect sizes within the same cluster) with bias-reduced linearization small-sample adjustment based on Satterthwaite approximated degrees of freedom (df_{Satt}) was employed. A Satterthwaite-adjusted *p*-value (P_{Satt}) of ≤0.05 was considered statistically significant. Heterogeneity point estimates (τ^2) were estimated using restricted maximum likelihood estimation (REML) for each level of random effect. Statistical significance of heterogeneity was tested using the *Q*-test statistic (Q_H) based on a chi-square distribution. Corresponding I^2 statistics represent the proportion of total heterogeneity that can be attributed to each level of random effect. CI, confidence interval; df, degrees of freedom; P_{Satt}, Satterthwaite-adjusted *p*-value.

[a]Removed studies with high or probably high risk of bias in each domain.

[b]Removed studies with high risk of confounding bias, defined as studies that did not account for both age and smoking.

[c]Removed studies that did not account for each effect potential modifier. "Coexposures" refer to known or measured exposures to other reproductive toxicants (besides OPs and NMCs) that may impact sperm concentration.

[d]Included studies that examined exposure to OP insecticides in isolation, regardless of whether other exposures were assessed in the study.

[e]Included studies that examined exposure to NMC insecticides in isolation, regardless of whether other exposures were assessed in the study.

[f]Included studies that examined exposure to a mixture of OP and NMC insecticides.

[g]Countries and regions represented on each continent: Asia (China, India, Iran, Malaysia), North America (Mexico, United States), South America (Brazil, Guadeloupe, Peru, Venezuela), and Europe (Denmark, Poland, Spain).

smoking) and effect modifiers (coexposures to other reproductive toxicants, abstinence time, and medical risk factors for low sperm concentration),[53,54,56–58,67,105,109] although it was reduced in magnitude and no longer statistically significant ($G_{Pooled} = -0.09$; 95% CI: −0.36, 0.18; $P_{Satt} = 0.25$). The lack of statistical significance could be due to limited power ($df_{Satt} = 1.68$). For this meta-analysis, overall heterogeneity beyond sampling error was statistically significant (Q_H, $p < 0.01$); 0% of overall heterogeneity was attributable to between-study population differences, whereas 37% was attributable to within-study population differences such as different exposure definitions assigned to the same study participants.

Subgroup meta-analyses results. All of the examined subgroup moderators—insecticide class, exposure setting, recruitment setting, and continent—substantially modified the observed pooled effect estimate (Table 3). Owing to limited statistical power of the examined subgroups, the pooled effect estimates for each of the subgroup analyses were not statistically compared. The discrepancies discussed below are observational in nature.

The negative pooled effect estimate across OP insecticide exposure studies was double the magnitude of that for NMC insecticide exposure studies ($G_{Pooled} = -0.32$; 95% CI: −0.59, −0.06; $P_{Satt} = 0.02$ vs. $G_{Pooled} = -0.16$; 95% CI: −0.60, 0.28; $P_{Satt} = 0.36$) (Table 3). Although the negative pooled effect estimate across NMC studies was not statistically significant, this may be explained by the fact that the subgroup had limited statistical power ($df_{Satt} < 4$). Notably, the negative pooled effect estimate across studies that assessed exposure to a mixture of OP and NMC insecticides demonstrated the largest pooled magnitude of effect among all meta-analyses performed ($G_{Pooled} = -0.50$; 95% CI: −1.34, 0.35; $P_{Satt} = 0.13$), although this subgroup of studies was also likely too small to detect a statistically significant association ($df_{Satt} < 4$).

The negative pooled effect estimate across occupational exposure studies was greater in magnitude than that of environmental exposure studies ($G_{Pooled} = -0.43$; 95% CI: −0.71, −0.16; $P_{Satt} = 0.01$ vs. $G_{Pooled} = -0.03$; 95% CI: −0.31, 0.25; $P_{Satt} = 0.75$) (Table 3). Although a significant association between higher OP and NMC insecticide exposure and lower sperm concentration was not discovered across environmental exposure studies, this subgroup meta-analysis lacked sufficient statistical power to draw any firm conclusions ($df_{Satt} < 4$).

The negative pooled effect estimate across studies that recruited adult men from infertility clinics was smaller in magnitude than that of studies that recruited adult men from population-based settings ($G_{Pooled} = -0.14$; 95% CI: −0.62, 0.33; $P_{Satt} = 0.27$ vs. $G_{Pooled} = -0.30$; 95% CI: −0.54, −0.07; $P_{Satt} = 0.02$) (Table 3).

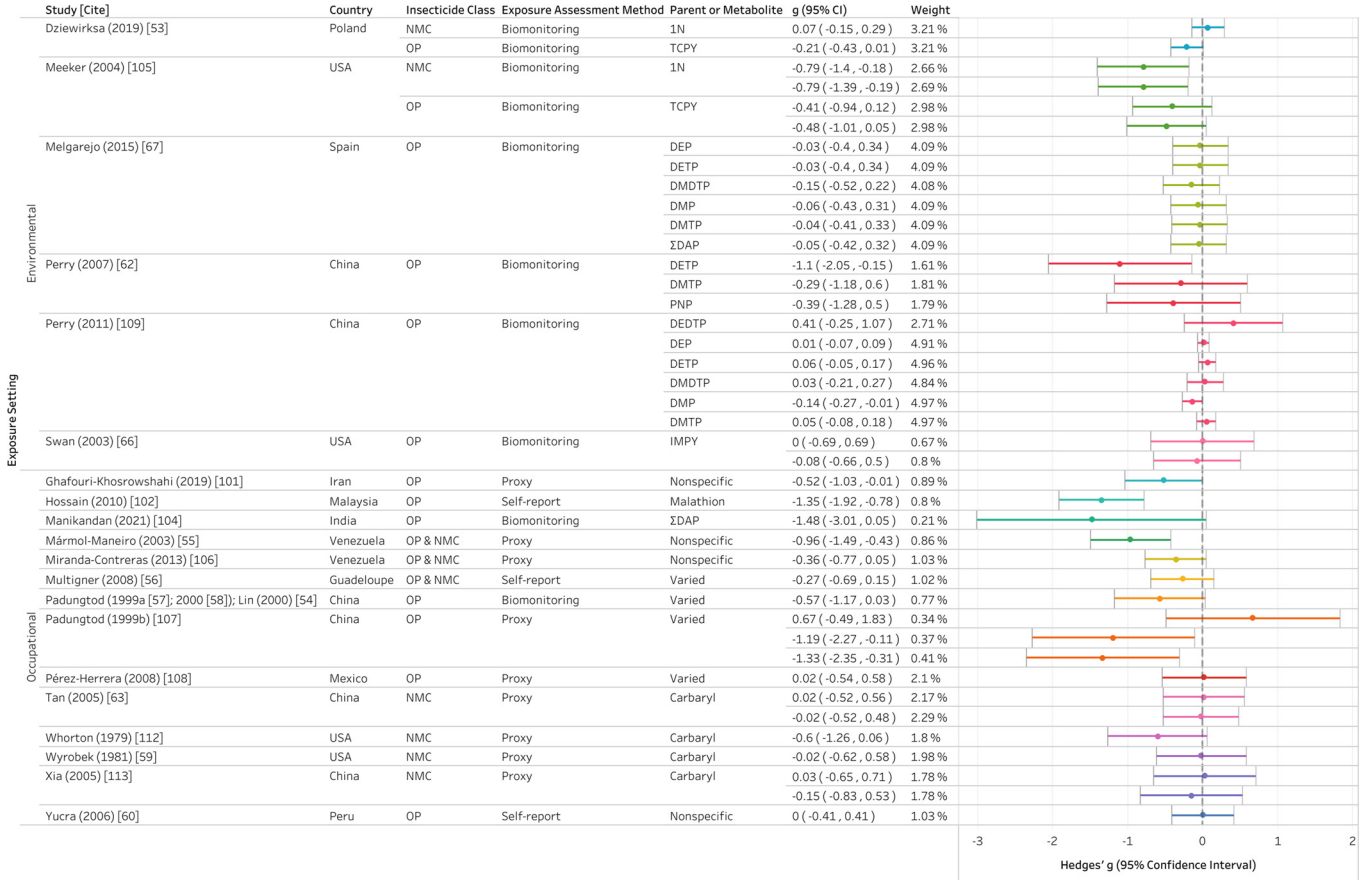


Figure 3. Forest plot of study-specific estimates of the bias-adjusted standardized mean difference in sperm concentration between adults more- and less-exposed to organophosphate (OP) and *N*-methyl carbamate (NMC) insecticides (Hedges' *g* effect sizes) included in the primary meta-analysis, stratified by exposure setting. Study is defined by the first author's last name and publication year followed by the numbered citation in brackets. "Parent" refers to the parent insecticide compound, whereas "metabolite" refers to the breakdown product of a parent compound. Weights of each study included in the primary meta-analysis were calculated using the generic inverse variance method. The extracted study results and R code used to calculate study-specific Hedges' *g* effect sizes and meta-analysis weights are available on GitHub. Estimated effect sizes were exported from R and imported into Tableau Public to generate this forest plot visualization. An interactive version of this forest plot is available on Tableau Public. Note: 1N, 1-naphthol; DAP, dialkylphosphate; DEDTP, diethyldithiophosphate; DEP, diethylphosphate; DETP, diethylthiophosphate; DMDTP, dimethyldithiophosphate; DMP, dimethylphosphate; DMTP, dimethylthiophosphate; IMPY, 2-isopropoxy-4-methyl-pyrimidinol; PNP, *para*-nitrophenol; ΣDAP, sum of dialkylphosphates; TCPY, 3,5,6-trichloropyridinol.

Like many of the other subgroups analyzed, the infertility clinic recruitment setting subgroup was insufficiently powered to detect a significant association ($df_{Satt} < 4$).

Pooled effect estimates were consistently negative in direction across continents but differed in magnitude and statistical significance (Table 3). The Asian study subgroup showed the largest negative pooled effect estimate ($G_{Pooled} = -0.48$; 95% CI: −0.95, 0.00; $P_{Satt} = 0.05$) and was the only geographic subgroup large enough to detect a statistically significant association ($df_{Satt} > 4$). The European study subgroup showed the smallest negative pooled effect estimate ($G_{Pooled} = -0.07$; 95% CI: −0.14, 0.01; $P_{Satt} = 0.05$), which reached statistical significance despite limited sample size.

Heterogeneity findings. The performed moderator sensitivity and subgroup meta-analyses did not sufficiently explain the estimated heterogeneity, or inconsistency in study results beyond sampling error (Table 3). Significant heterogeneity in study results beyond sampling error (Q_H, $p \le 0.05$) was apparent in all meta-analyses except for the OP and NMC insecticide subgroup (Q_H, $p = 0.11$) and the North American subgroup (Q_H, $p = 0.06$).

Certain subgroups showed no heterogeneity attributable to between-study population variation, including studies that accounted for abstinence time, studies that recruited men from infertility clinics, studies that controlled for confounders and effect modifiers, as well as studies conducted in North America and Europe (Table 3). These results indicate that differences in these variables between study populations may help explain at least some of the heterogeneity seen across the meta-analyses performed in this review.

Publication bias results. The review team did not find evidence of an association between-study precision ($1/\sqrt{n}$) and effect size ($P_{Satt} = 0.31$), ruling out small-study effects. It is possible that some degree of publication bias remains because small-study effects are only one source of publication bias. Moreover, one of the inaccessible reports identified in the systematic literature search relates to OP exposure and semen quality and could therefore be relevant to the results of this review; however, the abstract and report were unavailable.[116]

Summary of Results Excluded from the Meta-Analysis

Three OP occupational biomonitoring studies were excluded from the meta-analysis owing to the inability to back-transform relevant results to the raw scale,[110,114] in addition to selective outcome reporting.[111] All three studies found statistically insignificant associations between urinary dialkyl phosphate (DAP) metabolites in adult male farmers and sperm concentration. Despite not reaching levels of statistical significance, which may

be explained by limited sample size, two of these studies[110,114] reported lower sperm concentration among those with higher OP exposure, whereas the remaining study did not report the direction of effect.[111]

Two studies that used proxy exposure assessment methods to measure OP and NMC insecticide exposure were also excluded owing to the inability to back-transform relevant results to the raw scale,[100] in addition to selective outcome reporting.[103] Similar to the excluded biomonitoring studies, both studies found statistically insignificant associations between OP and NMC insecticide exposure and sperm concentration.

Most of the studies excluded from the meta-analysis accounted for abstinence time and medical risk factors, but only one excluded study[114] was free of known or measured coexposures to other reproductive toxicants. Risk of bias trends among studies excluded from the meta-analysis were similar to the entire body of evidence, with the recruitment strategy domain showing the greatest risk of bias across studies excluded from the meta-analysis (Figure 2B).

In addition to the studies excluded from the meta-analysis, relevant results for certain OP and NMC insecticide exposure definitions from studies included in the meta-analysis could not be quantified and were therefore excluded from the meta-analysis. For example, three relevant exposure definitions (3,5,6-trichloropyridinol, *para*-nitrophenol, and 1-naphthol) from Swan et al.[66] were not reported quantitively and thus could not be included in the meta-analysis. These unreported linear regression results (which were more adjusted than the odds ratio analysis results presented in Table 4 of the study by Swan et al.[66]) are presumably insignificant, given that the study authors reported results for only the five pesticides most associated with semen quality within Missouri study participants; however, the direction of effect for these metabolites is unclear. Moreover, it was not feasible to back-transform the statistic for DEDTP reported in Melgarejo et al.,[67] which demonstrated an insignificant association with lower sperm concentration. Finally, Lin et al.,[54] which was combined with Padungtod et al.[57,58] as a single study prior to analysis, analyzed the exposure and outcome data using two statistical models (maximum likelihood and working parameter), the results of which were consistent with the extracted results from the Padungtod et al.[58] report, demonstrating an association between occupational OP insecticide exposure and lower sperm concentration.

Quality and Strength of Evidence

The overall quality of evidence of an association between higher adult exposure to OP and NMC insecticides and lower sperm concentration was neither upgraded nor downgraded from moderate (Table 4). The overall strength of evidence was considered sufficient based on a moderate quality of evidence, consistent negative direction of effect across all sensitivity and subgroup analyses, confidence in the overall pooled effect estimate, and other compelling attributes such as statistical significance in nearly all sufficiently powered meta-analyses (Table 4).

Discussion

Key Findings

Understanding how OP and NMC insecticides affect sperm concentration is critical given their previously documented reproductive hazards and that exposures are ubiquitous. To our knowledge, this investigation is the most comprehensive systematic review on this topic to date, and the first to use multilevel and multivariate meta-analysis methods to quantitatively synthesize decades of epidemiological literature and statistically explore sources of heterogeneity. Based on the results of this investigation, the strength of evidence of an association between higher adult OP and NMC insecticide exposure and lower sperm concentration is sufficient enough to warrant concern, particularly in light of observed downward trends in semen quality.[1–6] Health-protective policy and engineering solutions are needed now to reduce exposures to OP and NMC insecticides and prevent continued male reproductive harm.

Results from the insecticide class subgroup analysis suggest that OP insecticides may present a greater risk to sperm concentration than NMC insecticides. Although this finding may be the result of the uniquely irreversible nature of OP cholinesterase inhibition,[117] the limited number of NMC studies prevents drawing any firm conclusions about whether NMCs are in fact safer than OPs regarding effects on sperm concentration.

As expected, occupational exposures demonstrated a stronger association with lower sperm concentration than environmental exposures. This finding represents a dose–response relationship given that workers exposed in occupational settings generally experience higher exposures than the general population.[118] It should be noted that the pooled effect estimate across environmental exposure studies was small and statistically insignificant, which may be related to the limited number of available environmental exposure studies. Nonetheless, given widespread exposure to these chemicals in the environment,[27] even a small magnitude of effect could have consequential impacts on sperm concentration at a population level.

Significant heterogeneity, or variation in study results beyond sampling error, was apparent across most sensitivity and subgroup meta-analyses performed. This was expected given the biological variability of the outcome, as well as the diverse study designs and exposure scenarios included in this review. Moreover, high statistical heterogeneity is more frequent in meta-analysis of continuous outcomes compared to that of binary outcomes.[119] Including all types of study designs and relying on author-determined groupings could explain some of this heterogeneity, given that there were differences in how different study authors recruited men into their study, measured exposure, and defined exposure groups. The magnitude of the REML heterogeneity point estimates should be interpreted with caution because they can be imprecise when the number of studies (i.e., <10 studies) or when the sample sizes of studies included in the meta-analysis are small.[80]

Context of Existing Literature

It is difficult to specify a single biological mechanism of effect that explains the adverse associations observed in this comprehensive review because mechanistic and animal evidence indicate there are many ways OP and NMC insecticides can harm sperm concentration. Many OP and NMC insecticides have been shown to directly interfere with hormone (e.g., androgen, estrogen) receptors[120] and damage cells in the testes through oxidative stress and genotoxic pathways. It is also important to investigate how cholinesterase-inhibiting insecticides may indirectly impact spermatogenesis by altering levels of neurotransmitters, such as acetylcholine in the brain, and disrupting the release of gonadotropins that directly influence sperm production.[106,121–123]

Although multiple pathways of effect are more likely than a single mechanism, this investigation demonstrates a clear association between higher adult OP and NMC insecticide exposure and lower sperm concentration, building on earlier qualitative reviews that have reached similar conclusions.[33–36] The only existing meta-analysis on OP pesticide exposure known to the review team at the time of this study similarly found that sperm concentration was significantly lower in more-exposed men.[124] However, a limited number of studies were included in the meta-analysis, and NMC insecticides were not considered. The present

Table 4. Summary of the quality and strength of evidence of an association between adult organophosphate (OP) and *N*-methyl carbamate (NMC) insecticide exposure and sperm concentration.

Evidence factor	Rating	Rationale
Quality of evidence factor		
Downgrade		
Risk of bias across studies	0	There was substantial risk of bias across included studies in three domains: recruitment strategy, blinding, and exposure assessment (Figure 2). The most influential study in the meta-analysis, Hossain et al.,[102] presented a similar pattern of bias, with high risk of bias in the recruitment strategy and exposure assessment domains. Nevertheless, the pooled effect estimate remained negative in direction across all risk of bias sensitivity analyses. Further, the removal of studies with high or probably high risk of bias in the blinding, outcome assessment, selective outcome reporting, and conflict of interest domains increased the magnitude of the pooled association between higher OP and NMC exposure and lower sperm concentration (Table 3). Given these findings, the review team ruled out bias as the primary explanation behind the observed association and thus did not downgrade the quality of evidence based on risk of bias.
Indirectness	0	All studies included in the review directly assessed the relevant population (adult men), exposure (OP and NMC insecticides), and outcome (sperm concentration) of interest. However, eight studies did not control for known or measured coexposures to other chemical reproductive toxicants.[62,63,66,100,103,106,110,111] Nevertheless, removing studies with known or measured coexposures from the meta-analysis increased the magnitude of the pooled association (Table 3). The review team therefore did not downgrade the quality of evidence based on indirectness.
Inconsistency	0	Study-reported effect sizes were generally consistent across studies included in the review, although outlier and influential effect sizes were identified. Removing these effect sizes did not impact the significance of the pooled association between higher OP and NMC exposure and lower sperm concentration, although the magnitude of the association was reduced after removing outlier effects (Excel Table S4). The primary meta-analysis indicated significant heterogeneity in study results beyond sampling error (Q_H, $p < 0.01$; $I^2 = 82\%$), which was not sufficiently explained by sensitivity and subgroup meta-analyses (Table 3). However, the review team expected a high degree of heterogeneity and thus did not downgrade the quality of evidence based on inconsistency.
Imprecision	0	A combination of narrow and wide confidence intervals around effect sizes included in the meta-analysis were observed through visual assessment of the forest plot (Figure 3). Those judged as wide[62,107,113] could largely be explained by small sample size. Moreover, the confidence intervals around the pooled effect estimates across primary and secondary meta-analyses were determined to be sufficiently narrow. Regarding the five studies excluded from the meta-analysis that did not find statistically significant associations, this may also be explained by small sample size. Thus, the review team did not downgrade the quality of evidence based on imprecision.
Publication bias	0	Study results were generally consistent, regardless of sample size or funding source. The statistical test for small-study effects (cluster-robust modified Eggers' regression test) demonstrated that study precision ($1/\sqrt{n}$) was not related to study results ($P_{\text{Satt}} = 0.31$). In addition, the comprehensive literature search performed helps safeguard against residual publication bias. Thus, the review team did not downgrade the quality of evidence based on publication bias.
Upgrade		
Large magnitude of effect	0	Most of the pooled effect estimates from the primary and secondary meta-analyses were determined to be moderate in magnitude. However, the pooled effect estimates for the combined OP and NMC insecticide class ($G_{\text{Pooled}} = -0.50$) and the occupational exposure setting ($G_{\text{Pooled}} = -0.43$) subgroups were rather large. Nevertheless, given the focus on OP and NMC insecticide exposure as a whole, the review team determined the magnitude of effect was not large enough to warrant an upgrade in quality of evidence.
Dose response	0	Dose–response relationships were established in some but not all of the included studies that modeled exposure continuously or in ordinal dose groups (Excel Table S2). The lack of a dose–response relationship in certain studies may be the result of limitations in study methods, particularly related to exposure assessment. Nonetheless, the exposure setting subgroup meta-analysis demonstrated a pooled dose–response relationship, given that occupational exposure studies showed a greater magnitude of association with lower sperm concentration than environmental exposure studies. Ultimately, the review team did not judge the available dose–response evidence as compelling enough to upgrade the quality of evidence.
Confounding minimizes effect	0	Some studies reported an association despite the presence of residual confounding, biases, or effect modification, whereas others did not. Thus, the review team determined that an upgrade in quality of evidence was not warranted.
Overall quality of evidence	Moderate	The quality of the body of evidence was neither downgraded nor upgraded from the initial rating of human evidence as moderate.
Strength of evidence factor		
Quality of body of evidence	NA	Moderate (see rationale above)
Direction of effect estimate	NA	An association between higher OP and NMC insecticide exposure and lower sperm concentration in adults was observed across a wide range of moderator meta-analyses.
Confidence in effect estimate	NA	The prediction interval around the primary pooled effect estimate (−0.95 to 0.36) contains null and positive values, indicating that future studies may not find an association between higher OP and NMC exposure and lower sperm concentration. However, the potential for a null future finding does not indicate that the primary pooled effect estimate would become obsolete. Given the number of studies included in the primary meta-analysis, it is reasonable to assume that the addition of a future influential study would not impact the observed association between higher OP and NMC insecticide exposure and lower sperm concentration. This assumption was verified by the fact that the addition of two relevant studies published prior to 1991[59,112] that contributed insignificant findings to the meta-analysis did not impact the negative direction or statistical significance of the pooled effect estimate.

Table 4. (*Continued.*)

Evidence factor	Rating	Rationale
Other compelling attributes of the data that may influence certainty	NA	All but one of the sufficiently powered ($df_{Satt} > 4$) meta-analyses performed in this review were statistically significant at $P_{Satt} \leq 0.05$ (Table 3).
Overall strength of evidence	Sufficient	This systematic review and meta-analysis found sufficient evidence of an association between higher adult OP and NMC insecticide exposure and lower sperm concentration. The review team believes with reasonable confidence that chance, bias, and confounding can be ruled out as an explanation for the association. The available evidence includes results from one or more well-designed, well-conducted studies, and the review team does not believe that the results of future studies would impact the findings of this investigation.

Note: The pooled Hedges' g (G_{Pooled}) represents the bias-adjusted standardized mean difference in sperm concentration between adult men more- and less-exposed to OP and NMC insecticides pooled across studies included in a given meta-analysis. The review team assumed a strong correlation ($r = 0.8$) between dependent effect sizes originating from, or nested under, fully or partially overlapping study participants (the study population "cluster"). A cluster-robust variance estimator (robust to the assumed correlation between dependent effect sizes within the same cluster) with bias-reduced linearization small-sample adjustment based on Satterthwaite approximated degrees of freedom (df_{Satt}) was employed. A Satterthwaite-adjusted p-value (P_{Satt}) of ≤ 0.05 was considered statistically significant. Heterogeneity point estimates (τ^2) were estimated using restricted maximum likelihood estimation for each level of random effect. Statistical significance of heterogeneity was tested using the Q-test statistic (Q_H) based on a chi-square distribution. Corresponding I^2 statistics represent the proportion of total heterogeneity that can be attributed to each level of random effect. Detailed criteria and guiding questions used to assess the quality and strength of evidence are presented in Table S9. Strength of evidence rating definitions for human evidence are presented in Table S11. NA, not applicable; P_{Satt}, Satterthwaite-adjusted p-value.

investigation expands on this meta-analysis by including more studies, exploring both OP and NMC insecticides, and accounting for hierarchical and correlational dependencies in the epidemiological data.

Limitations

Limitations of the epidemiological literature reviewed here include *a*) inconsistent reporting across studies and time periods, *b*) missing data that could have informed the risk of bias assessments or meta-analysis, and *c*) the predominantly cross-sectional nature of the body of evidence, which makes causal inference challenging at this time. Limitations of this systematic review and meta-analysis include *a*) the exploratory nature of the statistical transformations of study-reported effect sizes to a common index for quantitative synthesis (Hedges' g), given that many of the effect size transformation methods in the meta-analytical literature relate to bivariate results only; *b*) the pooling of bivariate and multivariate effect sizes in the same meta-analysis; and *c*) the lack of a registered protocol.

Multiple efforts were taken to address the limitations of this systematic review and meta-analysis. These efforts include the following: *a*) the potential impact of the effect size transformations was assessed via sensitivity analysis, which demonstrated that effect size transformations did not impact the overall finding that higher adult OP and NMC insecticide exposure is associated with lower sperm concentration; *b*) despite concerns about pooling unadjusted and adjusted results in the same meta-analysis, the review team determined that the benefit of including studies that controlled for key confounders outweighed the statistical uncertainty introduced through the synthesis of results from diverse models; and *c*) all adjustments made to the protocol during the review were carefully recorded and made available to the public.

Strengths

This systematic review and meta-analysis has several strengths. These strengths include: *a*) a comprehensive literature search for primary epidemiological studies published on or before 11 August 2022, with no statistical indication of publication bias, *b*) inclusion of a wide range of study-reported results using existing effect size transformation methods, *c*) data sensitivity analyses to assess the robustness of the three-level, multivariate meta-analytic model used to account for hierarchical and correlational dependencies in study results data, *d*) statistical exploration of heterogeneity through moderator sensitivity and subgroup meta-analyses, and *e*) use of robust and reproducible systematic review and meta-analysis methods.

Key Insights for Future Research

Based on the literature gaps identified in this review, future studies on OP and NMC insecticide exposure and sperm concentration should use prospective cohort study designs to capture temporal variability in exposure, particularly at environmentally relevant levels, and outcome. This would help to address causal inference.[125] Given the limited statistical power of the NMC insecticide subgroup meta-analysis and the lack of existing cohort studies, additional studies on NMC insecticide exposures should be prioritized over OP insecticides. Still, the large magnitude of effect observed across OP insecticide studies raises concerns about the potential impact of glyphosate, a widely used systematic OP herbicide known to be a weak inhibitor of cholinesterase.[126] To our knowledge, there are currently no observational human studies on glyphosate and sperm concentration, representing a major gap in the literature.

The review authors recommend that additional studies focus on whether NMCs are in fact safer than OPs in terms of impacts on sperm concentration, as the results of this review suggest. The answer to this question is essential to prevent regrettable substitution as OP insecticides are replaced with NMC alternatives developed, but not proven, to be safer to humans.[117] More research is also needed to explore whether other carbamate pesticide subclasses (thiocarbamates and dithiocarbamates) may cause similar impacts on sperm concentration as NMCs, given their structural and mechanistic similarities.[127]

Notably, exposure to a mixture of OP and NMC insecticides showed the greatest magnitude of association with lower sperm concentration across all meta-analyses performed in this investigation, indicating a potential synergistic effect of combined exposure to both classes of contemporary use insecticides. Additional studies on the potential synergistic effects of OP and NMC insecticides on sperm concentration would also be useful because people can be exposed to both simultaneously.

Future studies should be careful to reduce bias in two domains—recruitment strategy and exposure assessment—given that these two domains had the greatest impact on the findings of this investigation. Interestingly, in terms of exposure assessment, insights from Sánchez-Peña et al.[111] and Yucra et al.[60,114] suggest that biomonitoring exposure assessment methods may be less sensitive in detecting an association between OP insecticide exposure and sperm concentration, despite being judged as less-biased than self-report or proxy exposure assessment methods.

This may be a result of the nonpersistent nature of OP insecticide metabolites.[34,111,114] To further explore this issue, future studies should consider employing self-report, proxy, and biomonitoring exposure assessment methods to allow for within-study comparison of the sensitivity of each respective exposure assessment method. Novel biomarkers of chronic exposure, such as OP hair analysis methods,[128,129] may also provide additional clarity moving forward. Moreover, standardizing exposure categories would enable between-study exposure comparisons and reduce variations in how authors characterize OP and NMC insecticide exposure.

Finally, the review team has several recommendations to facilitate future quantitative syntheses of epidemiological evidence. The review team encourages primary study authors to report *a*) all results from planned analyses regardless of direction of effect or statistical significance to avoid selective outcome reporting bias, *b*) statistically transformed effect estimates on the transformed and raw scale to enable synthesis of effect sizes on the same scale, and *c*) R^2_{YZ} statistics with all regression model results to facilitate the conversion of beta coefficients to partial or semi-partial effect sizes for use in meta-analysis.

Conclusion

This systematic review and meta-analysis concludes that primary epidemiological studies published through 11 August 2022 demonstrate sufficient evidence that higher adult male exposure to OP and NMC insecticides is associated with lower sperm concentration. Mechanistic and animal evidence indicate that this association may be causal, but the mainly cross-sectional nature of the epidemiological literature makes causal inference challenging at this time. Although additional cohort studies would be beneficial to fill data gaps and address causal inference, action should be taken now to reduce exposure to OP and NMC insecticides and prevent continued reproductive harm.

Acknowledgments

We thank D. Magalhães, S. Arrington, and E. Gomez (three former Perry lab members affiliated with the George Washington University) for their contributions to the risk of bias assessments.

This research did not receive funding from agencies in the public, commercial, or not-for-profit sectors. M.J.P. provided expert consultation in legal cases pertinent to occupational COVID-19 and pesticide health effects in 2021. The other authors of this paper declare they have no actual or potential competing financial interest and that their freedom to design, conduct, interpret, and publish research is not compromised by any controlling sponsor as a condition of review or publication.

The protocol, study results data, and R code used for meta-analysis are available on GitHub (https://github.com/Lauren-Ellis/Ellis-et-al.-2023-OP-NMC-Insecticides-and-Sperm-Concentration). An interactive version of Figure 3 is available on Tableau Public (https://public.tableau.com/app/profile/lauren.ellis3833/viz/Ellisetal_2023ForestPlot_16952586826050/ForestPlot).

References

1. Carlsen E, Giwercman A, Keiding N, Skakkebaek NE. 1992. Evidence for decreasing quality of semen during past 50 years. BMJ 305(6854):609–613, PMID: 1393072, https://doi.org/10.1136/bmj.305.6854.609.
2. Swan SH, Elkin EP, Fenster L. 2000. The question of declining sperm density revisited: an analysis of 101 studies published 1934–1996. Environ Health Perspect 108(10):961–966, PMID: 11049816, https://doi.org/10.1289/ehp.00108961.
3. Sengupta P, Dutta S, Krajewska-Kulak E. 2017. The disappearing sperms: analysis of reports published between 1980 and 2015. Am J Mens Health 11(4):1279–1304, PMID: 27099345, https://doi.org/10.1177/1557988316643383.
4. Huang C, Li B, Xu K, Liu D, Hu J, Yang Y, et al. 2017. Decline in semen quality among 30,636 young Chinese men from 2001 to 2015. Fertil Steril 107(1):83–88.e2, PMID: 27793371, https://doi.org/10.1016/j.fertnstert.2016.09.035.
5. Levine H, Jørgensen N, Martino-Andrade A, Mendiola J, Weksler-Derri D, Mindlis I, et al. 2017. Temporal trends in sperm count: a systematic review and meta-regression analysis. Hum Reprod Update 23(6):646–659, PMID: 28981654, https://doi.org/10.1093/humupd/dmx022.
6. Levine H, Jørgensen N, Martino-Andrade A, Mendiola J, Weksler-Derri D, Jolles M, et al. 2023. 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 29(2):157–176, PMID: 36377604, https://doi.org/10.1093/humupd/dmac035.
7. Bonde JP, Ernst E, Jensen TK, Hjollund NH, Kolstad H, Henriksen TB, et al. 1998. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. Lancet 352(9135):1172–1177, PMID: 9777833, https://doi.org/10.1016/S0140-6736(97)10514-1.
8. Sharpe RM. 2012. Sperm counts and fertility in men: a rocky road ahead. EMBO Rep 13(5):398–403, PMID: 22491033, https://doi.org/10.1038/embor.2012.50.
9. Murray KS, James A, McGeady JB, Reed ML, Kuang WW, Nangia AK. 2012. The effect of the new 2010 World Health Organization criteria for semen analyses on male infertility. Fertil Steril 98(6):1428–1431, PMID: 22921910, https://doi.org/10.1016/j.fertnstert.2012.07.1130.
10. Kumar N, Singh AK. 2015. Trends of male factor infertility, an important cause of infertility: a review of literature. J Hum Reprod Sci 8(4):191–196, PMID: 26752853, https://doi.org/10.4103/0974-1208.170370.
11. WHO (World Health Organization). 2021. *WHO Laboratory Manual for the Examination and Processing of Human Semen*. 6th ed. https://iris.who.int/bitstream/handle/10665/343208/9789240030787-eng.pdf?sequence=1 [accessed 25 July 2022].
12. Campbell MJ, Lotti F, Baldi E, Schlatt S, Festin MPR, Björndahl L, et al. 2021. Distribution of semen examination results 2020—a follow up of data collated for the WHO semen analysis manual 2010. Andrology 9(3):817–822, PMID: 33528873, https://doi.org/10.1111/andr.12983.
13. Ferlin A, Garolla A, Ghezzi M, Selice R, Palego P, Caretta N, et al. 2021. Sperm count and hypogonadism as markers of general male health. Eur Urol Focus 7(1):205–213, PMID: 31427194, https://doi.org/10.1016/j.euf.2019.08.001.
14. Chen T, Belladelli F, Del Giudice F, Eisenberg ML. 2022. Male fertility as a marker for health. Reprod Biomed Online 44(1):131–144, PMID: 34848151, https://doi.org/10.1016/j.rbmo.2021.09.023.
15. Jacobsen R, Bostofte E, Engholm G, Hansen J, Olsen JH, Skakkebæk NE, et al. 2000. Risk of testicular cancer in men with abnormal semen characteristics: cohort study. BMJ 321(7264):789–792, PMID: 11009515, https://doi.org/10.1136/bmj.321.7264.789.
16. Hanson HA, Anderson RE, Aston KI, Carrell DT, Smith KR, Hotaling JM. 2016. Subfertility increases risk of testicular cancer: evidence from population-based semen samples. Fertil Steril 105(2):322–328.e1, PMID: 26604070, https://doi.org/10.1016/j.fertnstert.2015.10.027.
17. Ventimiglia E, Capogrosso P, Boeri L, Serino A, Colicchia M, Ippolito S, et al. 2015. Infertility as a proxy of general male health: results of a cross-sectional survey. Fertil Steril 104(1):48–55, PMID: 26006735, https://doi.org/10.1016/j.fertnstert.2015.04.020.
18. Eisenberg ML, Li S, Behr B, Pera RR, Cullen MR. 2015. Relationship between semen production and medical comorbidity. Fertil Steril 103(1):66–71, PMID: 25497466, https://doi.org/10.1016/j.fertnstert.2014.10.017.
19. Jensen TK, Jacobsen R, Christensen K, Nielsen NC, Bostofte E. 2009. Good semen quality and life expectancy: a cohort study of 43,277 men. Am J Epidemiol 170(5):559–565, PMID: 19635736, https://doi.org/10.1093/aje/kwp168.
20. Eisenberg ML, Li S, Behr B, Cullen MR, Galusha D, Lamb DJ, et al. 2014. Semen quality, infertility and mortality in the USA. Hum Reprod 29(7):1567–1574, PMID: 24838701, https://doi.org/10.1093/humrep/deu106.
21. Harris ID, Fronczak C, Roth L, Meacham RB. 2011. Fertility and the aging male. Rev Urol 13(4):e184–e190, PMID: 22232567.
22. Benatta M, Kettache R, Buchholz N, Trinchieri A. 2020. The impact of nutrition and lifestyle on male fertility. Arch Ital Urol Androl 92(2):121–131, PMID: 32597116, https://doi.org/10.4081/aiua.2020.2.121.
23. Rehman S, Usman Z, Rehman S, AlDraihem M, Rehman N, Rehman I, et al. 2018. Endocrine disrupting chemicals and impact on male reproductive health. Transl Androl Urol 7(3):490–503, PMID: 30050807, https://doi.org/10.21037/tau.2018.05.17.
24. Frazier LM. 2007. Reproductive disorders associated with pesticide exposure. J Agromedicine 12(1):27–37, PMID: 18032334, https://doi.org/10.1300/J096v12n01_04.
25. Mehrpour O, Karrari P, Zamani N, Tsatsakis AM, Abdollahi M. 2014. Occupational exposure to pesticides and consequences on male semen and fertility: a review. Toxicol Lett 230(2):146–156, PMID: 24487096, https://doi.org/10.1016/j.toxlet.2014.01.029.

26. Sharma A, Kumar V, Shahzad B, Tanveer M, Sidhu GPS, Handa N, et al. 2019. Worldwide pesticide usage and its impacts on ecosystem. SN Appl Sci 1:1446, https://doi.org/10.1007/s42452-019-1485-1.
27. Mdeni NL, Adeniji AO, Okoh AI, Okoh OO. 2022. Analytical evaluation of carbamate and organophosphate pesticides in human and environmental matrices: a review. Molecules 27(3):618, PMID: 35163876, https://doi.org/10.3390/molecules27030618.
28. Whorton D, Krauss R, Marshall S, Milby T. 1977. Infertility in male pesticide workers. Lancet 2(8051):1259–1261, PMID: 73955, https://doi.org/10.1016/s0140-6736(77)92665-4.
29. WHO Task Group. 1986. *Organophosphorus Insecticides: A General Introduction*. Environmental Health Criteria 63. Geneva, Switzerland: World Health Organization.
30. Gupta RC, Doss RB, Yurdakok-Dikmen B, Malik JK, Zaja-Milatovic S, Milatovic D. 2022. Chapter 33 - Organophosphates and carbamates. In: *Reproductive and Developmental Toxicology. Third*. Gupta RC, ed. Cambridge, MA: Academic Press, 617–639.
31. U.S. EPA (U.S. Environmental Protection Agency). 2000. *The Use of Data on Cholinesterase Inhibition for Risk Assessments of Organophosphorous and Carbamate Pesticides*. https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/use-data-cholinesterase-inhibition-risk-assessments [accessed 4 September 2022].
32. Wilson BW, Arrieta DE, Henderson JD. 2005. Monitoring cholinesterases to detect pesticide exposure. Chem Biol Interact 157–158:253–256, PMID: 16298353, https://doi.org/10.1016/j.cbi.2005.10.043.
33. Perry MJ. 2008. Effects of environmental and occupational pesticide exposure on human sperm: a systematic review. Hum Reprod Update 14(3):233–242, PMID: 18281240, https://doi.org/10.1093/humupd/dmm039.
34. Koureas M, Tsakalof A, Tsatsakis A, Hadjichristodoulou C. 2012. Systematic review of biomonitoring studies to determine the association between exposure to organophosphorus and pyrethroid insecticides and human health outcomes. Toxicol Lett 210(2):155–168, PMID: 22020228, https://doi.org/10.1016/j.toxlet.2011.10.007.
35. Martenies SE, Perry MJ. 2013. Environmental and occupational pesticide exposure and human sperm parameters: a systematic review. Toxicology 307:66–73, PMID: 23438386, https://doi.org/10.1016/j.tox.2013.02.005.
36. Knapke ET, Magalhaes DP, Dalvie MA, Mandrioli D, Perry MJ. 2022. Environmental and occupational pesticide exposure and human sperm parameters: a Navigation Guide review. Toxicology 465:153017, PMID: 34756984, https://doi.org/10.1016/j.tox.2021.153017.
37. Woodruff TJ, Sutton P, Navigation Guide Work Group. 2011. An evidence-based medicine methodology to bridge the gap between clinical and environmental health sciences. Health Aff (Millwood) 30(5):931–937, PMID: 21555477, https://doi.org/10.1377/hlthaff.2010.1219.
38. Woodruff TJ, Sutton P. 2014. The Navigation Guide systematic review methodology: a rigorous and transparent method for translating environmental health science into better health outcomes. Environ Health Perspect 122(10):1007–1014, PMID: 24968373, https://doi.org/10.1289/ehp.1307175.
39. Lam J, Lanphear BP, Bellinger D, Axelrad DA, McPartland J, Sutton P, et al. 2017. Developmental PBDE exposure and IQ/ADHD in childhood: a systematic review and meta-analysis. Environ Health Perspect 125(8):086001, PMID: 28799918, https://doi.org/10.1289/EHP1632.
40. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. 2021. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 372:n160, PMID: 33781993, https://doi.org/10.1136/bmj.n160.
41. Lam J, Sutton P, McPartland J, Davidson L, Daniels N, Sen S, et al. 2019. Applying the navigation guide systematic review methodology. Case study #5: association between developmental exposures to PBDEs and human neurodevelopment. PROSPERO 2015 CRD42015019753. https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42015019753 [accessed 10 September 2022].
42. Morgan RL, Whaley P, Thayer KA, Schünemann HJ. 2018. Identifying the PECO: a framework for formulating good questions to explore the association of environmental and other exposures with health outcomes. Environ Int 121(pt 1):1027–1031, PMID: 30166065, https://doi.org/10.1016/j.envint.2018.07.015.
43. Johnson PI, Sutton P, Atchley DS, Koustas E, Lam J, Sen S, et al. 2014. The Navigation Guide—evidence-based medicine meets environmental health: systematic review of human evidence for PFOA effects on fetal growth. Environ Health Perspect 122(10):1028–1039, PMID: 24968388, https://doi.org/10.1289/ehp.1307893.
44. Higgins J, Savović J, Page M, Elbers R, Sterne JAC. 2023. Chapter 8: Assessing risk of bias in a randomized trial. In: *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.4. Sterne J, Higgins J, Thomas J, eds. Chichester, UK: John Wiley & Sons.
45. Viswanathan M, Ansari MT, Berkman ND, Chang S, Hartling L, McPheeters M, et al. 2008. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. In: *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville, MD: Agency for Healthcare Research and Quality (US).
46. Bero LA. 2013. Why the Cochrane risk of bias tool should include funding source as a standard item. Cochrane Database Syst Rev (12):ED000075, PMID: 24575439, https://doi.org/10.1002/14651858.ED000075.
47. Borenstein M, Hedges LV. 2019. Effect sizes for meta-analysis. In: *The Handbook of Research Synthesis and Meta-Analysis*. Cooper H, Hedges LV, Valentine JC, eds. 3rd ed. New York, NY: Russell Sage Foundation, 207–244.
48. Deeks J, Higgins JPT, Altman DG. 2023. Chapter 10: Analysing data and undertaking meta-analyses. In: *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.4. Sterne J, Higgins J, Thomas J, eds. Chichester, UK: John Wiley & Sons.
49. Bakker A, Cai J, English L, Kaiser G, Mesa V, Van Dooren W. 2019. Beyond small, medium, or large: points of consideration when interpreting effect sizes. Educ Stud Math 102(1):1–8, https://doi.org/10.1007/s10649-019-09908-4.
50. Cohen J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: L. Erlbaum Associates.
51. Aloe AM, Thompson CG, Liu Z, Lin L. 2022. Estimating partial standardized mean differences from regression models. J Exp Educ 90(4):898–915, https://doi.org/10.1080/00220973.2021.1966605.
52. Higgins JPT, Li T, Deeks JJ. 2023. Chapter 6: Choosing effect measures and computing estimates of effect. In: *Cochrane Handbook for Systematic Reviews of Interventions*. Version 6.4. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., eds. Cochrane. https://training.cochrane.org/handbook/current/chapter-06 [accessed 18 September 2023].
53. Dziewirska E, Radwan M, Wielgomas B, Klimowska A, Radwan P, Kałużny P, et al. 2019. Human semen quality, sperm DNA damage, and the level of urinary concentrations of 1N and TCPY, the biomarkers of nonpersistent insecticides. Am J Mens Health 13(1):1557988318816598, PMID: 30813854, https://doi.org/10.1177/1557988318816598.
54. Lin X, Ryan L, Sammel M, Zhang D, Padungtod C, Xu X. 2000. A scaled linear mixed model for multiple outcomes. Biometrics 56(2):593–601, PMID: 10877322, https://doi.org/10.1111/j.0006-341X.2000.00593.x.
55. Mármol-Maneiro L, Fernández-D'Pool J, Sánchez BJ, Sirit Y. 2003. Seminal profile in workers exposed to cholinesterase inhibitor insecticides [in Spanish.] Invest Clin 44(2):105–117, PMID: 12815841.
56. Multigner L, Kadhel P, Pascal M, Huc-Terki F, Kercret H, Massart C, et al. 2008. Parallel assessment of male reproductive function in workers and wild rats exposed to pesticides in banana plantations in Guadeloupe. Environ Health 7:40, PMID: 18667078, https://doi.org/10.1186/1476-069X-7-40.
57. Padungtod C, Hassold TJ, Millie E, Ryan LM, Savitz DA, Christiani DC, et al. 1999. Sperm aneuploidy among Chinese pesticide factory workers: scoring by the FISH method. Am J Ind Med 36(2):230–238, PMID: 10398931, https://doi.org/10.1002/(SICI)1097-0274(199908)36:2<230::AID-AJIM2>3.0.CO;2-6.
58. Padungtod C, Savitz DA, Overstreet JW, Christiani DC, Ryan LM, Xu X. 2000. Occupational pesticide exposure and semen quality among Chinese workers. J Occup Environ Med 42(10):982–992, PMID: 11039162, https://doi.org/10.1097/00043764-200010000-00004.
59. Wyrobek AJ, Watchmaker G, Gordon L, Wong K, Moore D II, Whorton D. 1981. Sperm shape abnormalities in carbaryl-exposed employees. Environ Health Perspect 40:255–265, PMID: 6791917, https://doi.org/10.1289/ehp.8140255.
60. Yucra S, Rubio J, Gasco M, Gonzales C, Steenland K, Gonzales GF. 2006. Semen quality and reproductive sex hormone levels in Peruvian pesticide sprayers. Int J Occup Environ Health 12(4):355–361, PMID: 17168223, https://doi.org/10.1179/oeh.2006.12.4.355.
61. Higgins JPT, White IR, Anzures-Cabrera J. 2008. Meta-analysis of skewed data: combining results reported on log-transformed or raw scales. Stat Med 27(29):6072–6092, PMID: 18800342, https://doi.org/10.1002/sim.3427.
62. Perry MJ, Venners SA, Barr DB, Xu X. 2007. Environmental pyrethroid and organophosphorus insecticide exposures and sperm concentration. Reprod Toxicol 23(1):113–118, PMID: 17011162, https://doi.org/10.1016/j.reprotox.2006.08.005.
63. Tan LF, Sun XZ, Li YN, Ji JM, Wang QL, Chen LS, et al. 2005. Effects of carbaryl production exposure on the sperm and semen quality of occupational male workers [in Chinese]. Zhonghua Lao Dong Wei Sheng Zhi Ye Bing Za Zhi 23(2):87–90, PMID: 16105445.
64. Aloe AM, Thompson CG. 2013. The synthesis of partial effect sizes. J Soc Work Res 4(4):390–405, https://doi.org/10.5243/jsswr.2013.24.
65. Rodríguez-Barranco M, Tobías A, Redondo D, Molina-Portillo E, Sánchez MJ. 2017. Standardizing effect size from linear regression models with log-transformed variables for meta-analysis. BMC Med Res Methodol 17(1):44, PMID: 28302052, https://doi.org/10.1186/s12874-017-0322-8.

66. Swan SH, Kruse RL, Liu F, Barr DB, Drobnis EZ, Redmon JB, et al. 2003. Semen quality in relation to biomarkers of pesticide exposure. Environ Health Perspect 111(12):1478–1484, PMID: 12948887, https://doi.org/10.1289/ehp.6417.
67. Melgarejo M, Mendiola J, Koch HM, Moñino-García M, Noguera-Velasco JA, Torres-Cantero AM. 2015. Associations between urinary organophosphate pesticide metabolite levels and reproductive parameters in men from an infertility clinic. Environ Res 137:292–298, PMID: 25601731, https://doi.org/10.1016/j.envres.2015.01.004.
68. Viechtbauer W. 2022. metafor: meta-analysis package for R. https://CRAN.R-project.org/package=metafor [accessed 10 September 2022].
69. Pustejovsky J. 2022. clubSandwich: cluster-robust (sandwich) variance estimators with small-sample corrections. https://CRAN.R-project.org/package=clubSandwich [accessed 10 September 2022].
70. Harrer M, Cuijpers P, Furukawa T, Ebert DD. 2019. dmetar: companion R Package for the guide "Doing Meta-Analysis in R." http://dmetar.protectlab.org/ [accessed 4 September 2022].
71. Konstantopoulos S, Hedges LV. 2019. Statistically analyzing effect sizes: fixed- and random-effects models. In: *The Handbook of Research Synthesis and Meta-Analysis*. Cooper H, Hedges LV, Valentine JC, eds. 3rd ed. New York, NY: Russell Sage Foundation, 245–280.
72. López-López JA, Page MJ, Lipsey MW, Higgins JPT. 2018. Dealing with effect size multiplicity in systematic reviews and meta-analyses. Res Synth Methods 9(3):336–351, PMID: 29971966, https://doi.org/10.1002/jrsm.1310.
73. Hedges LV. 2019. Stochastically dependent effect sizes. In: *The Handbook of Research Synthesis and Meta-Analysis*. Cooper H, Hedges LV, Valentine JC, eds. 3rd ed. New York, NY: Russell Sage Foundation, 281–298.
74. Pustejovsky JE, Tipton E. 2022. Meta-analysis with robust variance estimation: expanding the range of working models. Prev Sci 23(3):425–438, PMID: 33961175, https://doi.org/10.1007/s11121-021-01246-3.
75. Viechtbauer W. 2010. Conducting meta-analyses in R with the metafor package. J Stat Softw 36(3):1–48, https://doi.org/10.18637/jss.v036.i03.
76. Satterthwaite FE. 1946. An approximate distribution of estimates of variance components. Biometrics 2(6):110–114, PMID: 20287815, https://doi.org/10.2307/3002019.
77. McCaffrey DF, Bell RM, Botts CH. 2001. Generalizations of biased reduced linearization. In: *Proceedings of the Annual Meeting of the American Statistical Association*. 5–9 August 2001. Atlanta, GA: American Statistical Association. http://www.asasrms.org/Proceedings/y2001/Proceed/00264.pdf [accessed 29 October 2023].
78. Bell RM, McCaffrey DF. 2002. Bias reduction in standard errors for linear regression with multi-stage samples. Surv Methodol 28(2):169–181.
79. Tipton E. 2015. Small sample adjustments for robust variance estimation with meta-regression. Psychol Methods 20(3):375–393, PMID: 24773356, https://doi.org/10.1037/met0000011.
80. Langan D, Higgins JPT, Jackson D, Bowden J, Veroniki AA, Kontopantelis E, et al. 2019. A comparison of heterogeneity variance estimators in simulated random-effects meta-analyses. Res Synth Methods 10(1):83–98, PMID: 30067315, https://doi.org/10.1002/jrsm.1316.
81. Harrer M, Cuijpers P, Furukawa TA, Ebert DD. Chapter 10. "Multilevel" meta-analysis. In: *Doing Meta-Analysis in R: A Hands-on Guide*. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/multilevel-ma.html [accessed 30 April 2022].
82. DerSimonian R, Laird N. 1986. Meta-analysis in clinical trials. Control Clin Trials 7(3):177–188, PMID: 3802833, https://doi.org/10.1016/0197-2456(86)90046-2.
83. Viechtbauer W. 2007. Hypothesis tests for population heterogeneity in meta-analysis. Br J Math Stat Psychol 60(pt 1):29–60, PMID: 17535578, https://doi.org/10.1348/000711005X64042.
84. Cheung MWL. 2014. Modeling dependent effect sizes with three-level meta-analyses: a structural equation modeling approach. Psychol Methods 19(2):211–229, PMID: 23834422, https://doi.org/10.1037/a0032968.
85. Viechtbauer W. 2007. Confidence intervals for the amount of heterogeneity in meta-analysis. Stat Med 26(1):37–52, PMID: 16463355, https://doi.org/10.1002/sim.2514.
86. Tukey JW. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.
87. Cook RD. 1977. Detection of influential observation in linear regression. Technometrics 19(1):15–18, https://doi.org/10.2307/1268249.
88. Riley RD, Higgins JPT, Deeks JJ. 2011. Interpretation of random effects meta-analyses. BMJ 342:d549, PMID: 21310794, https://doi.org/10.1136/bmj.d549.
89. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. 2016. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 6(7):e010247, PMID: 27406637, https://doi.org/10.1136/bmjopen-2015-010247.
90. Rubio-Aparicio M, López-López JA, Viechtbauer W, Marín-Martínez F, Botella J, Sánchez-Meca J. 2020. Testing categorical moderators in mixed-effects meta-analysis in the presence of heteroscedasticity. J Exp Educ 88(2):288–310, https://doi.org/10.1080/00220973.2018.1561404.
91. Hartung J, Makambi KH, Argaç D. 2001. An extended ANOVA F-test with applications to the heterogeneity problem in meta-analysis. Biom J 43(2):135–146, https://doi.org/10.1002/1521-4036(200105)43:2<135::AID-BIMJ135>3.0.CO;2-H.
92. Egger M, Davey Smith G, Schneider M, Minder C. 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315(7109):629–634, PMID: 9310563, https://doi.org/10.1136/bmj.315.7109.629.
93. Rodgers MA, Pustejovsky JE. 2020. Evaluating meta-analytic methods to detect selective reporting in the presence of dependent effect sizes. Psychol Methods 26(2):141–160, PMID: 32673040, https://doi.org/10.1037/met0000300.
94. Zwetsloot PP, Van Der Naald M, Sena ES, Howells DW, IntHout J, De Groot JA, et al. 2017. Standardized mean differences cause funnel plot distortion in publication bias assessments. eLife 6:e24260, PMID: 28884685, https://doi.org/10.7554/eLife.24260.
95. Shi L, Lin L. 2019. The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses. Medicine (Baltimore) 98(23):e15987, PMID: 31169736, https://doi.org/10.1097/MD.0000000000015987.
96. IARC (International Agency for Research on Cancer). 2019. *IARC Monographs on the Evaluation of Carcinogenic Risks to Humans: Preamble*. https://monographs.iarc.who.int/wp-content/uploads/2019/07/Preamble-2019.pdf [accessed 15 September 2022].
97. Sawaya GF, Guirguis-Blake J, LeFevre M, Harris R, Petitti D, U.S. Preventive Services Task Force. 2007. Update on the methods of the U.S. Preventive Services Task Force: estimating certainty and magnitude of net benefit. Ann Intern Med 147(12):871–875, PMID: 18087058, https://doi.org/10.7326/0003-4819-147-12-200712180-00007.
98. U.S. EPA. 1991. *Guidelines for Developmental Toxicity Risk Assessment*. EPA/600/FR-91/001. https://www.epa.gov/sites/default/files/2014-11/documents/dev_tox.pdf [accessed 15 September 2022].
99. U.S. EPA. 1996. *Guidelines for Developmental Toxicity Risk Assessment*. EPA/630/R-96/009. https://www.epa.gov/sites/default/files/2014-11/documents/guidelines_repro_toxicity.pdf [accessed 15 September 2022].
100. Cremonese C, Piccoli C, Pasqualotto F, Clapauch R, Koifman RJ, Koifman S, et al. 2017. Occupational exposure to pesticides, reproductive hormone levels and sperm quality in young Brazilian men. Reprod Toxicol 67:174–185, PMID: 28077271, https://doi.org/10.1016/j.reprotox.2017.01.001.
101. Ghafouri-Khosrowshahi A, Ranjbar A, Mousavi L, Nili-Ahmadabadi H, Ghaffari F, Zeinvand-Lorestani H, et al. 2019. Chronic exposure to organophosphate pesticides as an important challenge in promoting reproductive health: a comparative study. J Educ Health Promot 8(1):149, PMID: 31544114, https://doi.org/10.4103/jehp.jehp_148_19.
102. Hossain F, Ali O, D'Souza UJA, Naing DKS. 2010. Effects of pesticide use on semen quality among farmers in rural areas of Sabah, Malaysia. J Occup Health 52(6):353–360, PMID: 20924153, https://doi.org/10.1539/joh.L10006.
103. Juhler RK, Larsen SB, Meyer O, Jensen ND, Spanò M, Giwercman A, et al. 1999. Human semen quality in relation to dietary pesticide exposure and organic diet. Arch Environ Contam Toxicol 37(3):415–423, PMID: 10473800, https://doi.org/10.1007/s002449900533.
104. Manikandan I, Bora S, Adole PS, Thyagaraju C, Nachiappa Ganesh R. 2021. Assessment of organophosphate pesticides exposure in men with idiopathic abnormal semen analysis: a cross-sectional pilot study. Int J Fertil Steril 15(3):219–225, PMID: 34155869, https://doi.org/10.22074/IJFS.2020.134650.
105. Meeker JD, Ryan L, Barr DB, Herrick RF, Bennett DH, Bravo R, et al. 2004. The relationship of urinary metabolites of carbaryl/naphthalene and chlorpyrifos with human semen quality. Environ Health Perspect 112(17):1665–1670, PMID: 15579410, https://doi.org/10.1289/ehp.7234.
106. Miranda-Contreras L, Gómez-Pérez R, Rojas G, Cruz I, Berrueta L, Salmen S, et al. 2013. Occupational exposure to organophosphate and carbamate pesticides affects sperm chromatin integrity and reproductive hormone levels among Venezuelan farm workers. J Occup Health 55(3):195–203, PMID: 23445617, https://doi.org/10.1539/joh.12-0144-fs.
107. Padungtod C, Niu T, Wang Z, Savitz DA, Christiani DC, Ryan LM, et al. 1999. Paraoxonase polymorphism and its effect on male reproductive outcomes among Chinese pesticide factory workers. Am J Ind Med 36(3):379–387, PMID: 10470002, https://doi.org/10.1002/(SICI)1097-0274(199909)36:3<379::AID-AJIM5>3.0.CO;2-8.
108. Pérez-Herrera N, Polanco-Minaya H, Salazar-Arredondo E, Solís-Heredia MJ, Hernández-Ochoa I, Rojas-García E, et al. 2008. PON1Q192R genetic polymorphism modifies organophosphorous pesticide effects on semen quality and DNA integrity in agricultural workers from southern Mexico. Toxicol Appl Pharmacol 230(2):261–268, PMID: 18430447, https://doi.org/10.1016/j.taap.2008.02.021.
109. Perry MJ, Venners SA, Chen X, Liu X, Tang G, Xing H, et al. 2011. Organophosphorous pesticide exposures and sperm quality. Reprod Toxicol 31(1):75–79, PMID: 20850521, https://doi.org/10.1016/j.reprotox.2010.08.006.

110. Recio-Vega R, Ocampo-Gómez G, Borja-Aburto VH, Moran-Martínez J, Cebrian-Garcia ME. 2008. Organophosphorus pesticide exposure decreases sperm quality: association between sperm parameters and urinary pesticide levels. J Appl Toxicol 28(5):674–680, PMID: 18046699, https://doi.org/10.1002/jat.1321.
111. Sánchez-Peña LC, Reyes BE, López-Carrillo L, Recio R, Morán-Martínez J, Cebrián ME, et al. 2004. Organophosphorous pesticide exposure alters sperm chromatin structure in Mexican agricultural workers. Toxicol Appl Pharmacol 196(1):108–113, PMID: 15050412, https://doi.org/10.1016/j.taap.2003.11.023.
112. Whorton MD, Milby TH, Stubbs HA, Avashia BH, Hull EQ. 1979. Testicular function among carbaryl-exposed employees. J Toxicol Environ Health 5(5):929–941, PMID: 117116, https://doi.org/10.1080/15287397909529802.
113. Xia Y, Cheng S, Bian Q, Xu L, Collins MD, Chang HC, et al. 2005. Genotoxic effects on spermatozoa of carbaryl-exposed workers. Toxicol Sci 85(1):615–623, PMID: 15615886, https://doi.org/10.1093/toxsci/kfi066.
114. Yucra S, Gasco M, Rubio J, Gonzales GF. 2008. Semen quality in Peruvian pesticide applicators: association between urinary organophosphate metabolites and semen parameters. Environ Health 7:59, PMID: 19014632, https://doi.org/10.1186/1476-069X-7-59.
115. National Research Council. 2012. Appendix B. Carbamate pesticide and methyl isocyanate timeline. In: *The Use and Storage of Methyl Isocyanate (MIC) at Bayer CropScience*. Washington, DC: National Academies Press.
116. Nicolle-Mir L. 2011. Organophosphate compounds, sperm quality and hormonal levels. Environ Risques Sante 10(2):86–87.
117. Vale JA, Bradberry SM. 2017. Organophosphate and Carbamate Insecticide. In: *Critical Care Toxicology: Diagnosis and Management of the Critically Poisoned Patient*. Brent J, Burkhart K, Dargan P, Hatten B, Mergarbane B, Palmer R, eds. Cham, Switzerland: Springer International Publishing, 1829–1853.
118. Alavanja MCR. 2009. Introduction: pesticides use and exposure, extensive worldwide. Rev Environ Health 24(4):303–309, PMID: 20384038, https://doi.org/10.1515/reveh.2009.24.4.303.
119. Alba AC, Alexander PE, Chang J, MacIsaac J, DeFry S, Guyatt GH. 2016. High statistical heterogeneity is more frequent in meta-analysis of continuous than binary outcomes. J Clin Epidemiol 70:129–135, PMID: 26386323, https://doi.org/10.1016/j.jclinepi.2015.09.005.
120. Kitamura S, Sugihara K, Fujimoto N. 2006. Chapter 34. Endocrine disruption by organophosphate and carbamate pesticides. In: *Toxicology of Organophosphate & Carbamate Compounds*. Gupta RC, ed. Cambridge, MA: Academic Press, 481–494.
121. Shtenberg AI, Rybakova MN. 1968. Effect of carbaryl on the neuroendocrine system of rats. Food Cosmet Toxicol 6(4):461–467, PMID: 5753604, https://doi.org/10.1016/0015-6264(68)90136-3.
122. Krsmanovic LZ, Mores N, Navarro CE, Saeed SA, Arora KK, Catt KJ. 1998. Muscarinic regulation of intracellular signaling and neurosecretion in gonadotropin-releasing hormone neurons. Endocrinology 139(10):4037–4043, PMID: 9751480, https://doi.org/10.1210/endo.139.10.6267.
123. Sarkar R, Mohanakumar KP, Chowdhury M. 2000. Effects of an organophosphate pesticide, quinalphos, on the hypothalamo-pituitary-gonadal axis in adult male rats. J Reprod Fertil 118(1):29–38, PMID: 10793623, https://doi.org/10.1530/reprod/118.1.29.
124. Giulioni C, Maurizi V, Scarcella S, Di Biase M, Iacovelli V, Galosi AB, et al. 2021. Do environmental and occupational exposure to pyrethroids and organophosphates affect human semen parameters? Results of a systematic review and meta-analysis. Andrologia 53(11):e14215, PMID: 34410018, https://doi.org/10.1111/and.14215.
125. Pearce N, Vandenbroucke JP, Lawlor DA. 2019. Causal inference in environmental epidemiology: old and new approaches. Epidemiology 30(3):311–316, PMID: 30789434, https://doi.org/10.1097/EDE.0000000000000987.
126. Costas-Ferreira C, Durán R, Faro LRF. 2022. Toxic effects of glyphosate on the nervous system: a systematic review. Int J Mol Sci 23(9):4605, PMID: 35562999, https://doi.org/10.3390/ijms23094605.
127. U.S. EPA. 2007. Revised *N*-methyl Carbamate Cumulative Risk Assessment. EPA-HQ-OPP-2007-0935-0003. https://www.regulations.gov/document/EPA-HQ-OPP-2008-0347-0029 [accessed 2 October 2022].
128. Knipe DW, Jayasumana C, Siribaddana S, Priyadarshana C, Pearson M, Gunnell D, et al. 2016. Feasibility of hair sampling to assess levels of organophosphate metabolites in rural areas of Sri Lanka. Environ Res 147:207–211, PMID: 26894816, https://doi.org/10.1016/j.envres.2016.02.011.
129. Tsatsakis AM, Barbounis MG, Kavalakis M, Kokkinakis M, Terzi I, Tzatzarakis MN. 2010. Determination of dialkyl phosphates in human hair for the biomonitoring of exposure to organophosphate pesticides. J Chromatogr B Analyt Technol Biomed Life Sci 878(17–18):1246–1252, PMID: 20226747, https://doi.org/10.1016/j.jchromb.2010.02.021.