Journal Pre-proofs

Effects of Radiofrequency Electromagnetic Field (RF-EMF) exposure on male fertility: A systematic review of experimental studies on non-human mammals and human sperm *in vitro*

Eugenia Cordelli, Lucia Ardoino, Barbara Benassi, Claudia Consales, Patrizia Eleuteri, Carmela Marino, Maurizio Sciortino, Paola Villani, Martin H. Brinkworth, Guangdi Chen, James P. McNamee, Andrew W. Wood, Lea Belackova, Jos Verbeek, Francesca Pacchierotti

PII: S0160-4120(24)00095-3
DOI: https://doi.org/10.1016/j.envint.2024.108509
Reference: EI 108509

To appear in: *Environment International*

Received Date: 12 July 2023
Revised Date: 2 February 2024
Accepted Date: 16 February 2024

Please cite this article as: E. Cordelli, L. Ardoino, B. Benassi, C. Consales, P. Eleuteri, C. Marino, M. Sciortino, P. Villani, M. H. Brinkworth, G. Chen, J. P. McNamee, A. W. Wood, L. Belackova, J. Verbeek, F. Pacchierotti, Effects of Radiofrequency Electromagnetic Field (RF-EMF) exposure on male fertility: A systematic review of experimental studies on non-human mammals and human sperm *in vitro*, *Environment International* (2024), doi: https://doi.org/10.1016/j.envint.2024.108509

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024 Published by Elsevier Ltd.

Effects of Radiofrequency Electromagnetic Field (RF-EMF) exposure on male fertility: A systematic review of experimental studies on non-human mammals and human sperm *in vitro*

Eugenia Cordelli[a], Lucia Ardoino[a], Barbara Benassi[a], Claudia Consales[a], Patrizia Eleuteri[a], Carmela Marino[a1], Maurizio Sciortino[b1], Paola Villani[a], Martin H. Brinkworth[c], Guangdi Chen[d], James P. McNamee[e], Andrew W. Wood[f1], Lea Belackova[g], Jos Verbeek[g], Francesca Pacchierotti[a1]

[a]Division Health Protection Technologies, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Rome, Italy;

[b]Department for Sustainability, ENEA, Rome, Italy;

[c]School of Chemistry and Bioscience, Faculty of Life Sciences, University of Bradford, Bradford, UK;

[d]Bioelectromagnetics Laboratory, Zhejiang University School of Medicine, Hangzhou, China;

[e]Non-Ionizing Radiation Health Sciences Division, Consumer and Clinical Radiation Protection Bureau, Health Canada, Ottawa, Canada;

[f]Department of Health Sciences and Biostatistics, Swinburne University of Technology, Hawthorn, Australia;

[g]University Medical Centers Amsterdam, Coronel Institute of Occupational Health, Cochrane Work, Amsterdam, The Netherlands.

[1]retired

Corresponding authors: Eugenia Cordelli, Division Health Protection Technologies, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Rome, Italy; e-mail address: eugenia.cordelli@enea.it. Francesca Pacchierotti, e-mail address: francesca.pacchierotti@gmail.com

Journal Pre-proofs

Abstract

Background: The World Health Organization is coordinating an international project aimed at systematically reviewing the evidence regarding the association between radiofrequency electromagnetic field (RF-EMF) exposure and adverse health effects. Reproductive health outcomes have been identified among the priority topics to be addressed.

Objectives: To evaluate the effect of RF-EMF exposure on male fertility of experimental mammals and on human sperm exposed *in vitro*.

Methods: Three electronic databases (PubMed, Scopus and EMF Portal) were last searched on September 17, 2022. Two independent reviewers screened the studies, which were considered eligible if met the following criteria: 1) Peer-reviewed publications of sham controlled experimental studies, 2) Non-human male mammals exposed at any stage of development or human sperm exposed *in vitro*, 3) RF-EMF exposure within the frequency range of 100 kHz-300 GHz, including electromagnetic pulses (EMP), 4) one of the following indicators of reproductive system impairment:

- *decrease of fertility:* rate of infertile males, rate of nonpregnant females, litter size and *in vitro* fertilization rate;
- *effects on semen quality*: in animal studies sperm count, in both animal and *in vitro* studies sperm vitality, morphology and DNA/chromatin alterations;
- *reproductive organ toxicity:* testis-epididymis weight, testis or epididymis histology, testis histomorphometry, testicular cell death, estimated testicular cell production;
- *hormonal effects*: testosterone level.

Two reviewers extracted study characteristics and outcome data. We assessed risk of bias (RoB) using the Office of Health Assessment and Translation (OHAT) guidelines. We categorized studies into 3 levels of overall RoB: low, some or high concern. We pooled study results in a random effects meta-analysis comparing average exposure to no-exposure and in a dose-response meta-analysis using all exposure doses. For experimental animal studies, we conducted subgroup analyses for species, Specific Absorption Rate (SAR) and temperature increase. We grouped studies on human sperm exposed *in vitro* by the fertility status of sample donors and SAR. We assessed the certainty of the evidence using the GRADE approach after excluding studies that were rated as “high concern” for RoB.

Results: One-hundred and seventeen papers on animal studies and 10 papers on human sperm exposed *in vitro* were included in this review. Only few studies were rated as “low concern” because most studies were at RoB for exposure and/or outcome assessment.

Subgrouping the experimental animal studies by species, SAR, and temperature increase partly accounted for the heterogeneity of individual studies in about one third of the meta-analyses. In no case was it possible to conduct a subgroup analysis of the few human sperm *in vitro* studies because there were always 1 or more groups including less than 3 studies.

Among all the considered endpoints, the meta-analyses of animal studies provided evidence of adverse effects of RF-EMF exposure in all cases but the rate of infertile males and the size of the sired litters. The assessment of certainty according to the GRADE methodology

assigned a moderate certainty to the reduction of pregnancy rate and to the evidence of no-effect on litter size, a low certainty to the reduction of sperm count, and a very low certainty to all the other meta-analysis results. Studies on human sperm exposed *in vitro* indicated a small detrimental effect of RF-EMF exposure on vitality and no-effect on DNA/chromatin alterations. According to GRADE, a very low certainty was attributed to these results. The few studies that used EMP exposure did not show effects on the outcomes. A low to very low certainty was attributed to these results.

Discussion: Many of the studies examined suffered of severe limitations that led to the attribution of uncertainty to the results of the meta-analyses and did not allow to draw firm conclusions on most of the endpoints. Nevertheless, the associations between RF-EMF exposure and decrease of pregnancy rate and sperm count, to which moderate and low certainty were attributed, are not negligible, also in view of the indications that in Western countries human male fertility potential seems to be progressively declining.

It was beyond the scope of our systematic review to determine the shape of the dose-response relationship or to identify a minimum effective exposure level. The subgroup and the dose-response fitting analyses did not show a consistent relationship between the exposure levels and the observed effects. Notably, most studies evaluated RF-EMF exposure levels that were higher than the levels to which human populations are typically exposed, and the limits set in international guidelines. For these reasons we cannot provide suggestions to confirm or reconsider current human exposure limits.

Considering the outcomes of this systematic review and taking into account the limitations found in several of the studies, we suggest that further investigations with better characterization of exposure and dosimetry including several exposure levels and blinded outcome assessment were conducted.

Protocol registration: Protocols for the systematic reviews of animal studies and of human sperm *in vitro* studies were published in Pacchierotti et al., 2021. The former was also registered in PROSPERO (CRD42021227729 https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=227729) and the latter in Open Science Framework (OSF Registration DOI https://doi.org/10.17605/OSF.IO/7MUS3).

Keywords:

Radiofrequency electromagnetic fields, Semen quality, Reproductive organ toxicity, Testosterone, Animal studies, Meta-analysis

Journal Pre-proofs

1. Introduction

1.1. Rationale

Spermatogenesis is a finely regulated process that is sensitive to chemical and physical stresses (Gupta 2022). Concern about a temporal trend of human sperm quality decline in Western countries has been raised (Levine et al., 2017) and more recently reconsidered with proposals of further prospective investigations to better understand its spatiotemporal distribution and underlying ecological factors (Auger et al., 2022, Boulicault et al., 2022).

The possibility of adverse effects of radiofrequency electromagnetic fields (RF-EMF) on male fertility was raised long ago and has been extensively investigated by human epidemiological studies, studies in laboratory animals under experimentally controlled exposure conditions and studies in which human sperm have been exposed *in vitro* (Yu et al., 2021, AGNIR 2012, ICNIRP 2020, SCENIHR 2015*)*. It is well known that heating may impair spermatogenesis and mammalian testes are physiologically maintained at a temperature lower than the body temperature, but it is unclear if RF-EMF may affect male fertility beyond a hyperthermic effect.

The World Health Organization (WHO) is coordinating an international project aimed at systematically reviewing the evidence about adverse RF-EMF effects on human health. An expert group convened by the WHO identified 6 priority topics on which focusing the reviews (Verbeek et al., 2021). Among them, male fertility was included and has been addressed by 2 projects, one regarding human observational studies (Kenny et al., 2022) and one regarding experimental studies. The results of the latter are reported in this paper.

The protocol for the systematic review was previously published (Pacchierotti et al., 2021). Studies on both experimental mammals *in vivo* and human sperm exposed *in vitro* are included. The body of literature comprises, in addition to direct evidence about animal fertility and semen quality, a variety of endpoints related with fertility impairment. We aimed at a comprehensive review of available data and for this reason we considered all these endpoints and organized them into a systematic review structured with multiple outcomes.

Other reviews assessing the impact of RF-EMF exposure on male fertility have recently been published, but these only partially assessed the available literature data (Kesari et al., 2018, Sciorio et al., 2022, Sterling et al., 2022, Vornoli et al., 2019). The few recently published systematic reviews on this topic suffered from some methodological limitations such as the lack of a Risk of Bias analysis (Jaffar et al., 2019, Kim et al., 2021), they limited analysis to only assess effects on semen parameters or were limited to exposure conditions relevant to mobile phone exposures thereby imposing a SAR cut-off (Yu et al., 2021). International committees on human health protection from electromagnetic fields were unable to draw firm conclusions on the possibility of an adverse effect of RF-EMF on male fertility at exposure levels where humans are typically exposed (ICNIRP 2020, SCENIHR 2015).

1.2. Objective

To overcome limitations in the current assessment of the scientific evidence on the RF-EMF effects on male fertility, we carried out a systematic review of peer-reviewed literature of studies in animals and human sperm *in vitro,* according to the guidelines of the WHO (WHO 2014) and of the National Toxicology Program/Office of Health Assessment and Translation (NTP/OHAT) Handbook (NTP 2015a). In particular, we aimed at answering the following two PECO (Population, Exposure, Comparator, Outcome) questions: 1) What is the effect of

exposure to RF-EMF on male fertility compared to either a sham control or an equal-temperature control in non-human mammals? 2) What is the effect of exposure to RF-EMF on male fertility compared to either a sham control or an equal-temperature control, as inferred from studies with human semen exposed *in vitro*?

2. Methods

The systematic reviews of animal and human sperm *in vitro* studies were conducted according to the protocols published in Pacchierotti et al., 2021. Where there were deviations, these are mentioned in Section 4.5.2 Deviations from the protocol.

2.1 Eligibility criteria

The eligibility criteria applied to select studies to be included in the systematic review were those published in the protocol. Notably, guidelines to evaluate the validity of the analytical methods to decide about study eligibility were provided in Supplementary File 2 of Pacchierotti et al., 2021. Studies in which the exposure level could only be inferred from assumed exposure conditions and not by a measurement or numerical modelling were assessed together with all the other studies because it was difficult to set boundaries in a continuum of exposure dosimetry reporting. This is a slight deviation from the protocol (see Section 4.5.2). Studies were screened in relation to each of the elements of the PECO question as reported in Table 1.

Table 1. Eligibility criteria

Journal Pre-proofs

PECO	*Type of studies*	*Inclusion criteria*	*Exclusion criteria*
Population	Animal studies	- Male subjects of experimental mammalian models exposed at any life stage	- Humans - Non-mammalian experimental models
	In vitro studies	- Human sperm exposed *in vitro*	- Non-human sperm - Cell types other than sperm
Exposure	Animal and *in vitro* studies	- RF-EMF (frequency range 100 kHz - 300 GHz) at any exposure level - Electromagnetic pulses (EMP)	- Static or extremely low-frequency magnetic and/or electric fields - Optical radiation - Ultrasounds - Magnetic Resonance Imaging (MRI) - Mobile phone not in GSM mode, and not controlled by hardware or software, unless supported by measured or calculated metrics, as specified in the protocol - Co-exposure to RF-EMF and other chemical or physical agents - Exposure signals with more than 10% of the total signal energy outside the considered frequency range 100 kHz – 300 GHz - Exposure levels for which a minimum contrast between exposed and comparator groups was not guaranteed, as detailed in the protocol
Comparator	Animal and *in vitro* studies	- Sham-exposed controls - Temperature controls	- Historical controls
Outcomes	Animal studies	- Decrease of fertility: rate of infertile males, nonpregnant females over paired females, litter size, *in vitro* fertilization rate - Effects on semen quality: sperm count, morphology, vitality, DNA/chromatin alterations - Reproductive organ toxicity: testis/epididymis weight, testis histomorphometry, testis or epididymis histology, testicular cell death, testicular sperm production - Hormonal effects: testosterone level	- Qualitative evidence of toxic effects on testis and epididymis - Endpoints not predictive for male fertility impairment - Endpoints measured with invalid methods

Journal Pre-proofs

	In vitro studies	- Effects on semen quality: sperm morphology, vitality, DNA/chromatin alterations	- Qualitative evidence of toxic effects on sperm - Sperm alterations not predictive for male fertility impairment - Sperm alterations measured with invalid method

Journal Pre-proofs

We considered only original, controlled experimental studies published in peer-reviewed journals. We excluded non-experimental studies (e.g., human epidemiologic or other observational studies), and studies of exposure of both males and females of a mating pair (additional decision and change from protocol, see Section 4.5.2). We excluded papers reporting reviews, opinions, proceedings or meeting abstracts. We did not impose any year-of-publication or language restriction.

2.2. Information sources

Three publication databases were searched for eligible studies: NCBI PubMed (https://pubmed.ncbi.nlm.nih.gov/), Scopus (https://www.scopus.com/) and EMF Portal (https://www.emf-portal.org/), a database maintained by the RWTH Aachen University, Germany, specifically focused on EMF studies. The three databases were last consulted on September 17, 2022.

2.3. Search strategy

We interrogated the NCBI PubMed and SCOPUS databases, without any limits on year and language, by search queries composed by English terms identifying the exposure, the outcome and the population. We combined these elements in the queries by the Boolean operators “AND/OR/NOT” as reported in the Supplementary Files 3 and 4 of Pacchierotti et al., 2021. Search terms were identified to retrieve all relevant peer-reviewed publications on studies of RF-EMF effects on male fertility in animal and in human sperm *in vitro* by reviewing PubMed Medical Subject Heading (MeSH) terms associated with relevant papers and testing these and other terms chosen by expert judgment through an iterative trial-and-error process. The removal of non-experimental and human studies was done manually rather than by the use of search filters because studies might have been incorrectly indexed in the databases. We searched the EMF Portal database selecting pre-defined domains for topics, frequency ranges and time span among the options, and combining appropriate key words chosen from those listed in the Glossary (Supplementary Files 3 and 4 of Pacchierotti et al., 2021). The search outputs were then aligned to exclude duplicates and the resulting list was screened for eligibility criteria.

The search strategy was peer-reviewed as part of the publication process of the protocol.

2.4. Selection process

Two reviewers independently evaluated the titles and abstracts of the identified papers to exclude records that were not relevant or did not fulfil at least one inclusion criterion for the PECO elements. In the case of disagreement between the reviewers, or when the abstract did not report enough information, the paper was passed to the full text evaluation phase.

Two reviewers independently evaluated the full texts of the identified papers, and any disagreement between the reviewers was resolved by discussion or through involving a third reviewer. If findings from a study were described in more than one article, these were considered as one study only.

Non-English language papers were either translated by the reviewers or through the use of Google Translate (https://translate.google.com/).

2.5. Data collection process

For all eligible studies, one reviewer extracted the study characteristics and results, and a second reviewer checked all the extracted information against the relevant article for completeness and accuracy as a quality control measure. If disagreement occurred between the reviewers, this was resolved through discussion or by consulting a third reviewer. In no case were the reviewers the authors of the scrutinized papers. When essential data were missing or there were inconsistencies in the reported information, the authors were contacted by e-mail and in case of no-reply, a reminder was sent. In cases reporting power density or other exposure metrics instead of the SAR, the latter was estimated, if possible, on the basis of the available information. Data shown in figures were extracted by using digital rulers. In some cases, quantitative results were re-calculated from other data reported to produce a form best suited to a meta-analysis, for example, converting standard errors into standard deviations or calculating means and variation parameters from raw data.

2.6. Data items (outcomes)

We extracted data considered most representative of male fertility impairment and most relevant for human health, following what was planned in the protocol. Data considered redundant and not extracted were: all testicular histomorphometric parameters other than seminiferous tubule diameter, epididymis weight when sperm number was reported, specific sperm motility and morphology parameters. Where the protocol did not provide sufficient details for which data to extract, additional choices were made before inspecting the results Data considered redundant and not extracted were: testis weight when epididymis weight was reported, percentage of dead sperm when percentage of immotile sperm was reported, sperm apoptosis and oxidative stress when sperm DNA/chromatin alterations were reported, decrease of testicular post-meiotic cell fraction when data on testicular sperm production were reported. Whenever possible, we tried to extract data in an inclusive form most suitable for the overall meta-analysis, as in the case of litter size results, where we extracted only data for the whole mating period and not those for specific mating intervals.

We organized animal experimental data into four outcome categories, each one including multiple endpoints. The definitions of outcomes and endpoints are slightly different from those used in the protocol to make them clearer for the readers of the review without changing their content (see Section 4.5.2 and Table 2). Figure 1 schematically illustrates the outcomes and endpoints included in the systematic review, showing the multiple possible biomarkers of male mammal fertility impairment.

Effects on male fertility. This category included reduction of *in vitro* fertilization and development, increase of the rate of infertile males, increase of nonpregnant females in matings with experimental males, and decrease of litter size.

Effects on semen quality. This category included reduction of sperm count, increase of % morphologically abnormal sperm, decrease of sperm vitality (increase of % dead or immotile sperm), increase of sperm DNA/chromatin alterations (including increase of oxidative stress, DNA/chromatin alterations or apoptosis biomarkers).

Reproductive organ toxicity. This category included testis or epididymis weight reduction, testis histomorphometrical or quantitative histopathological alterations, increase of testicular cell death, reduction of estimated testicular sperm production.

Hormonal effects. This category included reduction of testosterone level in testis or serum.

For studies on human sperm exposed *in vitro*, we extracted data on sperm vitality, % morphologically abnormal sperm, sperm DNA/chromatin alterations.

For the synthesis of results in animal studies, the primary outcomes were effects on male reproductive performance and semen quality, assessed by sperm count, vitality or morphology, because they are direct measurements of fertility and are equivalent to human sperm quality criteria established by WHO, respectively. Similarly, for *in vitro* studies, sperm vitality and morphology were considered as the primary endpoints.

Table 2. Summary of changes in the wording and organization of outcomes and endpoints with respect to the published protocol, with reasons for change

Outcome* **Protocol**	**Outcome** **Review**	**Endpoints*** **Protocol**	**Endpoints** **Review**
Decrease of fertilization rate and embryonic survival	Decrease of fertility *(for sake of clarity the name was changed into decrease of fertility, but the content is the same as in the protocol)*	*In vitro* fertilization rate	*In vitro* fertilization rate *(no change)*
		In vivo mating/fertilization rate	Rate of infertile males *(wording change to make the measured parameter more explicit)*
			Nonpregnant females over paired females *(wording change to make the measured parameter more explicit)*
		Pre- and post-implantation embryonic losses	Litter size *(with hindsight, we noticed that pre- and post- implantation losses were seldom measured but litter size was. We considered these both measurements of the same concept, so this was not a real deviation from the protocol)*
Alterations of WHO sperm quality parameters	Effects on semen quality** *(wording change to more explicitly point to effects on sperm by combining 2 original outcomes into 1)*	Sperm count	Sperm count *(no change)*
		Sperm viability	Sperm vitality** *(we considered these measurements both representing the same concept and we combined them in the same endpoint. We prevented a study be included more than once by only taking one of these measures; should both parameters be measured in one paper, we extracted only motility data)*
		Sperm motility	
		Sperm morphology	Sperm morphology** *(no change)*
Alterations of other sperm		Oxidative stress	Sperm DNA/chromatin alterations ** *(wording change following the assimilation of the 3 original endpoints into 1, because of afterthought consideration of their common involvement in DNA/chromatin alterations. When multiple*
		Apoptosis	

integrity biomarkers		DNA/chromatin alterations	*endpoints were measured in the same samples of a study, we prevented this study be included more than once by only extracting one of these measures according to the hierarchy DNA/chromatin alterations, apoptosis, oxidative stress)*
Reproductive organ toxicity	Reproductive organ toxicity *(no change)*	Weight of testes and epididymis	Testis-epididymis weight *(for sake of simplicity)*
		Quantitative alterations of testis or epididymis histology	Testis or epididymis histology *(for sake of simplicity)*
			Testis histomorphometry *(this endpoint was missing in protocol Table 2, but it had been considered in the extraction guidelines detailed in the protocol Supplementary File 2)*
		Decrease of testicular post-meiotic cell fraction	Testicular sperm production *(some studies measured the fraction of post-meiotic cells and some others the testicular sperm production. We considered these endpoints similar and combined them. Should both parameters be measured in the same animals, we extracted testicular sperm production only)*
		Testicular cell apoptosis	Testicular cell death *(wording change to more comprehensively describe cell death phenomena, irrespective of the mechanism)*
Alterations of reproductive hormones	Hormonal effects *(wording change for sake of simplicity)*	Testosterone level in serum or reproductive organs	Testosterone level *(wording change for sake of simplicity)*

* As reported in Table 2 of Pacchierotti et al., 2021

** The same changes regard both animal studies and human sperm *in vitro* studies

2.7. Data items (other variables)

In addition to outcome data, we also extracted information related to the populations of experimental animals or human sample donors, the exposure conditions and the comparator characteristics.

For animal studies, the species, strain, age and number of animals were recorded. Several variables were extracted to characterize exposure conditions and to assess the risk of bias: frequency, modulation, exposure system, exposure level, exposure duration, animal temperature, lifestage at exposure (prenatally, pre-puberty, post-puberty).

Particular attention was given to the extraction of dosimetric information that defined the exposure level. When data on the exposure level(s) in terms of SAR were not reported, a

SAR estimate was calculated based on other dosimetric information and biophysical assumptions. Where exposure was presented as Power Density values only, SAR estimates were made using appropriate species-specific graphs of W/kg per W/m^2 in the frequency range 10 – 10,000 MHz from the Radiofrequency Radiation Dosimetry Handbook (Durney et al., 1986). These graphs are based on prolate spheroid models, but with experimental confirmation. The conversions are subject to many factors, such as body size and orientation, so it is difficult to assign a specific uncertainty, although a figure of +/- 30% is typically used. Randomization of animals to study groups, allocation concealment and blinding during exposure and/or outcome assessment, sham exposure conditions and statistical methods applied were also considered as elements in the risk of bias assessment.

For *in vitro* studies, population characteristics included donor age and fertility status. Exposure data were essentially the same as those recorded in animal studies. Regarding comparator conditions, we considered whether matched sham and exposed samples were derived from the same donor(s).

We extracted but did not further analyse information on conflict of interest and funding sources, as initially planned in the protocol, since, in the vast majority of papers, public funding and absence of conflict of interest were declared (see Section 4.5.2).

2.8. Study risk of bias assessment

Risk of bias (RoB) was evaluated using the RoB Rating Tool developed by OHAT (NTP, 2015a, b), with minor modifications informed by RoB expertise developed within SYRCLE (Hooijmans et al., 2014). Six bias domains were considered: 1) Selection bias; 2) Performance bias; 3) Detection bias relative to confidence in the exposure and outcome assessment; 4) Attrition/Exclusion bias; 5) Selective reporting bias; 6) Other sources of bias. For each of these domains a set of predefined questions guided the reviewers in the assessment of the internal quality of data.

Questions were based on those proposed in the OHAT handbook; the question “Has possible RF-EMF induced temperature increase been adequately considered and assessed?” was added because this aspect is especially relevant in the case of RF-EMF exposure to assess confidence in the control of exposure conditions. A customized guide to RoB assessment in the frame of the specific systematic review topic was developed to assist the reviewers as reported in the Supplementary File 10 of Pacchierotti et al., 2021. For the systematic review on RF-EMF effects on human sperm, some specific adaptations for *in vitro* studies were introduced following the indications of Bodewein et al. (2019), Golbach et al. (2016), Romeo et al. (2021).

Following the 3-tier system of study classification proposed by the OHAT, the scores for the different questions were integrated to obtain the study overall RoB estimate. A study was labelled “high concern” when one or more questions were answered with “definitely high RoB”. A study was labelled “low concern” when none of the questions were answered with “probably high RoB” or “definitely high RoB”. All other studies were labelled as “some concern”.

Two reviewers independently analysed included papers for RoB assessment, and disagreements were resolved by discussion with a third reviewer. RoB was evaluated at the endpoint level, meaning that one paper that reported results for different endpoints received multiple RoB evaluations. Whenever necessary to clarify issues relevant for RoB assessment, authors were contacted, and their reply or absence of reply was considered in

the assigned scores. LA, BB, MHB, CC, EC, GC, PE, JPM, FP, PV and AW participated in the RoB assessment.

2.9. Effect measures

All endpoints were expressed as continuous variables, with the exception of binary data regarding the rate of infertile males and the number of nonpregnant females over paired females.

For each endpoint, the preferred effect size measure of continuous variables was the mean difference (MD) that could be applied whenever data were expressed by or could be converted into the same metrics. Standardized Mean Differences (SMD), calculated as MD/pooled SD, were used for data that used different metrics (percentages or numbers) to measure the same endpoint, or when the scale of measures was expected to widely differ, e.g., in the case of testis weight in different species. In the case of the size of litters sired by experimental males, mean values and corresponding variation parameters were referred to the number of males instead of the number of dams to account for intra-cluster correction. For binary variables, we calculated Odds Ratios (OR) instead of Risk Ratios as initially planned in the protocol (see Section 4.5.2). To calculate the OR for the number of nonpregnant females over paired females we applied an intracluster correction (ICC) factor of 0.2 in the design effect formula provided by Golub and Sobin (2020) to account for male clustering.

2.10. Synthesis methods

All included papers were organized in a tabular form by the first author surname in alphabetical order. Studies that were homogenous with regard to the PECO elements were synthesized. With respect to what was planned in the protocol, we could not synthesize results on *in vitro* fertilization rate because only one study was retrieved. We decided to synthesize results on sperm viability and sperm motility into a unique synthesis of sperm vitality because we better considered these endpoints representing the same concept, and we decided to synthesize results on sperm oxidative stress, DNA/chromatin alterations and apoptosis into a unique synthesis of sperm DNA/chromatin alterations because of afterthought consideration of the involvement of DNA damage in all these enpoints (Table 2). In conclusion, for animal studies we made syntheses of results for 14 different endpoints belonging to the 4 outcomes. For human sperm *in vitro* studies, we made a synthesis of results for 3 endpoints belonging to the outcome category of semen quality (Table 3).

Table 3. Distribution of papers and studies by investigated endpoint. Figures in italics correspond to papers on EMP.

Endpoint	**Metrics**	**Studies entered into a meta-analysis**		**Papers presented by a narrative synthesis**
		N° papers (N° studies)[1]	**Effect size measure**	**N° papers**

Decrease of fertility				
Rate of infertile males	Number of males with unsuccessful copulation	4 (7)	OR	
Nonpregnant females over paired females	Number of nonpregnant females over paired females	9 (22) *1 (5)*	OR	*1*
Litter size	Number of offspring per mated female	13 (17) *2 (6)*	SMD	
	Total number of offspring per female after multiple pregnancies	1 (6)		
In vitro fertilization rate	*% in vitro* blastocysts			1
	% in vitro fertilized oocytes			1
Effects on semen quality-experimental animal studies				
Sperm count	Sperm count[2]	41 (121) *3 (14)*	SMD	3
Sperm morphology	% abnormal sperm	27 (92) *3 (13)*	MD	3
Sperm vitality	% immotile or dead sperm	23 (44) *2 (10)*	MD	1
Sperm DNA/chromatin alterations	% DNA fragmented sperm	1 (3)		
	Mean level DNA damage	2 (2)		
	% apoptotic sperm	2 (2)		

Journal Pre-proofs

Effects on semen quality-experimental studies on human sperm in vitro				
Sperm morphology	% abnormal sperm	2 (3)	MD	
Sperm vitality	% immotile or dead sperm	7 (29)	MD	1
Sperm DNA/chromatin alterations	% oxidative stress positive cells	1(2)	SMD	
	% DNA fragmented sperm	5 (16)		
Reproductive organ toxicity				
Testis-epididymis weight	Testis weight (g)	39 (73) *2 (10)*	SMD	7
	mg testis/g body weight	1 (1)		
	% testis/body weight	4 (9)		
	epididymis weight (g)	2 (5)		
	mg epididymis/g body weight	1 (2)		
Testis histomorphometry	Seminiferous tubule diameter (µm)	26 (39) *2 (10)*	SMD	
	Seminiferous tubule area X 10000 μm^2	1 (2)		
	germinal epithelium height (µm)	1 (3)		
Testis or epididymis histology	Johnsen's histopathology score (#)	14 (27)	MD	3
	histopathology score (#)			

Journal Pre-proofs

Testicular cell death	% death[3]	9 (30)	SMD	3 *1*
	Apoptosis gene expression (arbitrary units)	4 (6)		
	Number of dead cells	4 (10) *1 (5)*		
Testicular sperm production	Number of testis sperm per tubule	3 (14)	SMD	
	Number of testis sperm per gram testis (x 10^6)	1 (6)		
	Number of testis sperm per ml (x 10^7)	1 (2)		
	% flow cytometric haploid cells	1 (1)		
	daily sperm production per g of testis (x10^6)	8 (33)		
Hormonal effects				
Testosterone level	Testis testosterone (ng/mg protein)	2 (3)	SMD	1 *2*
	Testis testosterone (ng/ml)	4 (11)		
	Serum testosterone (ng/ml)	25 (39) *2 (23)*		

[1] The number of studies here corresponds to the number of different exposure groups reported in the papers. This number may be higher than the number of studies analysed in the results synthesis because when multiple exposure groups shared the same comparator, their data were averaged and considered as one study only.

[2] Numbers in different studies correspond to a variety of often unclear metrics, including epididymal sperm count/ml (x10^6), N° sperm in 48 small Neubauer chamber squares, relative concentration of epididymal sperm, sperm count (x 0.02 mm^3). For this reason, it was only possible to estimate the RF-EMF impact in terms of SMD, but not to infer the RF-EMF impact on the absolute sperm number.

[3] Including % dead cells, % apoptotic index, % TUNEL positive tubules, % area of caspase-3 immunopositive cells.

Abbreviations: MD: Mean Difference; OR: Odds Ratio; SMD: Standardized Mean Difference.

For each endpoint, we first conducted a meta-analysis of exposed vs sham control comparisons. When a study had several exposure groups matched to the same comparator, the means and standard deviations of these exposed groups were combined into one exposed group using the formulas provided in the paragraph 6.5.2.10 of the Cochrane Handbook (Higgins and Li, 2022), so that each study was entered only once into the meta-analysis. The exposure level assigned to that combined exposed group was calculated as the average SAR of the exposed groups in that study weighed by the number of animals in each exposed group. In the forest plots this is indicated with an asterisk after the study ID. Studies that compared each exposed group to another separate sham control group were entered as separate studies in the meta-analysis. When multiple studies were reported in the same paper, this is indicated with a number after the study ID in the forest plot.

A random-effects meta-analysis model was applied because the underlying effect size was expected to differ between studies due to the explorative nature and diversity in animal studies. Statistical heterogeneity of results was assessed by measures of heterogeneity variance (τ^2, I^2). For the random-effects model, the DerSimonian and Laird between-study variance estimator was used.

All data subject to a meta-analysis were graphically synthesized by forest plots. A forest plot was drawn in which the studies were divided according to their overall RoB level as “low or some concern” or “high concern”. We decided to exclude from the assessment of the pooled effect sizes the studies rated at “high concern” for RoB in order to draw conclusions based upon the most robust data (see Section 4.5.2).

To explore possible causes of heterogeneity, we conducted sub-group analyses according to animal species or fertility status of donors, exposure levels (SAR < 0.1, 0.1 ≤ SAR < 5, SAR ≥ 5 W/kg) and measurements of animal temperature increase below or above 1°C. We limited the subgroup analysis to these 3 variables, because they were considered the most important variables to affect possible associations between exposure and outcomes and to keep the work manageable. This is a slight deviation from the protocol (see Section 4.5.2).

According to the protocol, subgroup analyses were only interpreted when all the sub-groups included at least 3 studies.

Studies that tested the effect of Electromagnetic Pulses (EMP) were analysed separately from other studies.

Next, we conducted a dose-response meta-analysis as described by Orsini and Spiegelhalter (2021) and implemented in STATA (STATA/BE 17.0 by StataCorp LLC, College Station, TX, USA, 2022). We specified a model based on an assumed linear relation between the SAR and the outcome. We also specified a non-linear model based on cubic splines. To assess if the non-linear model fit better than the linear model, we used the difference between the Akaike’s Information Criterion (AIC) of the models. Finally, we visualised the summary estimate of the linear and the non-linear model together with the individual study dose-response curves in one graph based on the best linear unbiased

prediction. We compared the predicted effects at a dose of 1 W/kg to the other doses over a range of 0 to 10 W/kg.

We used STATA 17 for the meta-analysis and the dose-response meta-analysis.

2.11. Reporting bias assessment

We assessed reporting publication bias in all the studies retrieved, irrespective of their overall RoB level of concern, to enlarge as much as possible the database and increase the sensitivity of our analysis. To visualise possible publication bias, funnel plots of the study effect size measures against their standard errors were produced when at least 5 studies were available. If the funnel plot, upon visual inspection, showed that more imprecise studies with non-harmful effects were missing, this was considered an indication of possible publication bias. If ten or more studies were included in the same meta-analysis, an Egger's test was applied to evaluate potential small study bias, otherwise a qualitative evaluation was made (Egger et al., 1997).

2.12. Unplanned analyses

To investigate the influence on the pooled effect size of the lack of blinding for the analysis of outcomes entailing a subjective component, we compared the average effect size of studies that were specifically considered reliable for outcome assessment (probably or definitely low RoB to question 3.3 of the OHAT RoB tool, see Supplementary File 2a) with the average effect size of studies that were not considered fully reliable for outcome assessment (probably high RoB to the same question). We did this analysis for the datasets of "low or some concern" studies that included a good number of entries in each group; they were studies on sperm morphology, testicular cell death and testicular sperm production.

We calculated the average exposure level tested in each endpoint. For studies in which multiple exposure groups shared the same comparator, we entered the average SAR of the exposed groups in that study weighed by the number of animals in each exposed group.

2.13. Certainty assessment

The GRADE (Grading of Recommendations, Assessment, Development and Evaluations) framework for developing and presenting summaries of evidence was used to judge the certainty in the evidence of the effects observed in the systematic review and to draw conclusions (https://www.gradeworkinggroup.org). GRADE was initially developed for clinical studies and its application to animal toxicological studies is still under development. Toxicological studies pose a challenge to the GRADE approach because they are much less standardised than clinical studies. We started the rating of the certainty of the evidence at high certainty as is performed in human experimental studies (Hooijmans et al., 2018). Five domains were considered: RoB in the studies, indirectness considering how well the PECO question has been addressed from both the animal and human perspective, inconsistency, imprecision and publication bias. Depending on which criteria for which domains were met, we downgraded the certainty of the evidence to moderate, low, or very low according to the Supplementary File 11 of Pacchierotti et al., 2021. The only upgrading factor considered was consistency among animal species. Although we explored dose-response relationships, we did not apply evidence of a dose dependent effect as a further upgrading factor, because assessment of dose dependency was not considered by the PECO question and GRADE evidence profiling already started from high certainty, as indicated for experimental animal studies.

We ranked the endpoints according to a scale from 1 (the lowest importance) to 10 (the highest importance) in relation to the ultimate human relevant outcome as proposed by Guyatt et al. (2011) and we considered the ranking in the assessment of the indirectness domain. This was an operational elaboration of what expressed in the protocol regarding the importance attributed to the various outcomes.

3. Results

3.1. Study selection

Figures 2 and 3 show the flow diagrams for animal and human sperm *in vitro* studies, respectively, from the initially retrieved references to the finally included papers, as per the PRISMA 2020 template (Page et al., 2021).

We retrieved a total of 1335 different papers reporting animal studies. Title/abstract selection reduced this number to 323. Despite our search of electronic databases and attempts to contact the authors, we could not retrieve 20 papers and could not appropriately translate 11. Full text selection further reduced the database to a total of 117 included papers.

We retrieved a total of 869 different papers reporting studies on human sperm exposed *in vitro*. After title/abstract and full text selection, the database included 10 studies.

3.2. Excluded studies

After reading the full text, 175 papers on animal studies were excluded. They are listed in Supplementary File 1a with a justification of the exclusion rationale together with those not retrieved or not translated. Over 45% of the animal studies were excluded because essential information was missing regarding exposure set-up and/or dosimetry, e.g., details on how the exposure system output was established and maintained or exposure frequency. A further 27% of the studies were excluded because outcome data were deemed out-of-scope or invalid. For instance, genotoxic effects in testicular cells and effects on hormones other than testosterone were considered out-of-scope because they are too weakly linked with male infertility. We considered invalid methods those which were insufficiently described, insufficiently validated or improperly applied. Examples were non-standardized cellular and molecular markers of stress in testicular cells, improperly applied flow cytometric analysis of the haploid testicular cell population, or testicular histomorphometry assessment in an insufficient number of seminiferous tubule sections. Other papers were excluded because the experimental design was invalid for the scope of the systematic review, e.g., when non-experimentally controlled environmental exposure was investigated or when both males and females of a mating pair were exposed, which did not allow sorting out of specific effects on the male reproductive system.

Regarding studies on human sperm *in vitro*, we excluded 33 papers after reading the full text (Supplementary File 1b). Most papers were excluded because they did not report peer-reviewed original results. Other papers could not be included in the systematic review because exposure conditions and/or dosimetry were insufficiently reported or because the exposure conditions did not provide a sufficient exposure contrast between RF-EMF exposed and sham-exposed samples.

3.3. Study characteristics

Tables 4a and 4b list the included papers on animal and human sperm *in vitro* studies, respectively. In these tables, characteristics of the studies regarding populations, exposure and outcomes are presented. In particular, the tables report information on the animal species or the fertility status of the sample donors, the average sample size, the exposed life-stage of the animals or the donor age, the RF-EMF frequency and level(s) tested and the duration of exposure. Table 4a also shows which papers investigated exposure to EMP. Additionally, a very brief description of the main results in scope for the systematic review is reported, based on the authors' interpretation and discussion, together with the evidence or not of a temperature increase in the exposed animals/samples, when determined. The tables also show the specific outcomes investigated and whether the data were entered into the meta-analyses or not. A few results could not be synthesized in the form of meta-analysis because actual data were not reported and could not be retrieved even after contacting the authors. Other results were extracted from some papers but were not entered into a meta-analysis because they were considered a less relevant measure of the outcome or pertained to endpoints assessed in one paper only. Nevertheless, all these results have been synthesized in a narrative way.

The papers reporting animal studies were published between 1962 and 2022 while the papers on human sperm *in vitro* were published between 1980 and 2016.

Population

Seventy-six papers reported studies in rats, 38 in mice, 1 in hamster, 1 in guinea pig and 1 in rabbit. In the majority of studies, animals were exposed only during adulthood, in 14 studies animals were exposed totally or partially before birth and in 15 studies animals were exposed totally or partially before puberty.

Almost all studies on human sperm exposed *in vitro* tested RF-EMF effects on samples collected from fertile donors, in 3 papers results were reported on samples from subfertile donors.

Exposure

Only 5 papers on experimental animals investigated frequencies equal to or higher than 6000 MHz. Ninety-seven percent of studies were conducted in the 100 MHz – 10 GHz frequency range and 81% clustered around the interval 900-2450 MHz, the frequencies used in mobile phone communication or wi-fi systems. Only 17 out of 117 papers investigated effects at more than 1 exposure level. The levels of exposure were expressed as SAR for all but a few papers. The SAR values were either reported by the authors or calculated based on the exposure characteristics and other available dosimetric information. They corresponded always to whole-body average (WBA) values, including the few cases of local exposure. There is no clear relationship between testes temperature and WBA-SAR. Exposure location (e.g., focussed on testis), frequency, duration of exposure and ambient temperature could all influence the temperature in testes, so WBA-SAR is to be regarded as a proxy to the exposure level to the target tissue and not an indicator of testis temperature. Exposure levels ranged between 0.000012 and 184 W/kg, with the oldest studies mainly interested in the hyperthermic RF-EMF effect induced at very high SAR levels. The exposure duration was also highly variable, spanning from 1 day to 1 year. Seven papers explored exposure to EMP.

Eight of the 10 papers on human sperm *in vitro* investigated the effects of frequencies applied in mobile phone communications. In most studies the exposure time was short,

between 1 and a few hours, while 1 study exposed sperm for 16 hours. Only 1 paper aimed at exploring the dose effect relationship.

Comparators

All the studies in experimental animals used as the comparator a sham-exposed group of animals. Similarly, all studies on human sperm used as the comparator a sham-exposed sample, which in all but one study was an aliquot of the same semen sample used for RF-EMF exposure.

No study included animals or sperm samples exposed to direct heating at a temperature comparable to that induced by RF-EMF.

Outcomes

Experimental animal studies

We considered 4 different outcomes as relevant for assessing animal male fertility potential: i) direct effects on fertility, ii) effects on semen quality assessed by WHO recommended parameters or other biomarkers, iii) reproductive organ toxicity, iv) hormonal effects assessed by testosterone level. Many papers reported results on more than one outcome.

Reduction of fertility

Twenty-one papers reported results of tests directly assessing male fertility. Endpoints diagnostic of male fertility were the rate of infertile males (4 papers), the incidence of nonpregnant females over paired females (11 papers), the litter size (16 papers) or *in vitro* fertilization and embryonic development (1 paper). With the exception of *in vitro* embryonic development that was investigated in 1 study only, we included all the other papers in meta-analyses.

Effects on semen quality

Fifty-five papers contained data on one or more parameters of sperm quality defined by WHO guidelines for andrological analyses. In particular, sperm count was reported in 47 papers, frequency of morphologically abnormal sperm was reported in 33 papers and 26 papers contained data on sperm vitality. In 5 papers sperm quality was assessed by DNA/chromatin alterations. Almost all the papers reported data in a form suitable for a meta-analysis.

Reproductive organ toxicity

Eighty-six papers reported data on one or more toxicity biomarkers detected in testis or epididymis. Fifty-six papers reported data on testis or epididymis weight (data on epididymis weight were reported in further papers, but they were considered redundant when sperm count was also reported). Thirty papers reported data on testis histomorphometry, mainly as seminiferous tubule diameter. Seventeen papers reported semi-quantitative data on testis histopathology, mainly expressed by Johnsen's score (no data on epididymis histopathology were retrieved). Twenty-two papers contained data on testicular cell death. Fourteen papers reported data on the measured or estimated testicular sperm production.

Hormonal effects

The level of testosterone was chosen as the most meaningful biomarker of possible hormonal RF-EMF effects. Testosterone was measured in either testis (6 papers) or serum (30 papers). Both endpoints were considered reliable estimates of the hormone synthesis and secretion.

Journal Pre-proofs

Table 4a. List of included papers on male fertility studies in experimental animals with main study characteristics.

Reference	**Population**		**Exposure**			**Outcome**				*Summary of paper results*
	Species (Average group size)	*Stage of development during exposure: pre-natal (PN), before puberty (BP), after puberty (AP)*	*Frequency (MHz)/ Modulation (M, CW) or EMP*	*Level(s) W/kg*	*Duration(s) Hours per day /N° of days*	*Fertility*	*Semen quality*	*Organ toxicity*	*Hormonal effects*	
Aitken et al., 2005	Mouse (5)	AP	900	0.09	12:00/7					No effect on sperm quality. No effect on testis and epididymis weight. Evidence of DNA damage in epididymal sperm by not internationally validated biomarkers.
Akdağ et al., 1999	Rat (10)	AP	9450/CW	1.8	1:00/13, 26, 39, 52		X	X		About 1°C temperature increase in some but not all exposed groups. Effect on sperm quality starting from 26 exposure days. Effect on testis weight only after 26 exposure days. Some qualitative evidence of histopathological effect in testis and epididymis.
Almášiová et al., 2017	Rat (10)	AP	2450/M	0.5*	3:00/21		X			No temperature increase. No effect on sperm motility. Some qualitative evidence of histopathological effect in testis

Journal Pre-proofs

Almášiová et al., 2021	Rat (6)	PN	2450/M	1.82	2:00/21			X		Effect on testis histomorphometry. Some qualitative evidence of histopathological effect in testis. Some evidence of oxidative stress in testis considered out-of-scope for the systematic review aims.
Andrašková et al., 2021	Rat (6)	PN	2450/M	1.73	2:00/21			X		Decrease of seminiferous tubule diameter. Descriptive epididymis histopathology suggestive of an effect.
Atasoy et al., 2013	Rat (5)	AP	2437	0.091	24:00/140			X		Histopathological effect in testis. Some evidence of oxidative stress in testis considered out-of-scope for the systematic review aims.
Azimzadeh et al., 2019	Rat (10)	AP	900	0.332	2:00, 4:00/30				X	Decrease of testis testosterone level. Changes in testis cytokine level considered out-of-scope for the systematic review aims.
Beechey et al., 1986	Mouse (4)	AP	2450/M	0.05, 5, 20	0:30/12		X	X		3°C temperature increase at 20 W/kg. Small increase of sperm count. No effect on sperm morphology. No effect on testis weight. No effects on chromosomal aberrations in spermatocytes considered out-of-scope for the systematic review aims.
Berman et al., 1980	Rat (12)	AP, PN+BP+AP	2450/CW	1.35, 2, 5.6	4:00, 5:00/5, 20, 106	X	X	X		4.5°C temperature increase at 5.6 W/kg. Decrease of litter size only after 20 days of exposure of adult animal to the highest level. No effect on sperm count and viability. No effect on testis weight.

Journal Pre-proofs

Bilgici et al., 2018	Rat (11)	AP	2450/CW	0.023	1:00/30			X		Effect on testis histopathology. No change in testis cytokine level considered out-of-scope for the systematic review aims.
Cairnie and Harding 1981	Mouse (5)	AP	2450/CW	13*, 14*, 18*	1:00, 2:00, 4:00, 8:00, 16:00/1, 2, 4, 8, 30		X	X		No temperature increase. No effect on sperm quality or testis cell viability.
Cao et al., 2005	Mouse (10)	AP	947/CW	0.2*, 0.4*	2:00/35			X		No effect on testis or epididymis weight. Descriptive testis histopathology suggesting lack of effect.
Çetkin et al., 2017	Rat (8)	AP	900/M	0.96	2:00/70			X	X	No temperature increase. Effect on testis weight, seminiferous tubule diameter and histopathology. No effect on serum testosterone.
Chaturvedi et al., 2011	Mouse (5)	AP	2450/CW	0.036	2:00/30		X			No effect on sperm quality.
Chen et al., 2014	Mouse (5)	AP	1800/CW	0.222	2:00/32			X	X	Variabile effects on testicular sperm count and serum testosterone as a function of exposure time during the day.
Cobb et al., 2000	Rat (4)	PN	EMP	0.045	0:02/16	X				No temperature increase. Decrease of % mated females and no difference in % fertile matings and litter size.

Journal Pre-proofs

Dasdag et al., 2003	Rat (8)	AP	900/M	0.52	0:20/30		X	X		No temperature increase. No effect on sperm count and morphology. No effect on testis histopathology.
Dasdag et al., 2008	Rat (10)	AP	900/M	0.32	2:00/305			X		No increase of apoptosis in testis.
Dasdag et al., 1999	Rat (6)	AP	900/M	0.141	0:03/30		X			0.2°C temperature increase. No effect on sperm count and morphology. Evidence of effect on testis histopathology of insufficient quality to be considered for the systematic review.
Dasdag et al., 2015	Rat (8)	AP	2400/CW	0.002	24:00/365		X	X		No effect on sperm quality. No effect on testis weight and histopathology; decrease of seminiferous tubule diameter.
Delavarifar et al., 2020	Mouse (6)	AP	2400/M	0.092	2:00/4		X	X		Increase of sperm count; no effect on sperm viability and motility. No effect on seminiferous tubule diameter.
Dong et al., 2021	Mouse (10)	AP	EMP	50, 100, 300 W/m^2	0:30/1		X			Temperature increase between 0.2 and 0.5°C as a function of exposure level. Effect on sperm motility at the two highest exposure levels. No effect on serum testosterone.

Journal Pre-proofs

Er et al., 2022	Rat (6)	AP	2100/CW	1.159	2:00/5, 50			X		No temperature increase. Increase of apoptosis in testis after 5 but not 50 exposure days. Descriptive testis histopathology suggesting lack of effect.
Erdemli et al., 2017	Rat (6)	AP	2100/CW	0.36	0:30/24, 48			X		No temperature increase. No effect on epididymis weight. Descriptive epididymis histopathology suggestive of an effect.
Fahim et al., 1975	Rat (10)	AP	2450	30 at 100% power, other groups at 20% power	0:01, 0:05, 0:15/1				X	Testis temperature increased up to 65°C. No changes of sex organ weights (testis, epididymis, prostate, and seminal vesicle). No effect on serum testosterone. Effect on spermatogenesis and fertility as a function of duration and intensity of exposure up to total sterility.
Forgács et al., 2006	Mouse (54)	AP	1800/M	0.02	2:00/10				X	Increase of serum testosterone. Descriptive testis histopathology suggesting lack of effect.
Gao et al., 2016	Rat (10)	AP	900	0.15*	2:00, 4:00/30			X		Increase of apoptosis in testis. Descriptive testis histopathology suggestive of an effect.
Gautam et al., 2021	Rat (6)	AP	2115/M	0.159	2:00/45		X	X		No effect on sperm count, motility and morphology. Effect on sperm viability. No effect on testis weight and seminiferous tubule diameter. Descriptive testis histopathology suggestive of an effect.

Journal Pre-proofs

Gautam et al., 2019	Rat (8)	AP	1915/M	0.26	2:00/45		X	X		No effect on sperm count and morphology. Effect on sperm viability. Effect on testis weight. Descriptive testis histopathology suggestive of an effect.
Ghanbari et al., 2013	Rat (7)	AP	915/M	0.6*	6:00/14, 21		X			No effect on sperm count and morphology; effect on sperm viability and motility.
Goud et al., 1982	Mouse (16)	AP	2450/CW	85*	0:01/1	X	X			Effect on sperm morphology. Decrease of male fertility; increase of pre- and post-implantation losses.
Guo et al., 2019	Rat (12)	AP	220/M	0.03	1:00/30		X	X	X	<0.5°C temperature increase. No effect on sperm morphology; effect on sperm count and viability. No effect on testis weight and histomorphometry; increase of apoptosis in testis. Decrease of testis testosterone. Descriptive testis histopathology suggesting lack of effect.
Gur et al., 2021	Rat (8)	BP	900	0.01	1:00/26			X		No effect on testis histomorphometry and apoptosis. Descriptive testis histopathology suggesting lack of effect.
Hanci et al., 2018	Rat (8)	BP+AP	900/CW	0.007	1:00/39			X		Effect on testis weight, histomorphometry and histopathology; increase of apoptosis in testis.
Hanci et al., 2013	Rat	PN	900/CW	0.01*	1:00/9			X		No effect on testis weight; increase of apoptosis in testis. Descriptive testis histopathology suggestive of an effect.

Journal Pre-proofs

	(10)									
Houston et al., 2019	Mouse (6)	AP	905	2.2	12:00/7, 21, 35		X			No effect on in vitro fertilization. Effect on sperm viability and motility. Increase of DNA damage and oxidative stress in sperm. Descriptive testis histopathology suggesting lack of effect.
Huai and Min 1984	Mouse (7)	AP	2450/CW	9.5, 15	0:30/1		X			Temperature increase as a function of exposure level. Effect on sperm morphology at the high exposure level.
Imai et al., 2011	Rat (24)	BP+AP	1950/M	0.08, 0.4	5:00/35		X	X		No effect on sperm quality. No effect on testis weight; increase of testicular sperm number. Descriptive testis histopathology suggesting lack of effect.
Jensh 1984	Rat (15)	PN	6000	7.28	8:00/13	X		X		No temperature increase. No effect on litter size. No effect on testis weight.
Jensh et al., 1982	Rat (15)	PN	915/CW	3.5*	8:00/14	X		X		No temperature increase. No effect on litter size. No effect on testis weight.
Jensh et al., 1983	Rat (17)	PN	2450/CW	4*	8:00/14	X		X		No temperature increase. No effect on litter size. No effect on testis weight.

Journal Pre-proofs

Jin et al., 2013	Rat (20)	AP	849, 849+1950/M	4	0:45/20, 40				X	No effect on serum testosterone. No effect on other serum hormones out of scope for the systematic review aims.
Johnson et al., 1984	Rat (14)	AP	1300/M	6.3	6:00/9			X		1.5°C temperature increase. No effect on testis weight and daily sperm production. Descriptive testis histopathology suggesting lack of effect.
Jonwal et al., 2018	Mouse (8)	AP	2450/CW	0.09	2:00/30				X	Decrease of serum testosterone. Descriptive testis histopathology suggestive of an effect.
Kesari and Behari 2012	Rat (6)	AP	900/M	0.9	2:00/45				X	Decrease of serum testosterone.
Khillare and Behari 1998	Rat (9)	AP	200/M	1.82	2:00/32	X	X		X	Decrease of mating rate and litter size. Effect on sperm count and motility. No effect on serum testosterone. Descriptive testis histopathology suggestive of an effect.
Kim et al., 2007	Rat (10)	BP+AP	2450	1.4	1:00, 2:00/56		X	X	X	1°C temperature increase. No effect on sperm count. No effect on testis weight, histomorphometry and histopathology. Increase of serum testosterone after 2 hour exposure.
Kismali et al., 2009	Guinea pig	AP	NR	0.81	0:20/7				X	No effect on serum testosterone.

Journal Pre-proofs

	(6)									
Kowalczuk et al., 1983	Mouse (5)	AP	2450	43.6	0:30/1		X			8°C temperature increase. Effect on sperm count. Effect on sperm morphology after exposure of meiotic and postmeiotic stages.
Kumar et al., 2013	Rat (6)	AP	10000/CW	0.014	2:00/45		X	X	X	Increase of sperm DNA damage; suggestive evidence of increase of apoptosis in sperm. Effect on testis weight and histomorphometry. Decrease of serum testosterone.
Kumar et al., 2011	Rat (3)	AP	2450/M	0.014	2:00/60				X	Decrease of serum testosterone. Suggestive evidence of increase of apoptosis in sperm.
Kumar et al., 2014	Rat (6)	AP	1910/M	0.15	2:00/60		X	X		Effect on sperm count. Increase of sperm DNA damage. Effect on testis weight and histomorphometry. Descriptive testis histopathology suggesting lack of effect.
L'Abbate et al., 1982	Rabbit (7)	AP	100	3*	8:00/90			X		Suggestive evidence of a decrease of seminiferous tubule diameter. Descriptive testis histopathology suggestive of an effect.
Lebovitz and Johnson 1983	Rat (4)	AP	1300/M	6.6	6:00/9		X	X		2°C temperature increase. No effect on sperm count and morphology. No effect on testis weight and daily sperm production.

Journal Pre-proofs

Lebovitz and Johnson 1987a	Rat (4)	AP	1300/CW	9	8:00/1		X	X		3°C temperature increase. Effect on sperm count after exposure of late meiotic stages. No effect on testis weight and daily sperm production.
Lebovitz et al., 1987b	Rat (5)	AP	1300/M	4.2, 7.7, 8.5	1:30/1		X	X		Temperature increase between 1 and 4°C as a function of exposure level. Effect on sperm count after exposure of meiotic and postmeiotic stages to the highest level. Variable effects on testis weight and daily sperm production as a function of post-exposure time.
Lee et al., 2012	Rat (20)	BP+AP	848.5+1950/M	4	0:45/60		X	X	X	0.4°C temperature increase. No effect on sperm count and morphology. No effect on testis weight and apoptosis. No effect on serum testosterone. Descriptive testis histopathology suggesting lack of effect.
Lee et al., 2010	Rat (20)	BP+AP	848.5/M	2	1:30/60		X	X		No temperature increase. No effect on sperm count. No effect on testis weight, histomorphometry and apoptosis. Descriptive testis histopathology suggesting lack of effect.
Lee et al., 2005	Mouse (3)	AP	848.5, 1763/M	0.4	1:30/20, 40, 50					Suggestion of lack of effect on testis histopathology and apoptosis.
Lerchl et al., 2008	Hamster (120)	AP	383, 900, 1800/M	0.08	24:00/60					No effect on testis weight.

Journal Pre-proofs

Li et al., 2017	Mouse (10)	AP	EMP	NR	/14	X	X	X	X	0.2°C temperature increase. Effect on number of pregnant females after exposure of early and middle postmeiotic stages. No effect on litter size. No effect on sperm count and morphology. No effect on testis weight. Effect on seminiferous tubule diameter. Increase of serum testosterone.
Liu et al., 2015	Rat (24)	AP	900/M	0.66	2:00/50		X			Effect on sperm count; no effect on sperm morphology. Increased apoptotic sperm percentage.
Luo et al., 2013	Mouse (8)	AP	EMP	NR	/1					Increase of testis apoptosis at the highest exposure level. Descriptive testis histopathology suggestive of a dose dependent effect.
Ma et al., 2015	Rat (10)	AP	900	0.1	4:00/18			X		Increase of testis apoptosis. Descriptive testis histopathology suggestive of an effect.
Ma et al., 2014	Rat (6)	AP	900/CW	0.12*	4:00/16	X				Effect on number of pregnant females. No effect on litter size.
Meena et al., 2014	Rat (6)	AP	2450/M	0.14	2:00/45			X	X	Effect on testis weight. Decrease of testis testosterone. Descriptive testis histopathology suggestive of an effect.

Journal Pre-proofs

Miao et al., 2017	Mouse (6)	BP+AP	EMP	NR	/20		X	X		Effect on sperm count and morphology. Effect on testis weight at late recovery times. Decrease of seminiferous tubule diameter at early recovery times. Effect on testis apoptosis variable at different recovery times.
Odaci and Ozyilmaz 2015	Rat (8)	AP	900	0.025	1:00/30			X		No effect on testis weight; decrease of seminiferous tubule diameter; increase of testis apoptosis; worsening of histopathology scores.
Odacı et al., 2016	Rat (9)	PN	900	0.025	1:00/9		X	X		Effect on sperm count, motility and viability. Decrease of testis weight and seminiferous tubule diameter; increase of testis apoptosis. Descriptive testis histopathology suggestive of an effect.
Oh et al., 2018	Rat (4)	BP+AP	2104/CW	3	6:00, 18:00/28		X	X		No temperature increase. Effect on sperm count after the longest exposure at the shortest distance. No effect on testis weight; worsening of the histopathological score after the longest exposure at the shortest distance.
Oksay et al., 2014	Rat (8)	AP	2450/M	0.143	1:00/30			X		No effect on testis weight.
Ozguner et al., 2005	Rat (10)	AP	900/CW	0.35*	0:30/20			X	X	No effect on testis weight and histopathology; effect on histomorphometry. Decrease of serum testosterone.

Journal Pre-proof

Ozlem Nisbet et al., 2012	Rat (11)	BP+AP	900, 1800/CW	0.000012, 0.002	2:00/90		X		X	No effect on sperm count; improvement of sperm forward motility; decrease of sperm abnormal morphology at 900 MHz. Worsening of histopatological score at 1800 MHz. Increase of serum testosterone.
Pandey and Giri 2018	Mouse (5)	AP	902.4/M	0.029	6:00/35		X	X		Effect on sperm count and morphology. Decrease of haploid cells in testis. Descriptive testis histopathology suggestive of an effect.
Pandey et al., 2017	Mouse (5)	AP	902.4/M	0.029	4:00, 8:00/35		X	X		Effect on sperm number at the longest daily exposure; increase of percent abnormal sperm. Decrease of testis weight. Descriptive testis histopathology suggestive of an effect.
Pardhiya et al., 2022	Rat (6)	AP	2002/CW	1.2	2:00/48		X	X	X	Effect on sperm number, viability and morphology. Decrease of testis weight; no effect on testis histomorphometry. Decrease of serum testosterone.
Pedrosa et al., 2021	Rat (8)	AP	27.12	NR	0:15/15, 30, 60			X	X	No effect on testis weight and histomorphometry; increased daily sperm production at the shortest exposure. No effect on serum testosterone. Descriptive testis histopathology suggesting lack of effect except for some effects on Leydig cells.
Poulletier de Gannes et al., 2013	Rat (12)	AP	2450/M	0.08, 4	1:00/36			X		No effect on testis weight. Descriptive testis histopathology suggesting lack of effect.

Journal Pre-proofs

Prausnitz and Susskind 1962	Mouse (40)	AP	9300/M	1000 W/m^2	0.05/295					3.3°C temperature increase. Comments on testis weight and descriptive testis histopathology suggestive of an effect.
Qin et al., 2018	Mouse (6)	AP	1800/CW	0.055	1:00, 2:00/32				X	Decrease of testis testosterone.
Qin et al., 2014	Rat (6)	AP	1800/CW	0.041	2:00/32		X	X	X	Variable effects on sperm motility as a function of exposure time during the day. Effect on daily sperm production. Variable effects on testis testosterone as a function of exposure time during the day.
Qin et al., 2012	Rat (6)	AP	1800/CW	0.576	2:00/32				X	Variable effects on serum testosterone as a function of exposure time during the day and blood sampling daytime.
Qin et al., 2021	Mouse (9)	BP+AP	1800/CW	0.5	2:00/21			X	X	Variable effects on testis weight as a function of exposure time during the day. Decrease of daily sperm production. Variable effects on testis testosterone as a function of exposure time during the day.
Ren et al., 2020	Rat (10)	AP	900/CW	3.7 W/m^2	4:00, 8:00/15			X		No effect on testis weight.

Journal Pre-proofs

Ren et al., 2002	Mouse (NR)	AP	2450	100 W/m^2	0:30, 1:00, 1:30/6, 12, 18					No effect on sperm count; effects on sperm morphology. No effect on testis weight.
Ribeiro et al., 2007	Rat (8)	BP+AP	1840/M	0.001*	1:00/77		X	X	X	No temperature increase. No effect on sperm count. No effect on testis weight and histomorphometry. No effect on serum testosterone. Descriptive testis histopathology suggesting lack of effect.
Rugh and McManaway 1978	Mouse (20)	PN, AP	2450/CW	104.53, 106.25, 106.33, 109.87, 156.62, 183.89	0:02, 0:04/1	X				2-4°C temperature increase. No effect on litter size.
Saunders et al., 1983	Mouse (21)	AP	2450	43.4	0:30/1	X				7.4°C temperature increase over anesthetized control. Decrease of pregnancy rate mainly after exposure of meiotic and pre-meiotic stages. Effect on pre- but not on post-implantation survival.
Saunders and Kowalczuk 1981a	Mouse (4)	AP	2450/CW	18, 29.5, 36.7, 45.7, 57.3, 74.7	0:30/1		X	X		Temperature increase from 2.7 to 9.6°C as a function of the exposure level. Some decrease of sperm number. Dose dependent decrease of testis spermatids. Descriptive testis histopathology suggestive of an effect.
Saunders and Kowalczuk 1981b	Mouse (4)	AP	2450/M	7, 17, 33, 50, 66	0:05, 0:15, 0:45, 2:15, 4:20/1		X	X		No temperature increase at the lowest exposure level, then increase from 1.5 to 4.5°C as a function of the exposure level and duration. No effect on sperm count. No effect on the number of testis spermatids. Descriptive testis histopathology suggesting lack of effect.

Journal Pre-proofs

Saunders et al., 1988	Mouse (10)	AP	2450/CW	5	6:00/20	X		X		0.18°C temperature increase. No effect on pregnancy rate, pre-implantation and post-implantation losses. No effect on testis weight.
Saygin et al., 2016	Rat (12)	AP	2450/M	3.21	3:00/30				X	No effect on testis testosterone. Descriptive testis histopathology suggestive of an effect.
Saygin et al., 2011	Rat (6)	AP	2450/M	3.21	1:00/28			X		No effect on testis histomorphometry. Worsening of testis histopathology score.
Saygin et al., 2015	Rat (6)	AP	2450/M	3.21	1:00/28					Increase of apoptosis in testis.
Sepehrimanesh et al., 2014	Rat (5)	AP	900/CW	0.7	1:00, 2:00, 4:00/30				X	No effect on serum testosterone.
Shahin et al., 2019	Rat (8)	AP	900	1.075	2:00/56		X	X	X	Effect on sperm count, motility and morphology. No effect on testis histomorphometry; increase of testis apoptosis. Decrease of serum testosterone. Descriptive testis histopathology suggestive of an effect.
Shahin et al., 2014	Mouse (15)	AP	2450/CW	0.018	2:00/30		X			No temperature increase. Effect on sperm count and viability. Decrease of serum testosterone. Descriptive testis histopathology suggestive of an effect.

Journal Pre-proofs

Shahin et al., 2018a	Mouse (10)	AP	2450/CW	0.015	2:00/15, 30, 60		X	X	X	No temperature increase. Effect on sperm count and viability. Effect on testis histomorphometry; increase of testis apoptosis. Decrease of serum testosterone. Descriptive testis histopathology suggestive of an effect.
Shibkova et al., 2015	Mouse (5)	AP	925/M	0.4*	0:10/5	X				No effect on litter size.
Shirai et al., 2014	Rat (15)	PN+BP	2140/M	0.065, 0.195	20:00/58	X		X		No effect on successful copulation rate. No effect on testis weight.
Shirai et al., 2017	Rat (12)	PN+BP	multiple frequencies/M	0.08, 0.402	20:00/58	X				No effect on successful copulation rate. No effect on testis weight.
Šimaiová et al., 2019	Rat (6)	BP	2450/M	0.6*	2:00/21			X		Effect on testis histomorphometry. Increase of testis apoptosis when exposure started at PND 21 but not when exposure started at PND 14. Descriptive testis histopathology suggestive of an effect.
Smialowicz et al., 1981	Rat (20)	PN+BP+AP	100/CW	2.59	4:00/106	X				No effect on pregnancy rate and litter size.
Sommer et al., 2009	Mouse	AP	1966/M	0.101, 0.504, 1.642	24:00/80			X		No effect on testis and epididymis weight; no effect on number of testis spermatids.

Journal Pre-proofs

	(31)									
Takahashi et al., 2010	Rat (11)	PN+BP	2140/M	0.051, 0.119	20:00/35	X				No effect on successful copulation rate. No effect on testis weight.
Tang et al., 2022	Mouse (5)	AP	34500	5*	2:00/28, 35, 42, 49, 56, 63		X	X		No temperature increase. Effect on sperm count only after the longest exposure. Effect on sperm morphology except for the shortest exposure duration. Effect on sperm motility after the longest exposure times. Decrease of seminiferous tubule diameter, worsening of histopathological score and increase of testis apoptosis after the longest exposure times.
Tas et al., 2014	Rat (7)	AP	900/M	0.04	3:00/365		X	X		No effect on sperm number and motility. Decrease of % morphologically abnormal sperm. No effect on testis weight and histomorphometry; worsening of testis histopathological score.
Trosic et al., 2013	Rat (9)	AP	915/M	0.6	1:00/14		X	X		No temperature increase. No effect on sperm count, motility and morphology. No effect on testis weight and histopathological score.
Tumkaya et al., 2016	Rat (6)	AP	900/M	0.48	1:00/45			X		No effect on testis weight and apoptosis. Descriptive testis histopathology suggesting lack of effect.
Wang et al., 2003	Mouse	AP	EMP	NR	0:02/1				X	Decrease of serum testosterone at all exposure levels.

Journal Pre-proofs

	(6)									
Xu et al., 2020	Mouse (7)	AP	1800	0.299	2:00/7		X	X		Effect on sperm motility and morphology. No effect on daily sperm production.
Xue et al., 2022	Rat (9)	AP	5800/CW	1.15	1:00/30		X	X	X	Less than 1°C temperature increase. No effect on sperm count and morphology. No effect on testis weight. No effect on serum testosterone level.
Yahyazadeh and Altunkaynak 2019	Rat (6)	AP	900/M	2	1:00/28		X	X	X	Effect on sperm morphology. Effect on testis weight. Decrease of serum testosterone. Descriptive testis histopathology suggestive of an effect.
Yan et al., 2022	Mouse (6)	AP	2000/CW	0.31	3:00/98	X	X			No effect on fecundity. No effect on sperm count. No effect on sperm apoptosis. No effect on testis apoptosis.
Yu et al., 2020	Rat (15)	AP	2605	1.05	0:36/50, 100, 150	X	X	X		No temperature increase. No effect on successful copulation rate and litter size. Effect on sperm quality at the longest exposure durations. Effect on testis histomorphometry and worsening of the histopathological score after the longest exposure duration.
Zeng et al., 2011	Rat (7)	AP	EMP	0.03	0:43, 1:40, 2:13/1		X			No temperature increase. No effect on sperm count, motility and morphology. Erratic effects on serum testosterone under the various exposure conditions.

Journal Pre-proofs

Zhu et al., 2015	Mouse (10)	AP	900/CW	0.731	4:00/15	X		X		No temperature increase. No effect on the number of pregnant females. No effect on testis weight.

* SAR calculated for 1 or more of the exposed groups from different original exposure metrics

Gray cells mark outcomes for which each paper contributed data. When data were entered into a meta-analysis this is marked by X. Information on modulation is not reported when it was ambiguously provided in the paper. Exposure level(s) are expressed in SAR unless this data was ambiguously provided in the paper or calculated on the basis of other information, in which case PD values are reported if provided.

Abbreviations: AP: after puberty; BP: before puberty; CW: continuous wave unmodulated signal; EMP: electromagnetic pulses; M: modulated signal; NR: not reported; PD: power density; PN: pre-natal; PND: post-natal day.

Journal Pre-proofs

Studies on human sperm *in vitro*

As with experimental animal studies, WHO sperm quality parameters and sperm DNA/chromatin alterations were considered reliable indicators of semen quality for human sperm *in vitro* studies.

Two papers reported data on percentage of morphologically abnormal sperm, 8 papers reported data on frequency of immotile or dead sperm and 6 papers reported data on sperm DNA/chromatin alterations.

Journal Pre-proofs

Table 4b. List of included papers on male fertility studies in *in vitro* exposed human sperm with main study characteristics.

N	*Reference*	**Population**		**Exposure**			**Outcome**	*Summary of paper results*
		Donor fertility status (number of donors)	*Mean donor age (y)*	*Frequency (MHz)/ Modulation (M, CW)*	*Level(s) W/kg*	*Duration(s) Hours per day*	*Semen quality*	
1	Agarwal et al., 2009	F, SF (23, 9)	NR	850/CW	0.01-0.4 W/m^2	1:00	**X**	Effect on viability and motility in sperm of healthy donors and no effect in sperm of subfertile donors. No effect on sperm DNA damage.
2	Avendaño et al., 2012	F (29)	34.1	2400/CW	0.00003	4:00	**X**	No temperature increase. No effect on sperm viability; decrease of sperm motility. Increase of sperm DNA damage.
3	De Iuliis et al., 2009	F (4)	24.1	1800/CW	0.4, 1, 2.8, 4.3, 10.1, 27.5	16:00	**X**	< 0.4°C temperature increase. Dose-dependent effects on sperm viability and motility. Dose-dependent effect on oxidative DNA damage and DNA fragmentation.
4	Falzone et al., 2011	F (12)	23	900/M	2	1:00	**X**	No temperature increase. Decrease of sperm head area. Decrease of number of sperm binding to oocytes.
5	Falzone et al., 2008	F (12)	23	900/M	2, 5.7	1:00	**X**	No temperature increase. No effect on motility.

Journal Pre-proofs

6	Falzone et al., 2010	F (12)	23	900/M	2, 5.7	1:00	**X**	No temperature increase. No effects on DNA fragmentation, apoptosis, oxidative stress.
7	Makler et al., 1980	F, SF (8, 5)	NR	27	NR	0:30		No information on power density or SAR values; exposure level expressed as 0.6 V/m. No temperature increase. Decrease of sperm motility for both fertile and subfertile donor samples.
8	Nakatani-Enomoto et al., 2016	F (25)	23.6	1950/M	2, 6	1:00	**X**	No temperature increase. No effect on sperm motility. No effect on oxidative DNA damage.
9	Veerachari and Vasan, 2012	F (20)	NR	900/M	0.01-0.4 W/m^2	1:00	**X**	Effects on sperm motility and viability. Effect on DNA fragmentation.
10	Wang et al., 2015	SF (97)	31.2	1950	3	3:00	**X**	Effects on sperm viability and motility; no effect on sperm morphology.

Gray cells mark outcomes for which each paper contributed data. When data were entered into a meta-analysis this is marked by X. Information on modulation is not reported when it was not unambiguously provided in the paper. Exposure level(s) are expressed in SAR unless this data was not unambiguously provided in the paper, in which case PD values are reported if provided.

Abbreviations: CW: continuous wave unmodulated signal; F: fertile donors; M: modulated signal; NR: not reported; SF: sub-fertile donors

Journal Pre-proofs

Table 3 lists the endpoints considered for each outcome, showing the number of papers and studies for each of them, the metrics in which the results were expressed, and the effect size measures used for the synthesis of results by meta-analysis.

3.4. Risk of bias in studies

Supplementary Files 2a and 2b show, for animal and human sperm *in vitro* studies, respectively, the heatmaps of the consensus scores assigned to each RoB question together with the overall level of concern. These files are organised in different sheets corresponding to specific endpoints and each line is relative to a specific study that might include multiple comparisons between sham and exposed groups. The justifications of the assigned scores are reported in Supplementary Files 3a and 3b. Irrespectively of the analysed endpoint, the main reasons for concern for animal studies were related to the lack of blinding during experiment performance, the poor exposure dosimetry and characterization and the low confidence in the outcome assessment, mostly due to the lack of blinding during the analysis of endpoints. For human sperm *in vitro* studies, the main reasons for concern were lack of blinding during experimental performance and/or during outcome assessment (except in those few cases in which automated methods of analysis were applied) and poor exposure characterization. In general, very few studies were classified at “low concern”.

3.5. Effects of the exposure

3.5.1. Results of individual studies

Results of individual studies are reported in Supplementary Files 4a and 4b, along with experimental design and exposure conditions applied in each of them.

3.5.2. Results of the syntheses

3.5.2.1. Effects on fertility

Few studies directly assessed male fertility by the number of males that did not get at least one female pregnant after mating (rate of infertile males), the number of nonpregnant females over total number of paired females, or the size of the litters sired by experimental males.

Rate of infertile males

The pooled effect size of 4 studies that evaluated fertility by the frequency of males that did not get at least 1 female pregnant after mating yielded an OR value of 1.38 (95% CI 0.32 to 5.94), showing no association of exposure with this effect (Figure 4). All these studies were rated at “low or some concern” RoB level. The average SAR tested in these studies was 0.38 W/kg (SD 0.45, 0.09-1.05 min-max). Due to the paucity of studies, no subgroup and dose-response analyses were done.

Number of nonpregnant females over paired females

Nineteen studies, rated at “low or some concern” RoB level, evaluated male fertility by the number of nonpregnant females over the number of paired females. The pooled OR value of these studies was 2.39 (95% CI 1.52 to 3.74), suggesting an RF-EMF exposure-associated decrease of male fertility (Figure 5). The average SAR tested in these studies was 23.87 W/kg (SD 21.19, 0.12-43.4 min-max). Three further studies rated at “high concern” RoB level yielded an OR value of 2.69 (95% CI 1.62 to 4.45).

Number of nonpregnant females over paired females: subgroup analysis

No significant difference was observed between the studies performed on mice and the studies performed on rats. Similarly, no significant difference was observed between the studies testing SAR between 0.1 and 5 W/kg and the studies testing SAR over 5 W/kg, although only the latter yielded a statistically significant pooled OR value. Differences in levels of animal temperature increase among the studies could not be used to explore sources of result heterogeneity because the group of studies in which less than 1°C temperature increase was measured included only 2 studies (Supplementary File 5).

Number of nonpregnant females over paired females: dose response analysis

The linear fitting of the dose response relationship showed a small but statistically significant increase of the OR of 0.03 per unit W/kg increase. The cubic spline model did not seem to fit the data better than the linear model (Supplementary File 6).

Litter size

The pooled SMD of 0.04 (95% CI -0.15 to 0.23) for the 16 studies at "low or some concern" RoB level showed that there was not a statistically significant decrease of litter size after RF-EMF exposure. The average SAR tested in these studies was 24.22 W/kg (SD 46.89, 0.12-141.4 min-max). Four further studies rated at "high concern" RoB level yielded an SMD value of 4.15 (95% CI -1.02 to 9.31) (Figure 6).

Litter size: subgroup analysis

Heterogeneity among individual study results could not be explained by subgroup analyses according to experimental animal species, SAR or animal temperature increase, since test of group differences never produced a significant result (Supplementary File 5).

Litter size: dose response analysis

The dose-response analysis confirmed a lack of association of RF-EMF exposure with decrease of the litter sired by the experimental males (Supplementary File 6).

Further studies on male reproductive performance not included in the meta-analysis

In the study by Houston et al. (2019), the fertility of experimental mice exposed for 35 days to 2.2 W/kg was assessed by *in vitro* fertilization of oocytes with their sperm and measurement of the percentage of fertilized oocytes and blastocyst development, without any evidence of RF-EMF associated effect.

3.5.2.1.1. Effects on fertility after EMP exposure

All results on EMP exposure studies are shown in Supplementary File 7.

The pooled OR for the incidence of nonpregnant females for 5 studies rated at "low or some concern" RoB level was 0.89 (95% CI 0.38 to 2.1) showing no impact of EMP exposure for this endpoint. No study was rated at "high concern" RoB level. All studies, from a single paper, were conducted in mice at the same SAR associated to a temperature increase lower than 1°C.

The pooled SMD for litter size for 5 studies rated at “low or some concern” RoB level was -0.38 (95% CI -0.91 to 0.15) showing no impact of EMP exposure for this endpoint. All studies, from a single paper, were conducted in mice at the same exposure level associated to a temperature increase lower than 1°C. The only study rated at “high concern” RoB level had an SMD of 0.5 (95% CI -0.91 to 1.91).

In addition to studies contributing to the meta-analysis, Cobb et al. (2000) reported a significant decrease of mated females and a non-significant decrease of fertile matings after prenatal EMP exposure of male rats to 0.045 W/kg without animal temperature increase.

3.5.2.2. *Effects on semen quality*

3.5.2.2.1. Experimental animal studies

Sperm count

Eighty-seven studies assessed RF-EMF effects by the number or concentration of sperm collected from animal epididymis. The individual study results were expressed by a variety of metrics; in some cases, information on dilution factors was missing and sometimes it was unclear if figures referred to 1 or 2 epididymides. Notwithstanding this study limitation, we decided to include all studies that clearly applied the same metrics to the comparator and the exposed groups. However, this choice produced a large variability of results and made it necessary to use SMD as the pooled effect size. Pooled SMD values of 0.74 (95% CI 0.51 to 0.98, 80 studies) and 0.54 (95% CI -0.06 to 1.15, 7 studies) were obtained for studies rated at “low or some concern” and “high concern”, respectively. The meta-analysis of the former group of studies showed a statistically significant decrease of sperm quantity in RF-EMF exposed animals (Figure 7). The average SAR tested in these studies was 12.2 W/kg (SD 16.03, 0.001-43.6 min-max).

Sperm count: subgroup analysis

Heterogeneity among individual study results could not be explained by subgroup analyses according to experimental animal species, SAR or animal temperature increase, since test of group differences never produced a significant result (Supplementary File 5).

Sperm count: dose response analysis

Neither the linear nor the cubic spline curves fit the individual study results satisfactorily with AIC values of 846 and 758, respectively (Supplementary File 6).

Sperm morphology

The meta-analysis of 65 studies measuring the percentage of morphologically abnormal sperm, which were rated at “low or some concern” RoB level, yielded a statistically significant MD of -0.94 (95% CI -1.28 to -0.59). Considering an average percentage of abnormal sperm in the comparator group of 7.7% and the extremely large variability of this parameter among individual studies, from 0 to 24.6%, a less than 1% increase of the spontaneous percentage seems a modest effect. The average SAR tested in these studies was 13.59 W/kg (SD 14.52, 0.001-43.6 min-max). Five further studies rated at “high concern” RoB level had a pooled MD of -5.62 (95% CI -8.75 to -2.50) (Figure 8).

Sperm morphology: subgroup analysis

The subgroup analyses by animal species, SAR and animal temperature increase showed statistically significant results for the test of group differences, showing that these independent variables could partly explain the variability among the individual study results (Supplementary File 5).

Sperm morphology: dose response analysis

Both the linear and the cubic spline curves suggested a dose-dependent increase of the percentage of morphologically abnormal sperm, but the high AIC values (693 and 535, respectively) indicated that both models fit the data poorly (Supplementary File 6).

Sperm vitality

Thirty-eight studies assessed semen quality by measuring sperm vitality as a percentage of either immotile or dead sperm. The pooled MD in the 32 studies rated at "low or some concern" RoB level was -10.83 (95% CI -15.20 to -6.47), showing a significant decrease of sperm vitality in RF-EMF exposed animals. This increase of immotile/dead sperm should be regarded in relation to a spontaneous incidence of about 32% averaged across largely variable individual studies from 8 to 53%. The average SAR tested in these studies was 1.5 W/kg (SD 1.87, 0.001-5.0 min-max). The MD value in the 6 "high concern" RoB studies was -18.74 (95% CI -29.33 to -8.16) (Figure 9).

Sperm vitality: subgroup analysis

Heterogeneity among individual study results could not be explained by subgroup analyses according to experimental animal species or SAR since in both cases test of group differences did not show a significant result. No subgroup analysis could be conducted by animal temperature increase because all studies in which temperature was measured showed an increase of less than 1°C (Supplementary File 5).

Sperm vitality: dose response analysis

The linear and the cubic spline curves showed significant coefficients suggesting a dose dependent effect, but both analyses fitted the data poorly, with AIC values of 441 and 389, respectively (Supplementary File 6).

Sperm DNA/chromatin alterations

Few studies assessed possible effects on semen quality by markers other than those routinely used in human andrological studies. These markers might have an impact on the sperm fertilising capacity but have not yet been standardized and could be considered a more indirect evidence of semen quality deterioration. They are essentially markers of sperm DNA/chromatin alterations. We conducted an overall meta-analysis of studies using these markers by SMD as the pooled effect size (Figure 10). Six studies rated at "low or some concern" RoB level yielded an SMD of -1.92 (95% CI -2.78 to -1.05), showing an adverse effect of RF-EMF exposure on these markers. The average SAR tested in these studies was 1.59 W/kg (SD 0.92, 0.15-2.2 min-max). One further study rated at "high concern" RoB level had an SMD of -1.45 (95% CI -2.72 to -0.18).

No subgroup analysis could be conducted because only 2 studies were carried out in a species different from mice, all studies tested a SAR comprised between 1 and 5 W/kg and no study reported animal temperature measurement (Supplementary File 5).

Due to the paucity of studies and the variety of markers no dose response relationship was explored.

Further studies on semen quality not included in the meta-analysis

In the studies described by Aitken et al. (2005), no evidence of RF-EMF effects was reported on sperm count, percentage of abnormal sperm and sperm vitality after a 900 MHz, 12 hours/day, 7 days, 0.09 W/kg exposure of adult mice. The study by Pardhiya et al. (2022) on sperm morphological abnormalities reported an adverse effect of a 2002 MHz, 2 hours/day, 48 days, 1.2 W/kg exposure of adult rats. Ren et al. (2002) reported an adverse effect of RF-EMF on sperm morphology but not on sperm number after exposure of adult mice to 2450 MHz, 100 W/m^2 for 30, 60 or 90 minutes/day for 6-18 days. No effect on sperm count was reported by Yan et al. (2022) after exposure of adult mice to 2000 MHz, 0.31 W/kg, 3 hours/day for 98 days. All these results could not be included in the meta-analyses because of insufficient data reporting.

3.5.2.2.2. Human sperm in vitro studies

Morphology

There were only 2 studies assessing human sperm morphology after *in vitro* RF-EMF exposure. One, at “some concern” RoB level, measured sperm head area showing a detrimental RF-EMF effect by an SMD of 6.04 (95% CI 4.48 to 7.60). Another one, rated at “high concern” RoB level because of poor confidence in outcome assessment, reported a non-significant increase of the percentage of morphologically abnormal sperm (SMD -0.14, 95% CI -0.42 to 0.14).

Vitality

Twenty-four studies measured sperm vitality as percentage of either immotile or dead sperm. The pooled MD in the 23 studies rated at “low or some concern” RoB level was -1.37 (95% CI -2.46 to -0.28), showing a significant decrease of sperm vitality in RF-EMF exposed samples (Figure 11). This increase of immotile/dead sperm should be regarded in relation to a spontaneous incidence of about 30% averaged across largely variable individual studies from about 5 to 55%. The study rated at “high concern” for RoB reported an MD of -9.08, 95% CI -13.84 to -4.32. No subgroup analysis could be performed because, for the fertility status of donors and the SAR tested, some of the groups included less than 3 studies (Supplementary File 5), and for temperature increase all studies in which the parameter was measured reported less than 1°C increase (data not shown).

Sperm vitality: dose response analysis

Neither the linear nor the cubic spline dose response curves fitted the data satisfactorily yielding high AIC values and not significant coefficients (Supplementary File 6).

Sperm DNA/chromatin alterations

Thirteen studies, all at “low or some concern” for RoB, measured sperm DNA/chromatin alterations showing a pooled SMD of -0.17 (95% CI -0.48 to 0.13) suggesting the absence of RF-EMF effect on this endpoint (Figure 12). No subgroup analysis could be performed because, for the fertility status of donors and the SAR tested, some of the groups included

less than 3 studies (Supplementary File 5), and for temperature increase all studies in which the endpoint was measured reported less than 1°C increase (data not shown).

Sperm DNA/chromatin alterations: dose response analysis

Both the linear and the cubic spline dose response curves seem to fit the data with AIC values of 53.97 and 50.09, respectively, and significant coefficients (Supplementary File 6).

In vitro human sperm studies not included in the meta-analysis

Makler et al. (1980) tested the effects of 30 minutes exposure to 27 MHz, 0.6 V/m electric field strength on 8 samples from fertile donors and 5 samples from subfertile donors, reporting a decrease of sperm motility in exposed samples, in spite of no temperature increase.

3.5.2.2.3. Effects on semen quality after EMP exposure

All results on EMP exposure studies are shown in Supplementary File 7.

Relative to sperm count, pooled SMD values of 0.23 (95% CI -0.09 to 0.56, 10 studies) and 0.02 (95% CI -0.75 to 0.8, 1 study) were obtained for studies rated at “low or some concern” and “high concern”, respectively, showing no detrimental impact of EMP exposure on this endpoint. All studies at “low or some concern” for RoB, published in 2 papers, were conducted in mice; when animal temperature was measured, an increase lower than 1°C was detected.

Relative to sperm morphology, the pooled MD value for 10 studies rated at “low or some concern” for RoB was -0.4 (95% CI -0.57 to -0.23), showing an increase of morphologically abnormal sperm in EMP exposed animals. The MD of 1 study rated at “high concern” was -0.53 (95% CI -3.2 to 2.14). All studies at “low or some concern” for RoB, published in 2 papers, were conducted in mice; when animal temperature was measured, an increase lower than 1°C was detected.

Relative to sperm vitality, pooled SMD values of -2.5 (95% CI -7.89 to 2.9, 2 studies) and -5 (95% CI -30.57 to 20.57, 1 study) were obtained for studies rated at “low or some concern” and “high concern”, respectively, showing no detrimental impact of EMP exposure on this endpoint. The studies at “low or some concern for RoB were conducted in mice under the same exposure conditions associated to an animal temperature increase of less than 1°C.

3.5.2.3. Reproductive organ toxicity

Several different biomarkers of toxic effects, commonly applied in experimental reproductive toxicity studies, were reported in the various papers. They included measurement of testis or epididymis weight, histological measurement of seminiferous tubule diameter, Johnsen’s score as a synthetic marker of histopathological alterations, testicular cell death or assessment of testicular sperm production.

Testis or epididymis weight

Sixty-eight studies reported data on testis or epididymis weight. The SMD value of the 55 studies rated at “low or some concern” RoB level was 0.29 (95% CI 0.10 to 0.47), showing

a small but significant decrease of weight in RF-EMF exposed animals. The average SAR tested in these studies was 3.62 W/kg (SD 3.3, 0.002-5.0 min-max). The SMD value for the 13 studies rated at “high concern” RoB level was 0.42 (95% CI -0.03 to 0.87) (Figure 13).

Testis or epididymis weight: subgroup analysis

Heterogeneity among individual study results could not be explained by subgroup analyses according to experimental animal species, SAR or animal temperature increase, since test of group differences never produced a significant result (Supplementary File 5).

Testis or epididymis weight: dose response analysis

The linear fitting of results showed a dose-dependent decrease of testis or epididymis weight, but the slope of the curve was not statistically significant. The cubic spline curve, with a similar AIC, suggested a change in the direction of the dose-effect relationship at about 5 W/kg with statistically significant slopes in the two portions of the curve (Supplementary File 6).

Testis histomorphometry

Forty-one studies reported data on testis histomorphometry. The pooled SMD value of the 24 studies rated at “low or some concern” RoB level was 0.90 (95% CI 0.32 to 1.49), showing an association between exposure to RF-EMF and a decrease of tubule diameter. The average SAR tested in these studies was 2.5 W/kg (SD 2.18, 0.002-5.0 min-max). Seventeen studies rated at “high concern” RoB level yielded an SMD value of 0.84 (95% CI 0.27 to 1.40) (Figure 14).

Testis histomorphometry: subgroup analysis

Heterogeneity among individual studies could not be explained by SAR level. No subgroup analysis could be performed by species or animal core temperature increase because some of the groups included less than 3 studies (Supplementary File 5).

Testis histomorphometry: dose response analysis

The linear fitting showed a dose-dependent decrease of testicular tubules diameter. The cubic spline curve did not improve the fitting.

Testis or epididymis histology

Twenty-four studies reported data on Johnsen’s score, as an indicator of testis histopathology. The pooled MD value of the 17 studies rated at “low or some concern” RoB level was 0.69 (95% CI 0.45 to 0.92), showing a small but statistically significant decrease of the Johnsen’s score reflecting an association between exposure to RF-EMF and testis histological alterations. The average SAR tested in these studies was 2.95 W/kg (SD 2.35, 0.002-5.0 min-max). Seven studies rated at “high concern” RoB level yielded an MD value of 0.59 (95% CI 0.17 to 1.01) (Figure 15).

Testis or epididymis histology: subgroup analysis

Heterogeneity among individual studies could be partially explained by differences in species but not by SAR. Of note, all the data in mice were obtained at a SAR higher than

5 W/kg and were produced in one laboratory. All but 1 study measured a temperature increase lower than 1°C thus subgrouping by temperature increase could not be exploited to explore sources of heterogeneity (Supplementary File 5).

Testis or epididymis histology: dose response analysis

The linear fitting of results showed a dose-dependent decrease of Johnsen's score. The cubic spline curve, with a similar AIC, suggested a change in the dose-effect relationship at about 2 W/kg, SAR at which the curve seems to reach a plateau (Supplementary File 6).

Testicular cell death

Thirty-one studies reported data on the level of dead or apoptotic testicular cells. The pooled SMD value of the 23 studies rated at "low or some concern" RoB level was -1.18 (95% CI -1.82 to -0.54), showing a statistically significant increase of dead or apoptotic testicular cells in RF-EMF exposed animals. The average SAR tested in these studies was 6.51W/kg (SD 5.87, 0.007-18.0 min-max). Eight studies rated at "high concern" RoB level yielded an SMD value of -5.33 (95% CI -7.62 to -3.04) (Figure 16).

Testicular cell death: subgroup analysis

Heterogeneity among individual studies could be partially explained by differences in SAR levels. Test of group differences did not show differences among species. All studies reported animal temperature increases lower than 1°C (Supplementary File 5).

Testicular cell death: dose response analysis

The slope of the linear fitting of results was not statistically significant. The cubic spline curve, with an AIC of 198, showed a complex dose response relationship suggesting a change in the direction of the dose-effect relationship at about 5 W/kg (Supplementary File 6).

Testicular sperm production

Forty studies reported data on testicular sperm production. The pooled SMD value of the 36 studies rated at "low or some concern" RoB level was 0.87 (95% CI 0.51 to 1.22), showing a statistically significant decrease of sperm production in RF-EMF exposed animals. The average SAR tested in these studies was 5.79 W/kg (SD 9.05, 0.03-43.65 min-max). Four studies rated at "high concern" RoB level yielded an SMD value of -0.26 (95% CI -1.19 to 0.67) (Figure 17).

Testicular sperm production: subgroup analysis

Heterogeneity among individual studies could be partially explained by differences in animal species and in SAR. All studies in which animal temperature was measured reported an increase above 1°C (Supplementary File 5).

Testicular sperm production: dose response analysis

The very high AIC values of both linear and cubic spline fittings, associated to not statistically significant coefficients, did not shed light on the shape of the dose effect relationship (Supplementary File 6).

Further studies on toxic effects on reproductive organs not included in the meta-analysis

Several studies reported no effects on testis or epididymis weight under a variety of exposure conditions. The study by Aitken et al. (2005) exposed adult mice to 900 MHz, 0.09 W/kg, 12 hours/day for seven days. Fahim et al. (1975) exposed adult rats to 2450 MHz, at an exposure level inducing a temperature increase in the testis up to 65°C for a few minutes. Lerchl et al. (2008) exposed adult hamsters to 0.08 W/kg of different radiofrequencies for 24 hours/day for 60 days. Ren et al. (2002) exposed adult mice to 2450 MHz, 100 W/m^2, for variable times, between 0.5 and 1.5 hours/day for 6-18 days. Shirai et al. (2017) exposed rats in the pre-natal and pre-puberal periods to multiple frequencies, 0.08 or 0.402 W/kg, 20 hours/day for 58 days. Takahashi et al. (2010) exposed rats in the pre-natal and pre-puberal periods to 2140 MHz, 0.051 or 0.119 W/kg, 20 hours/day for 35 days. A detrimental effect on testis weight was suggested by Prausnitz and Susskind (1962) after exposure of mice to 9300 MHz, high power density leading to a 3.3°C body temperature increase for 5 minutes/day for 295 days. In the same study some histopathological effects on testis were also reported. Effects on testis histopathology were reported by Ozlem Nisbet et al. (2012) after exposure of rats from adolescence to adulthood to 900 or 1800 MHz, SAR equal to or lower than 0.002 W/kg, 2 hours/day for 90 days. Lee et al. (2005) reported no effects on testis histopathology and number of testicular dead cells after RF-EMF exposure of mice to various radiofrequencies, 0.4 W/kg, 1.5 hours/day for 20, 40 or 50 days. No effect on testicular cell apoptosis was observed by Yan et al. (2022) after exposure of mice to 2000 MHz, 0.31 W/kg, 3 hours/day for 28 days. Finally, Saygin et al. (2015) reported an increase of testicular cell apoptosis after exposure of rats to 2450 MHz, 3.21 W/kg, 1 hour/day for 28 days. All these results could not be included in the meta-analyses because of insufficient data reporting.

3.5.2.3.1. *Effects on reproductive organ toxicity after EMP exposure*

All results on EMP exposure studies are shown in Supplementary File 7.

Relative to testis or epididymis weight, the pooled SMD value of 10 studies rated at “low or some concern” RoB level was 0.16 (95% CI -0.14 to 0.45), showing no impact of EMP exposure. There was no study rated at “high concern” for RoB. All studies, published in 2 papers, were conducted in mice; when animal temperature was measured, an increase less than 1°C was detected.

Relative to testis histomorphometry, the pooled MD value of 10 studies rated at “low or some concern” RoB level was 0.66 (95% CI 0.33 to 0.99), showing a small reduction of the seminiferous tubule diameter in the exposed animals. There was no study rated at “high concern” for RoB. All studies, published in 2 papers, were conducted in mice; when animal temperature was measured, an increase less than 1°C was detected.

Relative to testicular cell death, the pooled SMD value of 5 studies rated at “low or some concern” RoB level was -0.05 (95% CI -0.9 to 0.8), showing no impact of EMP exposure. There was no study rated at “high concern” for RoB. All studies, published in a single paper, were conducted in mice under the same exposure conditions.

In addition to studies contributing to the meta-analyses, Luo et al. (2013) reported a significant increase of testicular apoptotic cells after exposure of adult mice to 200 kV/m EMP at the highest number of pulses tested.

3.5.2.4. *Hormonal effects*

Testosterone level

Forty studies reported data on testosterone level in either serum or testis. The pooled SMD value of the 29 studies rated at “low or some concern” RoB level was 0.87 (95% CI 0.43 to 1.30), showing a statistically significant decrease of testosterone level in RF-EMF exposed animals. The average SAR tested in these studies was 0.9 W/kg (SD 1.31, 0.001-4.0 min-max). Eleven studies rated at “high concern” RoB level yielded an SMD value of 0.50 (95% CI -0.22 to 1.23 (Figure 18).

Testosterone level: subgroup analysis

Differences in animal species and SAR could not explain variations in individual study results. No subgroup analysis could be conducted for animal core temperature increase because there were not enough studies in the subgroups (Supplementary File 5).

Testosterone level: dose response analysis

The linear curve, with an AIC value of 199, showed a significant decrease of testosterone level with an increase of the exposure level with a coefficient of 1.45. Statistically significant coefficients of the cubic spline fitting seemed to describe the shape of the dose effect relationship with better fitting (AIC=149) (Supplementary File 6), suggesting a non-linear relationship with a decrease of testosterone at low exposure levels, followed by an increase at higher SAR, especially above 8 W/kg.

Further studies on testosterone level not included in the meta-analysis

Shahin et al. (2014) reported a decrease of serum testosterone level after exposure of mice to 2450 MHz, 0.018 W/kg, 2 hours/day for 30 days.

3.5.2.4.1. *Hormonal effects after EMP exposure*

All results on EMP exposure studies are shown in Supplementary File 7.
Relative to testosterone level, the pooled SMD value of 6 studies rated at “low or some concern” RoB level was -1.74 (95% CI -3.92 to 0.43), showing no impact of EMP exposure. There was no study rated at “high concern” for RoB. All studies, published in 2 papers, were conducted in mice; when animal temperature was measured, an increase less than 1°C was detected.
In addition to the studies contributing to the meta-analysis, no effect on serum testosterone level was described by Dong et al. (2021) after EMP exposure of mice to 50-300 W/m^2 for 30 minutes, and erratic variations of serum testosterone levels were described by Zeng et al. (2011) after EMP exposure of adult rats to 0.03 W/kg for 43-133 minutes.

3.6. Reporting Bias assessment

Based on the funnel plots and the corresponding Egger’s tests, the study datasets indicated a reporting bias for all endpoints, with the exception of experimental animal studies on rate of pregnancies, litter size and testicular sperm production, and of *in vitro* experimental studies on sperm DNA/chromatin alterations (Supplementary File 8). Few studies were retrieved on EMP exposure and for many endpoints there were less than 10 studies so the Egger’s test was not done to support visual inspection of funnel plots. When it was conducted, evidence of reporting bias assessment was shown for sperm morphology and testis or epididymis weight, while such evidence was not shown for sperm count and testis histomorphometry (Supplementary File 7).

3.7. Results of additional unplanned analyses

The comparison between the pooled effect size of the 29 studies on sperm morphology considered less reliable for outcome assessment and the pooled effect size of the 36 studies on sperm morphology considered more reliable for outcome assessment showed that the former was about twice as high as the latter (-4.7 vs -2.6). This difference could even be greater, considering that the average exposure level in the former was about 10 times lower than that in the latter (2.9 W/kg vs 22.2 W/kg).

The comparison between the pooled effect size of the 14 studies on testicular cell death considered less reliable for outcome assessment and the pooled effect size of the 9 studies on testicular cell death considered more reliable for outcome assessment showed that the average SMD for the former was -4.38 and that of the latter was 0.2.

Similarly, in the case of studies on testicular sperm production, with 21 and 15 studies of worse and better quality, respectively, the average SMD for the former was 1.57 and that of the latter was 0.56.

3.8. Certainty assessment

Findings have been evaluated according to a GRADE approach as shown in Table 5 and, for EMP exposure, in Table 6.

Effects on fertility

The OR for infertile males, spanning from 0.32 to 5.94, is consistent with null. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, and for imprecision by 2 levels, due to the width of the 95% CI of the pooled effect size crossing the null hypothesis and the limited number of studies.

The SMD for litter size is similarly consistent with no effect, spanning from -0.15 to 0.23. We downrated the certainty to moderate, for RoB by one level, all but one study being at “some concern”. We did not downgrade for publication bias because of the pooled effect size consistent with null, the inspection of the funnel plot and the borderline statistical significance of the Egger’s test. Despite the pooled effect sizes of studies in different species were not significantly different, the result was not upgraded for consistency across species because the p value was borderline significant.

The OR for nonpregnant females, spanning from 1.52 to 3.74, is consistent with a detrimental effect of exposure. We downrated the certainty for RoB by one level, all but one study being at “some concern”, and for inconsistency by one level due to the variability among individual studies (I^2=60%) not explained by subgroup analysis. We upgraded the certainty to moderate for consistency among species.

Effects on semen quality

Effects on semen quality were evaluated by both experimental animal studies and studies on human sperm exposed *in vitro*. In animal experiments, the SMD for sperm count, spanning from 0.51 to 0.98, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty for RoB by one level, most studies being at “some concern”, for inconsistency by one level due the variability among individual studies (I^2=71%) not explained by subgroup analysis, and for publication bias by one level after inspection of the

funnel plot and the significance of Egger's test. We upgraded the certainty to low for consistency among species.

The MD for sperm morphology, spanning from -1.28 to -0.59, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty to very low, for RoB by one level, most studies being at “some concern”, for inconsistency by one level due to the variability among individual studies (I^2=90%) partly explained by subgroup analysis, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test.

The MD for sperm vitality, spanning from -15.2 to -6.47, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty for RoB by one level, most studies being at “some concern”, for inconsistency by two levels due to the variability among individual studies (I^2=96%) not explained by subgroup analysis, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test. We upgraded the certainty for consistency among species. As a whole, the certainty was very low.

The SMD for sperm DNA/chromatin alterations, spanning from -2.78 to -1.05, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, for inconsistency by one level due to the variability among individual studies (I^2=66%) not explained by subgroup analysis, and for indirectness by one level, this endpoint being weakly linked with male infertility by adverse outcome pathways.

In studies on human sperm exposed *in vitro*, only one study not rated at “high concern” RoB level reported data on sperm morphology. The MD for sperm vitality, spanning from -2.46 to -0.28, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty to very low, for RoB by one level, most studies being at “some concern”, for inconsistency by one level due to the variability among individual studies (I^2=65%) not explained by subgroup analysis, for indirectness by one level, because we considered all *in vitro* studies to assess indirectly the impact of RF-EMF exposure on men, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test.

The SMD for sperm DNA/chromatin alterations, spanning from -0.48 to 0.13, is consistent with null. We downrated the certainty to very low, for RoB by one level, most studies being at “some concern”, for inconsistency by one level, due to the variability among individual studies (I^2=53%) not explained by subgroup analysis, for indirectness by two levels, one because the endpoint is weakly linked with male infertility by adverse outcome pathways and one because we considered all *in vitro* studies to assess indirectly the impact of RF-EMF exposure on men.

Reproductive organ toxicity

The SMD for testis or epididymis weight, spanning from 0.10 to 0.47, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty for RoB by one level, most studies being at “some concern”, for inconsistency by one level, due to the variability among individual studies (I^2=44%) not explained by subgroup analysis, for indirectness by one level, because testis weight decrease in experimental rodents is considered weakly predictive of human male infertility, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test. We upgraded the certainty for consistency among species. As a whole the certainty was very low.

The SMD for testis histomorphometry, namely seminiferous tubule diameter, spanning from 0.32 to 1.49, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty to very low, for RoB by one level, most studies being at “some concern”, for inconsistency by two levels, due to the variability among individual studies (I^2=84%) not explained by subgroup analysis, for indirectness by one level, because decrease of seminiferous tubule diameter in experimental rodents is considered weakly predictive of human male infertility, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test.

The MD for testis histopathology measured by Johnsen’s score, spanning from 0.45 to 0.92, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty to very low, for RoB by one level, most studies being at “some concern”, for inconsistency by two levels, due to the variability among individual studies (I^2=93%) not explained by subgroup analysis, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test.

The SMD for testicular cell death, spanning from -1.82 to -0.54, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty for RoB by one level, most studies being at “some concern”, for inconsistency by one level, due to the variability among individual studies (I^2=86%) only partly explained by subgroup analysis, for indirectness by one level, because testicular cell death in experimental rodents is considered weakly predictive of human male infertility, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test. We upgraded the certainty for consistency among species. As a whole, the certainty was very low.

The SMD for testicular sperm production, spanning from 0.51 to 1.22, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty to very low, for RoB by one level, most studies being at “some concern”, for inconsistency by one level, due to the variability among individual studies (I^2=69%) only partially explained by subgroup analysis, and for indirectness by one level, because testicular sperm production in experimental rodents is considered weakly predictive of human male infertility.

Hormonal effects

The SMD for testosterone level, spanning from 0.43 to 1.30, is consistent with a detrimental effect of RF-EMF exposure. We downrated the certainty for RoB by one level, most studies being at “some concern”, for inconsistency by two levels, due to the variability among individual studies (I^2=86%) not explained by subgroup analysis, and for publication bias by one level after inspection of the funnel plot and the significance of Egger's test. We upgraded the certainty for consistency among species. As a whole, the certainty was very low.

Journal Pre-proofs

Table 5. GRADE Evidence Profile

Certainty assessment								Summary of findings					
								N° of participants		Effect		Certainty	Importance**
N° of studies	Design	RoB	Inconsistency	Indirectness	Imprecision	Publication bias*	Consistency across species	Exposure	Comparator	Relative (95% CI)	Absolute (95% CI)		
Decrease of fertility													
Rate of infertile males													
4	(a)	-1	0	0	-2	NA	NA	89	55	OR 1.38 (0.32 to 5.94)		Very low	4
Nonpregnant females over paired females													
19	(a)	-1	-1	0	0	0	+1	638	549	OR 2.39 (1.52 to 3.74)		Moderate	7
Litter size (an SMD positive value indicates a detrimental RF-EMF effect)													
16	(a)	-1	0	0	0	0	0	268	195	SMD 0.04 (-0.15 to 0.23)		Moderate	8

Journal Pre-proofs

In vitro fertilization rate: No meta-analysis was done because the database included only one paper												
Effects on semen quality-experimental animal studies												
Sperm count (an SMD positive value indicates a detrimental RF-EMF effect)												
80	(a)	-1	-1	0	0	-1	+1	752	569	SMD 0.74 (0.51 to 0.98)	Low	8
Sperm morphology (an MD negative value indicates a detrimental RF-EMF effect)												
65	(a)	-1	-1	0	0	-1	0	567	436	MD -0.94 (-1.28 to -0.59)	Very Low	7
Sperm vitality (an MD negative value indicates a detrimental RF-EMF effect)												
32	(a)	-1	-2	0	0	-1	+1	334	265	MD -10.83 (-15.2 to -6.47)	Very Low	8
Sperm DNA/chromatin alterations (an SMD negative value indicates a detrimental RF-EMF effect)												
6	(a)	-1	-1	-1	0	0	0	56	55	SMD -1.92 (-2.78 to -1.05)	Very Low	4

Journal Pre-proofs

Effects on semen quality-studies on human sperm in vitro												
Sperm morphology (an MD negative value indicates a detrimental RF-EMF effect): No meta-analysis was done because the database included only one study that was not at "high concern" for RoB												
Sperm vitality (an MD negative value indicates a detrimental RF-EMF effect)												
23	(b)	-1	-1	-1	0	-1	NA	455	455	MD -1.37 (-2.46 to -0.28)	Very Low	8
Sperm DNA/chromatin alterations (an SMD negative value indicates a detrimental RF-EMF effect)												
13	(b)	-1	-1	-2	0	0	NA	215	195	SMD -0.17 (-0.48 to 0.13)	Very Low	4
Reproductive organ toxicity												
Testis-epididymis weight (an SMD positive value indicates a detrimental RF-EMF effect)												
55	(a)	-1	-1	-1	0	-1	+1	725	503	SMD 0.29 (0.10 to 0.47)	Very Low	6
Testis histomorphometry (an SMD positive value indicates a detrimental RF-EMF effect)												
24	(a)	-1	-2	-1	0	-1	0	173	162	SMD 0.90	Very Low	2

Journal Pre-proofs

										(0.32 to 1.49)		
Testis or epididymis histology (an MD positive value indicates a detrimental RF-EMF effect)												
17	(a)	-1	-2	0	0	-1	0	125	108	MD 0.69 (0.45 to 0.92)	Very Low	5
Testicular cell death (an SMD negative value indicates a detrimental RF-EMF effect)												
23	(a)	-1	-1	-1	0	-1	+1	285	168	SMD -1.18 (-1.82 to -0.54)	Very Low	3
Testicular sperm production (an SMD positive value indicates a detrimental RF-EMF effect)												
36	(a)	-1	-1	-1	0	0	0	364	243	SMD 0.87 (0.51 to 1.22)	Very Low	4
Hormonal effects												
Testosterone level (an SMD positive value indicates a detrimental RF-EMF effect)												
29	(a)	-1	-2	0	0	-1	+1	462	321	SMD 0.87 (0.43 to 1.30)	Very low	6

(a) Controlled experimental animal studies

Journal Pre-proofs

(b) Controlled experimental *in vitro* studies

* All studies considered, irrespective of their RoB rating

** The importance of each endpoint in relation to human male infertility was rated on a scale 1-10 from the least to the most important

NA: Not applicable

Journal Pre-proofs

Effects after EMP exposure

All studies after EMP exposure were conducted in experimental animals.

Effects on fertility

The OR for nonpregnant females, spanning from 0.38 to 2.10, is consistent with null. We downrated the certainty to low, for RoB by one level, all studies being at “some concern”, and for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes do not reach the Optimal Information Size.

The SMD for litter size, spanning from -0.91 to 0.15, is consistent with null. We downrated the certainty to low, for RoB by one level, all studies being at “some concern”, and for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes do not reach the Optimal Information Size.

Effects on semen quality

The SMD for sperm count, spanning from -0.09 to 0.56, is consistent with null. We downrated the certainty to low, for RoB by one level, all studies being at “some concern”, and for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes do not reach the Optimal Information Size.

The MD for sperm morphology, spanning from -0.57 to -0.23, is consistent with a detrimental effect of EMP exposure. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, for inconsistency by one level, due to the variability among individual studies (I^2=83%) not explained by subgroup analysis, for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes do not reach the Optimal Information Size, and for publication bias after inspection of the funnel plot and the significance of Egger’s test.

The MD for sperm vitality, spanning from -7.89 to 2.9, is consistent with null. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, for inconsistency by one level, due to the variability among individual studies (I^2=95%) not explained by subgroup analysis, and for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes do not reach the Optimal Information Size.

Reproductive organ toxicity

The SMD for testis or epididymis weight, spanning from -0.14 to 0.45, is consistent with null. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, for indirectness, because testis weight decrease in experimental rodents is considered weakly predictive of human male infertility, for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes barely reach the Optimal Information Size, and for publication bias after inspection of the funnel plot and the significance of Egger’s test.

The SMD for testis histomorphometry, spanning from 0.33 to 0.99, is consistent with a detrimental effect of EMP exposure. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, for indirectness, because decrease of

seminiferous tubule diameter in experimental rodents is considered weakly predictive of human male infertility, and for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes barely reach the Optimal Information Size.

The SMD for testicular cell death, spanning from -0.9 to 0.8, is consistent with null. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, for indirectness, because testicular cell death in experimental rodents is considered weakly predictive of human male infertility, and for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes do not reach the Optimal Information Size.

Hormonal effects

The SMD for testosterone level, spanning from -3.92 to 0.43, is consistent with null. We downrated the certainty to very low, for RoB by one level, all studies being at “some concern”, for inconsistency by one level, due to the variability among individual studies (I^2=95%) not explained by subgroup analysis, and for imprecision by one level, because the 95% CI crosses the null hypothesis and the sample sizes do not reach the Optimal Information Size.

Journal Pre-proofs

Table 6. GRADE Evidence Profile – studies on EMP exposure

Certainty assessment								Summary of findings				
								N° of participants		Effect		Certainty Importance**
N° of studies	Design	RoB	Inconsistency	Indirectness	Imprecision	Publication bias*	Consistency across species	Exposure	Comparator	Relative (95% CI)	Absolute (95% CI)	
Decrease of fertility												
Nonpregnant females over paired females												
5	(a)	-1	0	0	-1	NA	NA	25	25	OR 0.89 (0.38 to 2.10)		Low 7
Litter size (an SMD positive value indicates a detrimental RF-EMF effect)												
5	(a)	-1	0	0	-1	NA	NA	50	50	SMD -0.38 (-0.91 to 0.15)		Low 8
Effects on semen quality-experimental animal studies												
Sperm count (an SMD positive value indicates a detrimental RF-EMF effect)												

Journal Pre-proofs

10	(a)	-1	0	0	-1	0	NA	75	75	SMD 0.23 (-0.09 to 0.56)	Low	8
Sperm morphology (an MD negative value indicates a detrimental RF-EMF effect)												
10	(a)	-1	-1	0	-1	-1	NA	75	75	MD -0.4 (-0.57 to -0.23)	Very Low	7
Sperm vitality (an MD negative value indicates a detrimental RF-EMF effect)												
2	(a)	-1	-1	0	-1	NA	NA	60	20	MD -2.5 (-7.89 to 2.9)	Very Low	8
Reproductive organ toxicity												
Testis-epididymis weight (an SMD positive value indicates a detrimental RF-EMF effect)												
10	(a)	-1	0	-1	-1	-1	NA	90	90	SMD 0.16 (-0.14 to 0.45)	Very Low	6
Testis histomorphometry (an SMD positive value indicates a detrimental RF-EMF effect)												
10	(a)	-1	0	-1	-1	0	NA	75	75	SMD 0.66 (0.33 to 0.99)	Very Low	2

Journal Pre-proofs

Testicular cell death (an SMD negative value indicates a detrimental RF-EMF effect)												
5	(a)	-1	0	-1	-1	NA	NA	25	25	SMD -0.05 (-0.9 to 0.8)	Very Low	3
Hormonal effects												
Testosterone level (an SMD positive value indicates a detrimental RF-EMF effect)												
6	(a)	-1	-1	0	-1	NA	NA	158	56	SMD -1.74 (-3.92 to 0.43)	Very Low	6

(a) Controlled experimental animal studies

* All studies considered, irrespective of their RoB rating

** The importance of each endpoint in relation to human male infertility was rated on a scale 1-10 from the least to the most important

NA: Not applicable

Journal Pre-proofs

4. Discussion

4.1. Summary of the evidence and interpretation of the results

From experimental animal studies there is moderate certainty of evidence that RF-EMF exposure reduces rate of pregnancy, moderate certainty of evidence that exposure does not reduce litter size, and low certainty of evidence that exposure lowers sperm count. All other results of animal studies and all results on human sperm exposed *in vitro* have very low certainty. We retrieved few independent studies reporting male reproductive effects after experimental animal exposure to EMP. For this source of exposure, results on pregnancy rate, litter size and sperm count, all consistent with null, have a low certainty. All other results have a very low certainty.

It can be asked whether the results of our meta-analyses are consistent with the hypothesis that higher exposure levels, especially those inducing an hyperthermic effect, are more biologically effective than lower exposure levels. The result on the decrease of pregnancy rate is consistent with this hypothesis, as shown by the observation that the pooled effect size is statistically significant only in the subgroup of studies exposed to SAR equal to or higher than 5 W/kg and the statistically significant slope of the linear dose-response relationship. On the other hand, the results on sperm count do not show an increase of the detrimental effect with increasing SAR and all the models of dose-response relationship tested fit the data poorly. Also for other endpoints (the results of which were rated at very low certainty), a direct relationship between the effect and the exposure level is not evident by the subgroup and dose-response analyses and, in some cases, even the possibility of an inverse relationship is suggested by the data. However, this suggestion is not sustained by a solid adverse outcome pathway, and, in some cases, it is based only on few independent studies. We tested if other variables unequally distributed among the subgroups could have a role in increasing the heterogeneity of the observed results and could confound any underlying dose-effect relationship. Indeed, we showed that the absence of blinding during outcome assessment could strongly influence the results for those endpoints that were not measured by automated methods, thus supporting this hypothesis.

Journal Pre-proofs

4.2. *Limitations in the evidence*

Of all the papers included in the database of animal studies after the title/abstract evaluation, about 60% had to be excluded for different reasons, with poor exposure characterization accounting for about 45% of them. The experimental designs of the included studies were very variable for stage, duration and level of exposure and this probably contributed to the inconsistency of results and the limited certainty attributed to the body of evidence. In addition, several different endpoints were investigated in the studies to assess male fertility, which reduced the study database in each specific meta-analysis. Moreover, few studies directly assessed experimental male reproductive performance (this would have required a very large number of animals) and the large majority used surrogate markers for it. As shown by the RoB assessment, in most studies the endpoint was not measured according to rigorous principles and recommended guidelines, including blinding of experimental conditions, which is especially critical for semen quality, histomorphometric and histopathological parameters when not measured by automated methods. Very few papers reported data of experiments aimed at assessing the shape of dose response relationships or comparing low, non-thermal vs high exposure levels, under the same experimental conditions. Few papers reported data on a comparator group exposed to direct heating for specifically assessing the possibility of a non-thermal mechanism of action (Lebovitz et al., 1987b, Saunders et al., 1981a,b), but since temperature increase was not strictly comparable to that induced by RF-EMF exposure, we did not use the data as further comparator additional to the sham controls. For almost all the endpoints considered, evidence of reporting publication bias was shown by funnel plots and Egger's tests. Very few studies were retrieved on human sperm exposed *in vitro.* Furthermore, these studies are of limited relevance for predicting the impact of RF-EMF exposure on human male fertility because they expose only the very late stage of germ cell development, disregarding the possible impact of exposure on earlier spermatogenesis stages and mechanisms of regulation. While the SR looked for evidence across the 0.1 MHz – 300 GHz frequency range, 97% of studies were conducted in the 100 MHz – 10 GHz frequency range and 81% clustered around the interval 900-2450 MHz, the frequencies applied in telecommunications. Few studies were retrieved on exposure to EMP, a limitation worsened by the lack of independent replication, since often they derived from only few laboratories.

4.3. Limitations in the review process

Relative to animal studies, we could not make a decision about the inclusion of 31 papers out of the 323 selected by title/abstract examination either because we could not retrieve 20 of them in spite of our extensive searching of electronic databases and attempts to contact the authors or were unable to translate 11 of them. Relative to human sperm *in vitro* studies, we were unable to translate 1 paper out of the 44 included after title/abstract selection.

For inclusion in the meta-analysis of studies in which a single comparator was matched to different exposed experimental groups, we averaged the exposure conditions and responses, renouncing to the independency of results among the exposed groups.

For most endpoints, SMD, instead of MD, was used as the effect size, due to the variety of metrics by which the endpoints were measured in the different papers and our aims of synthesis and inclusiveness.

We acknowledge limitations in our subgroup analyses. In particular, the choice of conducting subgroup analysis to investigate sources of heterogeneity among the studies even when the subgroups were small (N=3 studies) induces the risk of false negative outcomes. However,

for the interpretation of results we relied upon the statistical significance of the between-group difference, which takes into account the group size. In addition, we acknowledge the limitation of subgrouping the studies showing an increase of animal temperature below or above 1°C, since it does not take into account the magnitude, timing or duration of temperature increase. Furthermore, in most cases the authors measured rectal body temperature rather than testicular temperature, thus the distinction between studies above and below a 1°C temperature increase may not be accurate for the target tissue (depending on the exposure system design). Indeed, studies in experimental mice have shown that similar testicular damage can be induced by 12 hours exposure to 36°C ambient temperature as well to 30 minutes exposure to over 40°C localised scrotal heating (Paul et al. 2008).

Similarly, the choice of subgrouping the studies by exposure level in 3 groups only, due to the need to include a reasonable number of studies in each group, may have blurred the contribution of studies testing very high exposure levels.

Regarding the assessment of publication bias, we acknowledge the limits of the Egger's test and the possibility of false positive results when drawing funnel plots with large SMD effect sizes (Zwetsloot et al., 2017). However, considering that no better alternatives exist to the Egger's test, we think that our approach is still an acceptable approximation to the assessment of this important source of bias.

4.4. Implications for policy and research

In conclusion, our systematic review and meta-analyses indicate a possible detrimental effect of RF-EMF exposure on pregnancy rate and sperm count in experimental mammals, whereas the meta-analysis of data on litter size was consistent with null.

Although sperm count is not a functional indicator of male fertility, it is a well-standardised analysis routinely applied in clinical andrology. RF-EMF emitting devices are widely applied and epidemiological surveys seem to indicate that, in Western countries, male fertility potential is declining (Auger et al., 2022, Boulicault et al., 2022, Levine et al., 2017). For these reasons the results of our meta-analyses should not be overlooked at a policy level.

It was beyond the scope of our systematic review to determine the shape of the dose-response relationship or to identify a minimum effective exposure level. For these reasons, we cannot provide suggestions to confirm or reconsider current human exposure limits. Nevertheless, it is of note that most studies on male fertility, semen quality and reproductive organ toxicity investigated exposure levels which were rather high with respect to those relevant for human populations: 75-80% tested exposure levels above 0.4 W/kg (ICNIRP basic restriction for workers) and 46-53% tested exposure levels above 4 W/kg (ICNIRP health effect level) (ICNIRP 2020). Thus, it is not known the extent to which the conclusions of the SR meta-analysis can be applied to human exposure levels. Similarly, it is unknown how much our conclusion can be extrapolated to frequencies below 100 MHz and above 10000 MHz, for which only very few studies were retrieved.

During the systematic review, we identified several methodological limitations in the studies that should be overcome to improve the quality of future research. In particular, blinding during experiment performance and outcome assessment should always be applied to minimize bias, an adequate number of cytological or histological preparations should be analysed, automated methods of analysis should be applied whenever possible, a more standardized and complete reporting of technical methods and results should be adopted.

Many studies had to be excluded from the systematic review because of insufficient exposure characterization and a large proportion of included studies were rated at either 'some' or 'high concern' for RoB for similar reasons. We would recommend that future studies bear the reasons for exclusion or RoB concerns in mind in study design and implementation. There are several papers in the research literature with recommendations on how exposure characterisation concerns can be mitigated, for example Kuster and Schonborn (2000). Finally, studies investigating not just a single level but several exposure levels, spanning from low levels comparable to human exposure to higher levels where mild hyperthermic effects could be expected, should be conducted under the same experimental conditions and target tissue temperature monitoring should be employed.

As a final suggestion for future research, we consider it a priority to obtain a scientifically solid database of possible RF-EMF effects on the best predictive surrogate markers of male infertility in experimental rodents. Based on the results of this research, the possibility of testing directly the RF-EMF impact on male reproductive performance could be considered. In view of the limitations of the approach applying *in vitro* exposure of human sperm, we do not recommend further studies of this kind. Conversely, we suggest exploiting semen quality analysis in human biomonitoring investigations of RF-EMF exposed populations.

4.5. Registration and Protocol

4.5.1. Protocol registration

Protocols for the systematic reviews of animal studies and of human sperm *in vitro* studies were published in (Pacchierotti et al., 2021).The former was also registered in PROSPERO (CRD42021227729 https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=227729) and the latter in Open Science Framework (OSF Registration DOI https://doi.org/10.17605/OSF.IO/7MUS3).

4.5.2. Deviations from the protocol

In the course of the systematic review, a few amendments to the protocols were made.

1. Minor changes consisted of a different organization and a slight rewording of outcomes and endpoints as reported in Table 2, in which specific reasons for changes are also shown.
2. To base the assessment of possible RF-EMF exposure impact on the most solid set of data, we excluded studies at "high concern" for RoB from the summary of findings assessed for the certainty of evidence by the GRADE approach, even if this had not been explicitly stated in the protocol.
3. In relation to exposure eligibility criteria, the protocol specified that studies in which exposure level from mobile phones or other RF-EMF generating devices was not measured or estimated by reliable methods, but simply inferred from assumed exposure conditions from the RF-generating device type, were to be assessed as a separate group. We preferred to assess this group together with the other studies and rate the confidence in the exposure characterization by the RoB assessment, because it was difficult to set boundaries in a continuum of exposure dosimetry reporting. In addition, we specified the exclusion of studies on exposure to ultrasound.
4. In relation to study design eligibility criteria, we excluded studies of exposure of both males and females of a mating pair because it made impossible sorting out specific effects on male fertility.

5. Another exclusion criterium unforeseen at protocol stage was paper retraction.
6. For binary outcomes, we used Odds Ratio instead of Relative Risk as the effect size measure because it was more easily tractable by the applied data analysis software.
7. For parameters of statistical heterogeneity of results, we calculated I^2 and τ^2 but not prediction intervals because we deemed that the former were sufficient indicators of heterogeneity.
8. Information regarding conflict-of-interest declarations and funding sources were not analysed since, in the vast majority of papers, public funding and absence of conflict of interest were declared.
9. Among the factors envisaged in the Protocol, we limited our heterogeneity analyses for *in vivo* studies to exposure levels and animal temperature increase, because these are the variables most likely affecting RF-EMF biological effects, and experimental animal species, because inter-species consistency of results was to be considered as an upgrading factor for the certainty of evidence. For *in vitro* studies, possible explored sources of heterogeneity were exposure levels and fertility status of sample donors. We did not explore sources of heterogeneity by differences in tested radiofrequencies because only 5 papers on animal studies and no paper on human sperm *in vitro* studies assessed effects at frequencies above 6000 MHz. This was the upper range in which a different mechanism of biological interaction might be expected because of short penetration depth into superficial tissues (a few mm or less). We did not explore the exposed male reproductive system stage of development (differentiating, pre-natal, pre-puberal and post-puberal exposures) because the large majority of *in vivo* studies were conducted in post-puberal animals.

Financial support

This project was partially funded by the World Health Organization (contracts 2020/1026306-0, 2022/1275453-1). WHO provided the basis for the protocol and methodological support throughout the review process. Additional in-kind funds were provided by ENEA, Health Canada and Swinburne University of Technology.

Declaration of Competing Interest

AWW previously directed a research group, which included two technical associates who are telecommunications company employees. AWW has been member of the ICNIRP Scientific Expert Group (SEG) from 2013 until 2021 and collaborates with the Australian Radiation Protection and Nuclear Safety Agency. JPMN was a member for IARC Monograph 102 Working Group assessing the carcinogenicity of RF-EMF (Mechanistic Studies sub-group), a co-author of Canada's Safety Code 6 (which are the *de facto* national human exposure limits applied in Canada) and a member of the WHO EMF Project International Advisory Committee (Canadian representative). Health Canada financially contributed to the WHO EMF Project to support the completion of the systematic reviews on RF-EMF. CM has been member of Technical Consultation on the WHO RF Research Agenda (2010), member of ICNIRP main commission since May 2012, confirmed in 2016 and 2020, Italian delegate for the European Cost Actions BM0704 and BM1309 "EMF-MED". All other authors declare that they have no known conflicts of interest.

Acknowledgments

We are grateful to Emilie van Deventer, Maria Rosaria Scarfì and Eric van Rongen for advice regarding the protocols draft and for discussions to ensure consistency in approaches across the multiple ongoing WHO systematic reviews. We wish to thank Flavio Di Marzio for his appreciated graphical help.

Appendix A. Supplementary material

Supplementary data to this article can be found online at **….**

References

Agarwal, A., N. R. Desai, K. Makker, A. Varghese, R. Mouradi, E. Sabanegh and R. Sharma. 2009. "Effects of radiofrequency electromagnetic waves (RF-EMW) from cellular phones on human ejaculated semen: an in vitro pilot study." Fertil Steril 92(4): 1318-1325. doi: 10.1016/j.fertnstert.2008.08.022

Advisory Group on Non-ionising Radiation (AGNIR), 2012. “Health Effects from Radiofrequency Electromagnetic Fields.” Report of the independent Advisory Group on Non-ionising Radiation. UK Health Protection Agency

Aitken, R. J., L. E. Bennetts, D. Sawyer, A. M. Wiklendt and B. V. King. 2005. "Impact of radio frequency electromagnetic radiation on DNA integrity in the male germline." Int J Androl 28(3): 171-179. doi: 10.1111/j.1365-2605.2005.00531.x

Akdağ, M. Z., M. S. Çelik, A. Ketani, Y. Nergiz, M. Deniz and S. Daşdağ. 1999. "Effect of chronic low-intensity microwave radiation on sperm count, sperm morphology, and testicular and epididymal tissues of rats." Electro- and Magnetobiology 18(2): 133-145

Almášiová, V., K. Holovská, S. Andrašková, V. Cigánková, Z. Ševčíková, A. Raček, Z. Andrejčáková, K. Beňová, Š. Tóth, E. Tvrdá, J. Molnár and E. Račeková. 2021. "Potential influence of prenatal 2.45 GHz radiofrequency electromagnetic field exposure on Wistar albino rat testis." Histol Histopathol: 18331. doi: 10.14670/hh-18-331

Almášiová, V., K. Holovská, V. Šimaiová, K. Beňová, A. Raček, E. Račeková, M. Martončíková, J. Mihálik, F. Horváthová, L. Tarabová, T. Slanina and V. Cigánková. 2017. "The thermal effect of 2.45 GHz microwave radiation on rat testes." Acta Veterinaria Brno 86(4): 413-419. doi: 10.2754/avb201786040413

Andrašková, S., K. Holovská, Z. Ševčíková, Z. Andrejčáková, Š. Tóth, M. Martončíková, E. Račeková and V. Almášiová. 2021. "The potential adverse effect of 2.45 GHz microwave radiation on the testes of prenatally exposed peripubertal male rats." Histol Histopathol: 18402. doi: 10.14670/hh-18-402

Atasoy, H. I., M. Y. Gunal, P. Atasoy, S. Elgun and G. Bugdayci. 2013. "Immunohistopathologic demonstration of deleterious effects on growing rat testes of radiofrequency waves emitted from conventional Wi-Fi devices." J Pediatr Urol 9(2): 223-229. doi: 10.1016/j.jpurol.2012.02.015

Auger, J., F. Eustache, C. Chevrier, B. Jégou. 2022. "Spatiotemporal trends in human semen quality.” Nat Rev Urol 19: 597-626. doi:10.1038/s41585-022-00626-w

Avendaño, C., A. Mata, C. A. Sanchez Sarmiento and G. F. Doncel. 2012. "Use of laptop computers connected to internet through Wi-Fi decreases human sperm motility and increases sperm DNA fragmentation." Fertil Steril 97(1): 39-45.e32. doi: 10.1016/j.fertnstert.2011.10.012

Azimzadeh, M. and G. Jelodar. 2019. "Alteration of testicular regulatory and functional molecules following long-time exposure to 900 MHz RFW emitted from BTS." Andrologia 51(9): e13372. doi: 10.1111/and.13372

Beechey, C. V., D. Brooker, C. I. Kowalczuk, R. D. Saunders and A. G. Searle. 1986. "Cytogenetic effects of microwave irradiation on male germ cells of the mouse." Int J Radiat Biol Relat Stud Phys Chem Med 50(5): 909-918. doi: 10.1080/09553008614551321

Berman, E., H. B. Carter and D. House. 1980. "Tests of mutagenesis and reproduction in male rats exposed to 2,450-MHz (CW) microwaves." Bioelectromagnetics 1(1): 65-76. doi: 10.1002/bem.2250010107

Bilgici, B., S. Gun, B. Avci, A. Akar and K. E. B. 2018. "What is adverse effect of wireless local area network, using 2.45 GHz, on the reproductive system?" Int J Radiat Biol 94(11): 1054-1061. doi: 10.1080/09553002.2018.1503430

Bodewein, L., K. Schmiedchen, D. Dechent, D. Stunder, D. Graefrath, L. Winter, T. Kraus, S. Driessen. 2019. “Systematic review on the biological effects of electric, magnetic and electromagnetic fields in the intermediate frequency range (300 Hz to 1 MHz).” Environ. Res. 171: 247–259 doi:10.1016/j.envres.2019.01.015

Boulicault, M., M. Perret, J. Galka, A. Borsa, A. Gompers, M. Reiches, S. Richardson. 2022. "The future of sperm: a biovariability framework for understanding global sperm count trends.” Hum Fertil (Camb) 25: 888-902. doi:10.1080/14647273.2021.1917778

Cairnie, A. B. and R. K. Harding. 1981. "Cytological studies in mouse testis irradiated with 2.45-GHz continuous-wave microwaves." Radiat Res 87(1): 100-108

Cao, D., Y. Zhang and C. Zhou. 2005. "Effects of electromagnetic radiation of cellular telephone on lactic dehydrogenase-X isozyme of testis." Medical Journal of Wuhan University 26(1): 58-61

Çetkin, M., N. Kızılkan, C. Demirel, Z. Bozdağ, S. Erkılıç and H. Erbağcı. 2017. "Quantitative changes in testicular structure and function in rat exposed to mobile phone radiation." Andrologia 49(10). doi: 10.1111/and.12761

Chaturvedi, C. M., V. P. Singh, P. Singh, P. Basu, M. Singaravel, R. K. Shukla, A. Dhawan, A. K. Pati, R. K. Gangwar and S. P. Singh. 2011. "2.45GHz (CW) Microwave irradiation alters circadian organization, spatial memory, DNA structure in the brain cells and blood cell counts of male mice, mus musculus." Progress In Electromagnetics Research B(29): 23-42. doi: 10.2528/PIERB11011205

Chen, L., F. Qin, Y. Chen, J. Sun and J. Tong. 2014. "[Chronotoxicity of 1800 MHz microwave radiation on sex hormones and spermatogenesis in male mice]." Wei Sheng Yan Jiu 43(1): 110-115

Journal Pre-proofs

Cobb, B. L., J. R. Jauchem, P. A. Mason, M. P. Dooley, S. A. Miller, J. M. Ziriax and M. R. Murphy. 2000. "Neural and behavioral teratological evaluation of rats exposed to ultra-wideband electromagnetic fields." Bioelectromagnetics 21(7): 524-537. doi: 10.1002/1521-186x(200010)21:7<524::aid-bem6>3.0.co;2-j

Dasdag, S., M. Z. Akdag, F. Aksen, F. Yilmaz, M. Bashan, M. M. Dasdag and M. S. Celik. 2003. "Whole Body Exposure of Rats to Microwaves Emitted from a Cell Phone Does Not Affect the Testes." Bioelectromagnetics 24(3): 182-188. doi: 10.1002/bem.10083

Dasdag, S., M. Z. Akdag, E. Ulukaya, A. K. Uzunlar and D. Yegin. 2008. "Mobile phone exposure does not induce apoptosis on spermatogenesis in rats." Arch Med Res 39(1): 40-44. doi: 10.1016/j.arcmed.2007.06.013

Dasdag, S., M. A. Ketani, Z. Akdag, A. R. Ersay, I. Sari, O. C. Demirtas and M. S. Celik. 1999. "Whole-body microwave exposure emitted by cellular phones and testicular function of rats." Urol Res 27(3): 219-223. doi: 10.1007/s002400050113

Dasdag, S., M. Taş, M. Z. Akdag and K. Yegin. 2015. "Effect of long-term exposure of 2.4 GHz radiofrequency radiation emitted from Wi-Fi equipment on testes functions." Electromagn Biol Med 34(1): 37-42. doi: 10.3109/15368378.2013.869752

De Iuliis, G. N., R. J. Newey, B. V. King and R. J. Aitken. 2009. "Mobile phone radiation induces reactive oxygen species production and DNA damage in human spermatozoa in vitro." PLoS One 4(7): e6446. doi: 10.1371/journal.pone.0006446

Delavarifar, S., Z. Razi, A. Tamadon, F. Rahmanifar, D. Mehrabani, M. Owjfard, O. Koohi-Hoseinabadi and S. Zaker Abasali. 2020. "Low-power density radiations emitted from common wi-fi routers influence sperm concentration and sperm histomorphometric parameters: A new horizon on male infertility treatment." Journal of Biomedical Physics and Engineering 10(2): 167-176. doi: 10.31661/jbpe.v0i0.581

Dong, G., H. Zhou, Y. Gao, X. Zhao, Q. Liu, Z. Li, X. Zhao, J. Yin and C. Wang. 2021. "Effects of 1.5-GHz high-power microwave exposure on the reproductive systems of male mice." Electromagn Biol Med: 1-10. doi: 10.1080/15368378.2021.1891091

Durney, C.G., H. Massoudi, M.F. Iskander. 1986. “Radiofrequency Radiation Dosimetry Handbook”. 4th Edition. USAFSAM-TR-85-73, Brooks Air Force Base, TX 78235-5301. https://apps.dtic.mil/sti/pdfs/ADA180678.pdf

Egger, M., G.Davey Smith, M.Schneider, C.Minder. 1997. “Bias in meta-analysis detected by a simple, graphical test.” BMJ 315: 629–634. doi:10.1136/bmj.315.7109.629

Er, H., G. G. Tas, B. Soygur, S. Ozen and L. Sati. 2022. "Acute and Chronic Exposure to 900 MHz Radio Frequency Radiation Activates p38/JNK-mediated MAPK Pathway in Rat Testis." Reproductive Sciences. doi: 10.1007/s43032-022-00844-y

Erdemli, C., S. Omeroglu, B. Sirav, M. Colbay, N. Seyhan, S. Ozkan and I. Yetkin. 2017. "Effects of 2100 MHz radio frequency radiation on ductus epididymis tissue in rats." Bratisl Lek Listy 118(12): 759-764. doi: 10.4149/bll_2017_143

Fahim, M. S., Z. Fahim, R. Der, D. G. Hall and J. Harman. 1975. "Heat in male contraception (hot water 60 °C, infrared, microwave, and ultrasound)." Contraception 11(5): 549-562. doi: 10.1016/0010-7824(75)90109-2

Falzone, N., C. Huyser, P. Becker, D. Leszczynski and D. R. Franken. 2011. "The effect of pulsed 900-MHz GSM mobile phone radiation on the acrosome reaction, head morphometry and zona binding of human spermatozoa." Int J Androl 34(1): 20-26. doi: 10.1111/j.1365-2605.2010.01054.x

Falzone, N., C. Huyser, F. Fourie, T. Toivo, D. Leszczynski and D. Franken. 2008. "In vitro effect of pulsed 900 MHz GSM radiation on mitochondrial membrane potential and motility of human spermatozoa." Bioelectromagnetics 29(4): 268-276. doi: 10.1002/bem.20390

Falzone, N., C. Huyser, D. R. Franken and D. Leszczynski. 2010. "Mobile phone radiation does not induce pro-apoptosis effects in human spermatozoa." Radiat Res 174(2): 169-176. doi: 10.1667/rr2091.1

Forgács, Z., Z. Somosy, G. Kubinyi, J. Bakos, A. Hudák, A. Surján and G. Thuróczy. 2006. "Effect of whole-body 1800MHz GSM-like microwave exposure on testicular steroidogenesis and histology in mice." Reprod Toxicol 22(1): 111-117. doi: 10.1016/j.reprotox.2005.12.003

Gao, X. H., H. R. Hu, X. L. Ma, J. Chen and G. H. Zhang. 2016. "[Cellphone electromagnetic radiation damages the testicular ultrastructure of male rats]." Zhonghua Nan Ke Xue 22(6): 491-495

Gautam, R., E. Priyadarshini, J. P. Nirala, R. Meena and P. Rajamani. 2021. "Modulatory effects of Punica granatum L juice against 2115 MHz (3G) radiation-induced reproductive toxicity in male Wistar rat." Environ Sci Pollut Res Int. doi: 10.1007/s11356-021-14378-4

Gautam, R., K. V. Singh, J. Nirala, N. N. Murmu, R. Meena and P. Rajamani. 2019. "Oxidative stress-mediated alterations on sperm parameters in male Wistar rats exposed to 3G mobile phone radiation." Andrologia 51(3): e13201. doi: 10.1111/and.13201

Ghanbari, M., S. B. Mortazavi, A. Khavanin and M. Khazaei. 2013. "The Effects of Cell Phone Waves (900 MHz-GSM Band) on Sperm Parameters and Total Antioxidant Capacity in Rats." Int J Fertil Steril 7(1): 21-28

Golbach, L.A., L.A. Portelli, H.F.J. Savelkoul, S. R. Terwel, N. Kuster, R. B. M. de Vries, B. M. L. Verburg-van Kemenade. 2016. “Calcium homeostasis and low-frequency magnetic and electric field exposure: a systematic review and meta-analysis of in vitro studies.” Environ. Int. 92-93, 695–706 doi:10.1016/j.envint.2016.01.014

Golub, M. S., C. A. Sobin. 2020. Statistical modeling with litter as a random effect in mixed models to manage "intralitter likeness". Neurotoxicol Teratol. 77: 106841. doi: 10.1016/j.ntt.2019.106841

Goud, S. N., M. V. Usha Rani, P. P. Reddy, O. S. Reddi, M. S. Rao and V. K. Saxena. 1982. "Genetic effects of microwave radiation in mice." Mutation Research Letters 103(1): 39-42. doi: 10.1016/0165-7992(82)90084-7

Journal Pre-proofs

Guo, L., J. J. Lin, Y. Z. Xue, G. Z. An, J. P. Zhang, K. Y. Zhang, W. He, H. Wang, W. Li and G. R. Ding. 2019. "Effects of 220 MHz Pulsed Modulated Radiofrequency Field on the Sperm Quality in Rats." Int J Environ Res Public Health 16(7). doi: 10.3390/ijerph16071286

Gupta. 2022. Reproductive and Developmental Toxicology. Third Edition. Academic Press

Gur, F. M., A. I. Keles, H. S. Erol, C. Guven, E. Taskin, H. Kaya, H. E. Gur, E. Odaci, M. B. Halici and S. Timurkaan. 2021. "The effect of 900-MHz radiofrequency electromagnetic fields during the adolescence on the histological structure of rat testis and its androgen and estrogen receptors localization." International Journal of Radiation Research 19(1): 135-144. doi: 10.29252/IJRR.19.1.135

Guyatt, G., A.D. Oxman, E.A. Akl, R. Kunz, G. Vist, J. Brozek, S. Norris, Y. Falck-Ytter, P. Glasziou, H. DeBeer, R. Jaeschke, D. Rind, J. Meerpohl, P. Dahm, H.J. Schünemann. 2011. “GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables.” J Clin Epidemiol 64: 383-394. doi: 10.1016/j.jclinepi.2010.04.026

Hanci, H., G. Kerimoğlu, T. Mercantepe and E. Odaci. 2018. "Changes in testicular morphology and oxidative stress biomarkers in 60-day-old Sprague Dawley rats following exposure to continuous 900-MHz electromagnetic field for 1 h a day throughout adolescence." Reprod Toxicol 81: 71-78. doi: 10.1016/j.reprotox.2018.07.002

Hanci, H., E. Odaci, H. Kaya, Y. Aliyazicioglu, I. Turan, S. Demir and S. Colakoglu. 2013. "The effect of prenatal exposure to 900-megahertz electromagnetic field on the 21-old-day rat testicle." Reprod Toxicol 42: 203-209. doi: 10.1016/j.reprotox.2013.09.006

Higgins J. P. T., T. Li. 2022. Chapter 6: Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022). Cochrane,. Deeks JJ (editors). Available from www.training.cochrane.org/handbook

Houston, B. J., B. Nixon, K. E. McEwan, J. H. Martin, B. V. King, R. J. Aitken and G. N. De Iuliis. 2019. "Whole-body exposures to radiofrequency-electromagnetic energy can cause DNA damage in mouse spermatozoa via an oxidative mechanism." Sci Rep 9(1): 17478. doi: 10.1038/s41598-019-53983-9

Huai, C. and W. Min. 1984. "Morphological change in mouse sperm following microwave exposure." Journal of Bioelectricity 3(3): 367-372

Hooijmans, C.R., R.B.M.de Vries, M.Ritskes-Hoitinga,M.M. Rovers, M.M. Leeflang, J. IntHout, K.E.Wever, L. Hooft, H.de Beer, T. Kuijpers, M.R. Macleod, E. S. Sena, G. ter Riet, R.L. Morgan, K.A. Thayer, A.ARooney, G.H. Guyatt, H.J. Schünemann, M.W.Langendam,J. Minnerup. 2018. “Facilitating healthcare decisions by assessing the certainty in the evidence from preclinical animal studies.” PLoS One. 13 (1): e0187271. doi:10.1371/journal.pone.0187271

Hooijmans, C.R., M.M. Rovers, R.B. de Vries, M. Leenaars, M.Ritskes-Hoitinga, M.W.Langendam. 2014. “SYRCLE’s risk of bias tool for animal studies.“ BMC Med. Res. Methodol. 14: 43. doi:10.1186/1471-2288-14-43

Imai, N., M. Kawabe, T. Hikage, T. Nojima, S. Takahashi and T. Shirai. 2011. "Effects on rat testis of 1.95-GHz W-CDMA for IMT-2000 cellular phones." Syst Biol Reprod Med 57(4): 204-209. doi: 10.3109/19396368.2010.544839

International Commission on Non-Ionizing Radiation Protection (ICNIRP), 2020. Guidelines for Limiting Exposure to Electromagnetic Fields (100 kHz to 300 GHz). Health Phys. 118, 483–524

Jaffar, F.H.F., K. Osman, N.H. Ismail, K.Y. Chin, S.F. Ibrahim. 2019. “Adverse effects of Wi-Fi radiation on male reproductive system: a systematic review.” Tohoku J. Exp. Med. 248 (3): 169–179 doi:10.1620/tjem.248.169

Jensh, R. P., 1984. "Studies of the teratogenic potential of exposure of rats to 6000-MHz microwave radiation. II. Postnatal psychophysiologic evaluations." Radiat Res 97(2): 282-301

Jensh, R. P., W. H. Vogel and R. L. Brent. 1982. "Postnatal functional analysis of prenatal exposure of rats to 915 MHz microwave radiation." J Am Coll Toxicol 1(3): 73-90

Jensh, R. P., W. H. Vogel and R. L. Brent. 1983. "An evaluation of the teratogenic potential of protracted exposure of pregnant rats to 2450-MHz microwave radiation. II. Postnatal psychophysiologic analysis." J Toxicol Environ Health 11(1): 37-59. doi: 10.1080/15287398309530319

Jin, Y. B., H. D. Choi, B. C. Kim, J. K. Pack, N. Kim and Y. S. Lee. 2013. "Effects of simultaneous combined exposure to CDMA and WCDMA electromagnetic fields on serum hormone levels in rats." J Radiat Res 54(3): 430-437. doi: 10.1093/jrr/rrs120

Johnson, L., R. M. Lebovitz and W. K. Samson. 1984. "Germ cell degeneration in normal and microwave-irradiated rats: potential sperm production rates at different developmental steps in spermatogenesis." Anat Rec 209(4): 501-507. doi: 10.1002/ar.1092090410

Jonwal, C., R. Sisodia, V. K. Saxena and K. K. Kesari. 2018. "Effect of 2.45 GHz microwave radiation on the fertility pattern in male mice." Gen Physiol Biophys 37(4): 453-460. doi: 10.4149/gpb_2017059

Kenny, R.P.W., E. B Millar., A. Adesanya, C. Richmond, F. Beyer, C. Calderon, J. Rankin, M. Toledano, M. Feychting, M.S. Pearce, D. Craig, and F. Pearson. 2022. “The effects of radiofrequency exposure on male fertility and adverse reproductive outcomes: A protocol for two systematic reviews of human observational studies with meta-analysis.” Environ Int. 158: 106968. doi: 10.1016/j.envint.2021.106968

Kesari, K.K., A. Agarwal, R. Henkel. 2018. “Radiations and male fertility.” Reprod. Biol. Endocrinol. 16: 118. doi:10.1186/s12958-018-0431-1

Kesari, K. K. and J. Behari. 2012. "Evidence for mobile phone radiation exposure effects on reproductive pattern of male rats: role of ROS." Electromagn Biol Med 31(3): 213-222. doi: 10.3109/15368378.2012.700292

Khillare, B. and J. Behari. 1998. "Effect of amplitude-modulated radiofrequency radiation on reproduction pattern in rats." Electromagnetic Biology and Medicine 17(1): 43-55

Kim, J. Y., H. T. Kim, K. H. Moon and H. J. Shin. 2007. "Long-term exposure of rats to a 2.45 GHz electromagnetic field: Effects on reproductive function." Korean Journal of Urology 48(12): 1308-1314. doi: 10.4111/kju.2007.48.12.1308

Kim, S., D. Han, J. Ryu, K. Kim, Y.H. Kim. 2021. "Effects of mobile phone usage on sperm quality - No time-dependent relationship on usage: A systematic review and updated meta-analysis." Environ Res 202: 111784. doi:10.1016/j.envres.2021.111784

Kismali, G., E. Ozgur, S. Sayiner, B. Alpaslan, G. Guler, N. Seyhan and T. Sel. 2009. "The evaluation of epigallocatechin gallate and N-acetylcysteine on serum testosterone levels in male guinea pigs expose to cell phone microwave." Journal of Animal and Veterinary Advances 8(6): 1149-1151

Kowalczuk, C. I., R. D. Saunders and H. R. Stapleton. 1983. "Sperm count and sperm abnormality in male mice after exposure to 2.45 GHz microwave radiation." Mutat Res 122(2): 155-161. doi: 10.1016/0165-7992(83)90054-4

Kumar, S., J. Behari and R. Sisodia. 2013. "Influence of electromagnetic fields on reproductive system of male rats." Int J Radiat Biol 89(3): 147-154. doi: 10.3109/09553002.2013.741282

Kumar, S., K. K. Kesari and J. Behari. 2011. "The therapeutic effect of a pulsed electromagnetic field on the reproductive patterns of male Wistar rats exposed to a 2.45-GHz microwave field." Clinics (Sao Paulo) 66(7): 1237-1245. doi: 10.1590/s1807-59322011000700020

Kumar, S., J. P. Nirala, J. Behari and R. Paulraj. 2014. "Effect of electromagnetic irradiation produced by 3G mobile phone on male rat reproductive system in a simulated scenario." Indian J Exp Biol 52(9): 890-897

Kuster, N. and F. Schonborn. 2000. Recommended minimal requirements and development guidelines for exposure setups of bio-experiments addressing the health risk concerns of wireless communications. Bioelectromagnetics 21: 508-514. doi: 10.1002/1521-186x(200010)21:7<508::aid-bem4>3.0.co;2-f

L'Abbate, N., R. De Leo, R. Ricco and G. Sborgia. 1982. "Experimental results in rabbits exposed to 100 MHz radiofrequencies." Giornale Italiano di Medicina del Lavoro 4(1): 31-38

Lebovitz, R. M. and L. Johnson. 1983. "Testicular function of rats following exposure to microwave radiation." Bioelectromagnetics 4(2): 107-114. doi: 10.1002/bem.2250040202

Lebovitz, R. M. and L. Johnson. 1987a. "Acute, whole-body microwave exposure and testicular function of rats." Bioelectromagnetics 8(1): 37-43. doi: 10.1002/bem.2250080106

Lebovitz, R. M., L. Johnson and W. K. Samson. 1987b. "Effects of pulse-modulated microwave radiation and conventional heating on sperm production." J Appl Physiol (1985) 62(1): 245-252. doi: 10.1152/jappl.1987.62.1.245

Lee, H. J., Y. B. Jin, T. H. Kim, J. K. Pack, N. Kim, H. D. Choi, J. S. Lee and Y. S. Lee. 2012. "The effects of simultaneous combined exposure to CDMA and WCDMA electromagnetic fields on rat testicular function." Bioelectromagnetics 33(4): 356-364. doi: 10.1002/bem.20715

Lee, H. J., J. K. Pack, T. H. Kim, N. Kim, S. Y. Choi, J. S. Lee, S. H. Kim and Y. S. Lee. 2010. "The lack of histological changes of CDMA cellular phone-based radio frequency on rat testis." Bioelectromagnetics 31(7): 528-534. doi: 10.1002/bem.20589

Lee, J. S., T. Q. Huang, J. J. Lee, J. K. Pack, J. J. Jang and J. S. Seo. 2005. "Subchronic exposure of hsp70.1-deficient mice to radiofrequency radiation." Int J Radiat Biol 81(10): 781-792. doi: 10.1080/09553000500500188

Lerchl, A., H. Krüger, M. Niehaus, J. R. Streckert, A. K. Bitz and V. Hansen. 2008. "Effects of mobile phone electromagnetic fields at nonthermal SAR values on melatonin and body weight of Djungarian hamsters (Phodopus sungorus)." J Pineal Res 44(3): 267-272. doi: 10.1111/j.1600-079X.2007.00522.x

Levine H., N. Jørgensen, A. Martino-Andrade, J. Mendiola, D. Weksler-Derri, I. Mindlis, R. Pinotti, S. H. Swan. 2017. “Temporal trends in sperm count: a systematic review and meta-regression analysis.” Hum Reprod Update 23(6): 646-659. doi: 10.1093/humupd/dmx022

Li, J. H., D. P. Jiang, Y. F. Wang, J. J. Yan, Q. Y. Guo, X. Miao, H. Y. Lang, S. L. Xu, J. Y. Liu and G. Z. Guo. 2017. "Influence of electromagnetic pulse on the offspring sex ratio of male BALB/c mice." Environ Toxicol Pharmacol 54: 155-161. doi: 10.1016/j.etap.2017.06.015

Liu, Q., T. Si, X. Xu, F. Liang, L. Wang and S. Pan. 2015. "Electromagnetic radiation at 900 MHz induces sperm apoptosis through bcl-2, bax and caspase-3 signaling pathways in rats." Reprod Health 12: 65. doi: 10.1186/s12978-015-0062-3

Luo, Y., X. Wang, Y. Chen, S. Xu, G. Ding and C. Shi. 2013. "Effects of electromagnetic radiation on morphology and TGF-β3 expression in mouse testicular tissue." Toxicology 310: 8-14. doi: 10.1016/j.tox.2013.05.004

Ma, H. R., X. H. Cao, X. L. Ma, J. J. Chen, J. W. Chen, H. Yang and Y. X. Liu. 2015. "[Protective effect of Liuwei Dihuang Pills against cell phone electromagnetic radiation-induced histomorphological abnormality, oxidative injury, and cell apoptosis in rat testes]." Zhonghua Nan Ke Xue 21(8): 737-741

Ma, H. R., Y. Y. Li, Y. P. Luo, X. L. Ma and Z. Q. Gong. 2014. "[Effect of Guilingji Capsule on the fertility, liver functions, and serum LDH of male SD rats exposed by 900 mhz cell phone]." Zhongguo Zhong Xi Yi Jie He Za Zhi 34(4): 475-479

Makler, A., M. Tatcher, A. Vilensky and J. M. Brandes. 1980. "Factors affecting sperm motility. III. Influence of visible light and other electromagnetic radiations on human sperm velocity and survival." Fertil Steril 33(4): 439-444. doi: 10.1016/s0015-0282(16)44664-9

Meena, R., K. Kumari, J. Kumar, P. Rajamani, H. N. Verma and K. K. Kesari. 2014. "Therapeutic approaches of melatonin in microwave radiations-induced oxidative stress-mediated toxicity on male fertility pattern of Wistar rats." Electromagn Biol Med 33(2): 81-91. doi: 10.3109/15368378.2013.781035

Miao, X., Y. Wang, H. Lang, Y. Lin, Q. Guo, M. Yang, J. Guo, Y. Zhang, J. Zhang, J. Liu, Y. Liu, L. Zeng and G. Guo. 2017. "Preventing Electromagnetic Pulse Irradiation Damage on Testis Using Selenium-rich Cordyceps Fungi. A Preclinical Study in Young Male Mice." Omics 21(2): 81-89. doi: 10.1089/omi.2016.0151

Nakatani-Enomoto, S., M. Okutsu, S. Suzuki, R. Suganuma, S. J. Groiss, S. Kadowaki, H. Enomoto, K. Fujimori and Y. Ugawa. 2016. "Effects of 1950 MHz W-CDMA-like signal on human spermatozoa." Bioelectromagnetics 37(6): 373-381. doi: 10.1002/bem.21985

National Toxicology Program (NTP), 2015a. Handbook for Conducting a Literature based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. Office of Health Assessment and Translation

National Institute of Environmental Health Sciences. National Toxicology Program (NTP), 2015b. OHAT Risk of Bias Rating Tool for Human and Animal Studies. Office of Health Assessment and Translation

Odacı, E., H. Hancı, E. Yuluğ, S. Türedi, Y. Aliyazıcıoğlu, H. Kaya and S. Çolakoğlu. 2016. "Effects of prenatal exposure to a 900 MHz electromagnetic field on 60-day-old rat testis and epididymal sperm quality." Biotech Histochem 91(1): 9-19. doi: 10.3109/10520295.2015.1060356

Odaci, E. and C. Ozyilmaz. 2015. "Exposure to a 900 MHz electromagnetic field for 1 hour a day over 30 days does change the histopathology and biochemistry of the rat testis." Int J Radiat Biol 91(7): 547-554. doi: 10.3109/09553002.2015.1031850

Oh, J. J., S. S. Byun, S. E. Lee, G. Choe and S. K. Hong. 2018. "Effect of Electromagnetic Waves from Mobile Phones on Spermatogenesis in the Era of 4G-LTE." Biomed Res Int 2018: 1801798. doi: 10.1155/2018/1801798

Oksay, T., M. Naziroğlu, S. Doğan, A. Güzel, N. Gümral and P. A. Koşar. 2014. "Protective effects of melatonin against oxidative injury in rat testis induced by wireless (2.45 GHz) devices." Andrologia 46(1): 65-72. doi: 10.1111/and.12044

Orsini N. and D. Spiegelhalter. 2021. “Meta-analysis of dose-response relationships, chapter 18. In: Handbook of Meta-analysis.” Eds Schmid CH, Stijnen T, White IR. Publ CRC Press, Boca Raton FL USA

Ozguner, M., A. Koyu, G. Cesur, M. Ural, F. Ozguner, A. Gokcimen and N. Delibas. 2005. "Biological and morphological effects on the reproductive organ of rats after exposure to electromagnetic field." Saudi Med J 26(3): 405-410

Ozlem Nisbet, H., C. Nisbet, A. Akar, M. Cevik and M. O. Karayigit. 2012. "Effects of exposure to electromagnetic field (1.8/0.9 GHz) on testicular function and structure in growing rats." Res Vet Sci 93(2): 1001-1005. doi: 10.1016/j.rvsc.2011.10.023

Pacchierotti, F., L. Ardoino, B. Benassi, C. Consales, E. Cordelli, P. Eleuteri, C. Marino, M. Sciortino, M. H. Brinkworth, G. Chen, J. P. McNamee, A. W. Wood, C. R. Hooijmans, R. B.M. de Vries. 2021. “Effects of Radiofrequency Electromagnetic Field (RF-EMF) exposure on male fertility and pregnancy and birth outcomes: Protocols for a systematic review of experimental studies in non-human mammals and in human sperm exposed in vitro.” Environ Int 157: 106806. doi: https://doi.org/10.1016/j.envint.2021.106806

Page, M. J., J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer , J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher.

2021. “The PRISMA 2020 statement: an updated guideline for reporting systematic reviews.” Bmj. 372:n71. doi: 10.1136/bmj.n71

Pandey, N. and S. Giri. 2018. "Melatonin attenuates radiofrequency radiation (900 MHz)-induced oxidative stress, DNA damage and cell cycle arrest in germ cells of male Swiss albino mice." Toxicol Ind Health 34(5): 315-327. doi: 10.1177/0748233718758092

Pandey, N., S. Giri, S. Das and P. Upadhaya. 2017. "Radiofrequency radiation (900 MHz)-induced DNA damage and cell cycle arrest in testicular germ cells in swiss albino mice." Toxicol Ind Health 33(4): 373-384. doi: 10.1177/0748233716671206

Pardhiya, S., R. Gautam, J. P. Nirala, N. N. Murmu and P. Rajamani. 2022. "Modulatory role of Bovine serum albumin conjugated manganese dioxide nanoparticle on microwave radiation induced alterations in reproductive parameters of rat." Reprod Toxicol 113: 136-149. doi: 10.1016/j.reprotox.2022.09.003

Paul, C. Murray, A. A., Spears, N. Saunder, P. T. 2008. “A single, mild, transient scrotal heat stress causes DNA damage, subfertility and impairs formation of blastocysts in mice.” Reproduction 136(1):73-84. doi: 10.1530/REP-08-0036. Epub 2008 Apr 4. PMID: 18390691

Pedrosa, M. B., B. M. Tenorio, F. C. A. M. Tenorio, R. N. Morais, R. A. Nogueira and V. A. S. Junior. 2021. "Effects of the 27.12 mhz magnetic field emitted by short-wave equipment on spermatogenesis." Acta Scientiarum - Health Sciences 43: 1-12. doi: 10.4025/actascihealthsci.v43i1.53934

Poulletier de Gannes, F., B. Billaudel, E. Haro, M. Taxile, L. Le Montagner, A. Hurtier, S. Ait Aissa, H. Masuda, Y. Percherancier, G. Ruffié, P. Dufour, B. Veyret and I. Lagroye. 2013. "Rat fertility and embryo fetal development: influence of exposure to the Wi-Fi signal." Reprod Toxicol 36: 1-5. doi: 10.1016/j.reprotox.2012.11.003

Prausnitz, S. and C. Susskind. 1962. "Effects of Chronic Microwave Irradiation on Mice." IRE Transactions on Bio-Medical Electronics 9(2): 104-108. doi: 10.1109/TBMEL.1962.4322972

Qin, F., H. Cao, C. Feng, T. Zhu, B. Zhu, J. Zhang, J. Tong and H. Pei. 2021. "Microarray profiling of LncRNA expression in the testis of pubertal mice following morning and evening exposure to 1800 MHz radiofrequency fields." Chronobiology International 38(12): 1745-1760. doi: 10.1080/07420528.2021.1962902

Qin, F., H. Cao, H. Yuan, W. Guo, H. Pei, Y. Cao and J. Tong. 2018. "1800 MHz radiofrequency fields inhibits testosterone production via CaMKI /RORα pathway." Reprod Toxicol 81: 229-236. doi: 10.1016/j.reprotox.2018.08.014

Qin, F., J. Zhang, H. Cao, W. Guo, L. Chen, O. Shen, J. Sun, C. Yi, J. Li, J. Wang and J. Tong. 2014. "Circadian alterations of reproductive functional markers in male rats exposed to 1800 MHz radiofrequency field." Chronobiol Int 31(1): 123-133. doi: 10.3109/07420528.2013.830622

Qin, F., J. Zhang, H. Cao, C. Yi, J. X. Li, J. Nie, L. L. Chen, J. Wang and J. Tong. 2012. "Effects of 1800-MHz radiofrequency fields on circadian rhythm of plasma melatonin and testosterone in male rats." J Toxicol Environ Health A 75(18): 1120-1128. doi: 10.1080/15287394.2012.699846

Ren, D. D., X. X. Lu, W. Zhong, H. R. Ma, J. W. Chen and L. J. Sun. 2020. "[Guilingji Capsules reduce 900 MHz collphone electromagnetic radiation-induced testicular oxidative damage and downregulate Prdx2 protein expression in the rat testis]." Zhonghua Nan Ke Xue 26(10): 926-933

Ren, D. Q., W. Q. Yang and G. Y. Zeng. 2002. "Effects of microwave at 2 450 MHz on spermatic system in mice." Chinese Journal of Clinical Rehabilitation 6(6): 822-823

Ribeiro, E. P., E. L. Rhoden, M. M. Horn, C. Rhoden, L. P. Lima and L. Toniolo. 2007. "Effects of subchronic exposure to radio frequency from a conventional cellular telephone on testicular function in adult rats." J Urol 177(1): 395-399. doi: 10.1016/j.juro.2006.08.083

Romeo, S., O. Zeni, A. Sannino, S. Lagorio, M.Biffoni,M.R.Scarfì. 2021. “Genotoxicity of radiofrequency electromagnetic fields: protocol for a systematic review of in vitro studies.” Environ Int. 148: 106386. https://doi.org/10.1016/j.envint.2021.106386

Rugh, R. and M. McManaway. 1978. "Does pre- or post-natal microwave radiation sterilize mice?" Congenital Anomalies 18(2): 69-74

Saunders, R. D., S. C. Darby and C. I. Kowalczuk. 1983. "Dominant lethal studies in male mice after exposure to 2.45 GHz microwave radiation." Mutat Res 117(3-4): 345-356. doi: 10.1016/0165-1218(83)90134-9

Saunders, R. D. and C. I. Kowalczuk. 1981a. "Effects of 2.45 GHz microwave radiation and heat on mouse spermatogenic epithelium." Int J Radiat Biol Relat Stud Phys Chem Med 40(6): 623-632. doi: 10.1080/09553008114551611

Saunders, R. D. and C. I. Kowalczuk. 1981b. "The effect of acute far field exposure at 2.45 GHz on the mouse testis." Int J Radiat Biol Relat Stud Phys Chem Med 39(6): 587-596. doi: 10.1080/09553008114550711

Saunders, R. D., C. I. Kowalczuk, C. V. Beechey and R. Dunford. 1988. "Studies of the induction of dominant lethals and translocations in male mice after chronic exposure to microwave radiation." Int J Radiat Biol Relat Stud Phys Chem Med 53(6): 983-992. doi: 10.1080/09553008814551341

Saygin, M., H. Asci, O. Ozmen, F. N. Cankara, D. Dincoglu and I. Ilhan. 2016. "Impact of 2.45 GHz microwave radiation on the testicular inflammatory pathway biomarkers in young rats: The role of gallic acid." Environ Toxicol 31(12): 1771-1784. doi: 10.1002/tox.22179

Saygin, M., S. Caliskan, N. Karahan, A. Koyu, N. Gumral and A. Uguz. 2011. "Testicular apoptosis and histopathological changes induced by a 2.45 GHz electromagnetic field." Toxicol Ind Health 27(5): 455-463. doi: 10.1177/0748233710389851

Saygin, M., S. Caliskan, M. F. Ozguner, N. Gumral, S. Comlekci and N. Karahan. 2015. "Impact of L-carnitine and Selenium Treatment on Testicular Apoptosis in Rats Exposed to 2.45 GHz Microwave Energy." West Indian Med J 64(2): 55-61. doi: 10.7727/wimj.2014.205

Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR), 2015. Potential Health Effects of Exposure to Electromagnetic Fields (EMF). European Commission, Luxembourg. (https://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_041.pdf).

Sciorio, R., L. Tramontano, S.C. Esteves. 2022. "Effects of mobile phone radiofrequency radiation on sperm quality." Zygote 30: 159-168. doi:10.1017/s096719942100037x

Sepehrimanesh, M., M. Saeb, S. Nazifi, N. Kazemipour, G. Jelodar and S. Saeb. 2014. "Impact of 900 MHz electromagnetic field exposure on main male reproductive hormone levels: a Rattus norvegicus model." Int J Biometeorol 58(7): 1657-1663. doi: 10.1007/s00484-013-0771-7

Shahin, N. N., N. A. El-Nabarawy, A. S. Gouda and B. Mégarbane. 2019. "The protective role of spermine against male reproductive aberrations induced by exposure to electromagnetic field - An experimental investigation in the rat." Toxicol Appl Pharmacol 370: 117-130. doi: 10.1016/j.taap.2019.03.009

Shahin, S., V. Mishra, S. P. Singh and C. M. Chaturvedi. 2014. "2.45-GHz microwave irradiation adversely affects reproductive function in male mouse, Mus musculus by inducing oxidative and nitrosative stress." Free Radic Res 48(5): 511-525. doi: 10.3109/10715762.2014.888717

Shahin, S., S. P. Singh and C. M. Chaturvedi. 2018a. "2.45 GHz microwave radiation induced oxidative and nitrosative stress mediated testicular apoptosis: Involvement of a p53 dependent bax-caspase-3 mediated pathway." Environ Toxicol 33(9): 931-945. doi: 10.1002/tox.22578

Shibkova, D. Z., T. V. Shilkova and A. V. Ovchinnikova. 2015. "Early and Delayed Effects of Radio Frequency Electromagnetic Fields on the Reproductive Function and Functional Status of the Offspring of Experimental Animals." Radiatsionnaia biologiia, radioecologiia / Rossiĭskaia akademiia nauk 55(5): 514-519. doi: 10.7868/s0869803115050112

Shirai, T., N. Imai, J. Wang, S. Takahashi, M. Kawabe, K. Wake, H. Kawai, S. Watanabe, F. Furukawa and O. Fujiwara. 2014. "Multigenerational effects of whole body exposure to 2.14 GHz W-CDMA cellular phone signals on brain function in rats." Bioelectromagnetics 35(7): 497-511. doi: 10.1002/bem.21871

Shirai, T., J. Wang, M. Kawabe, K. Wake, S. I. Watanabe, S. Takahashi and O. Fujiwara. 2017. "No adverse effects detected for simultaneous whole-body exposure to multiple-frequency radiofrequency electromagnetic fields for rats in the intrauterine and pre- and post-weaning periods." J Radiat Res 58(1): 48-58. doi: 10.1093/jrr/rrw085

Šimaiová, V., V. Almášiová, K. Holovská, T. Kisková, F. Horváthová, Z. Ševčíková, Š. Tóth, A. Raček, E. Račeková, K. Beňová, P. Dvořák and V. Cigánková. 2019. "The effect of 2.45 GHz non-ionizing radiation on the structure and ultrastructure of the testis in juvenile rats." Histol Histopathol 34(4): 391-403. doi: 10.14670/hh-18-049

Smialowicz, R. J., J. S. Ali, E. Berman, S. J. Bursian, J. B. Kinn, C. G. Liddle, L. W. Reiter and C. M. Weil. 1981. "Chronic exposure of rats to 100-MHz (CW) radiofrequency radiation: Assessment of biological effects." Radiation Research 86(3): 488-505. doi: 10.2307/3575465

Sommer, A. M., K. Grote, T. Reinhardt, J. Streckert, V. Hansen and A. Lerchl. 2009. "Effects of radiofrequency electromagnetic fields (UMTS) on reproduction and development of mice: a multi-generation study." Radiat Res 171(1): 89-95. doi: 10.1667/rr1460.1

Sterling, L., L.R. Harris, K. Carroll. 2022. "The effects of wireless devices on male reproductive health: A literature overview." Rev Int Androl 20: 196-206. doi:10.1016/j.androl.2020.10.004

Takahashi, S., N. Imai, K. Nabae, K. Wake, H. Kawai, J. Wang, S. Watanabe, M. Kawabe, O. Fujiwara, K. Ogawa, S. Tamano and T. Shirai. 2010. "Lack of adverse effects of whole-body exposure to a mobile telecommunication electromagnetic field on the rat fetus." Radiat Res 173(3): 362-372. doi: 10.1667/RR1615.1

Tang, Z., J. Wei, L. Zhang, Q. Zhai, J. Yuan and Z. Li. 2022. "Effects of millimeter wave radiation on thereproductive system of male mice." Fushe Yanjiu yu Fushe Gongyi Xuebao/Journal of Radiation Research and Radiation Processing 40(1). doi: 10.11889/j.1000-3436.2021-0106

Tas, M., S. Dasdag, M. Z. Akdag, U. Cirit, K. Yegin, U. Seker, M. F. Ozmen and L. B. Eren. 2014. "Long-term effects of 900 MHz radiofrequency radiation emitted from mobile phone on testicular tissue and epididymal semen quality." Electromagnetic Biology and Medicine 33(3): 216-222. doi: 10.3109/15368378.2013.801850

Trosic, I., M. Matausic-Pisl, I. Pavicic and A. M. Marjanovic. 2013. "Histological and cytological examination of rat reproductive tissue after short-time intermittent radiofrequency exposure." Arh Hig Rada Toksikol 64(4): 513-519. doi: 10.2478/10004-1254-64-2013-2394

Tumkaya, L., Y. Kalkan, O. Bas and A. Yilmaz. 2016. "Mobile phone radiation during pubertal development has no effect on testicular histology in rats." Toxicol Ind Health 32(2): 328-336. doi: 10.1177/0748233713500820

Veerachari, S. B. and S. S. Vasan. 2012. "Mobile phone electromagnetic waves and its effect on human ejaculated semen: An in vitro study." International Journal of Infertility and Fetal Medicine 3(1): 15-21. doi: 10.5005/jp-journals-10016-1034

Verbeek, J., G. Oftedal, M. Feychting, E. van Rongen, M. R. Scarfì, S. Mann, R. Wong, E. van Deventer. 2021. "Prioritizing health outcomes when assessing the effects of exposure to radiofrequency electromagnetic fields: a survey among experts." Env. Int. 146: 106300. doi.org/10.1016/j.envint.2020.106300

Vornoli, A., L. Falcioni, D. Mandrioli, L. Bua, F. Belpoggi. 2019. "The contribution of in vivo mammalian studies to the knowledge of adverse effects of radiofrequency radiation on human health." Int. J. Environ. Res. Public Health. 16 (18): 3379. doi:10.3390/ijerph16183379

Wang, D., B. Li, Y. Liu, Y. F. Ma, S. Q. Chen, H. J. Sun, J. Dong, X. H. Ma, J. Zhou and X. H. Wang. 2015. "[Impact of mobile phone radiation on the quality and DNA methylation of human sperm in vitro]." Zhonghua Nan Ke Xue 21(6): 515-520

Wang, S. M., D. W. Wang, R. Y. Peng, Y. B. Gao, Y. Yang, W. H. Hu, H. Y. Chen, Y. R. Zhang and Y. Gao. 2003. "Effect of electromagnetic pulse irradiation on structure and function of Leydig cells in mice." Zhonghua nan ke xue = National journal of andrology 9(5): 327-330

World Health Organization (WHO), 2014. Handbook for Guideline Development, second ed. WHO Press, Geneva, Switzerland

Xu, Y., Z. A. Zheng, T. Zhu, B. Zhu, C. Feng, Y. Chen and F. Qin. 2020. "[Joint effects of nano-selenium and nano-cerium on the male reproductive function of mice exposed to microwave radiation]." Wei Sheng Yan Jiu 49(5): 795-801. doi: 10.19813/j.cnki.weishengyanjiu.2020.05.018

Xue, Y., L. Guo, J. Lin, P. Lai, G. Rui, L. Liu, R. Huang, Y. Jing, F. Wang and G. Ding. 2022. "Effects of 5.8 GHz Microwaves on Testicular Structure and Function in Rats." Biomed Res Int 2022: 5182172. doi: 10.1155/2022/5182172

Yahyazadeh, A. and B. Z. Altunkaynak. 2019. "Protective effects of luteolin on rat testis following exposure to 900 MHz electromagnetic field." Biotech Histochem 94(4): 298-307. doi: 10.1080/10520295.2019.1566568

Yan, S., Y. Ju, J. Dong, H. Lei, J. Wang, Q. Xu, Y. Ma, J. Wang and X. Wang. 2022. "Paternal Radiofrequency Electromagnetic Radiation Exposure Causes Sex-Specific Differences in Body Weight Trajectory and Glucose Metabolism in Offspring Mice." Front Public Health 10: 872198. doi: 10.3389/fpubh.2022.872198

Yu, G., Z. Tang, H. Chen, Z. Chen, L. Wang, H. Cao, G. Wang, J. Xing, H. Shen, Q. Cheng, D. Li, G. Wang, Y. Xiang, Y. Guan, Y. Zhu, Z. Liu and Z. Bai. 2020. "Long-term exposure to 4G smartphone radiofrequency electromagnetic radiation diminished male reproductive potential by directly disrupting Spock3-MMP2-BTB axis in the testes of adult rats." Sci Total Environ 698: 133860. doi: 10.1016/j.scitotenv.2019.133860

Yu, G., Z. Bai, C. Song, Q. Cheng, G. Wang, Z. Tang, S. Yang. 2021. "Current progress on the effect of mobile phone radiation on sperm quality: An updated systematic review and meta-analysis of human and animal studies." Environ Pollut 282: 116952. doi:10.1016/j.envpol.2021.116952

Zeng, L., X. Ji, Y. Zhang, X. Miao, C. Zou, H. Lang, J. Zhang, Y. Li, X. Wang, H. Qi, D. Ren and G. Guo. 2011. "MnSOD expression inhibited by electromagnetic pulse radiation in the rat testis." Electromagnetic Biology and Medicine 30(4): 205-218. doi: 10.3109/15368378.2011.587929

Zhu, S., J. Zhang, C. Liu, Q. He, Vijayalaxmi, T. J. Prihoda, J. Tong and Y. Cao. 2015. "Dominant lethal mutation test in male mice exposed to 900MHz radiofrequency fields." Mutat Res Genet Toxicol Environ Mutagen 792: 53-57. doi: 10.1016/j.mrgentox.2015.07.004

Zwetsloot P.P., M. Van Der Naald, E. S. Sena, D. W. Howells, J. IntHout, J. A. De Groot, S. A. Chamuleau, M. R. MacLeod, K. E. and Wever. 2017.Standardized mean differences cause funnel plot distortion in publication bias assessments. Elife. 6:e24260. doi: 10.7554/eLife.24260

Figure Legends

Figure 1: The multiple endpoints selected as targets of possible RF-EMF effects. Beside direct markers of the male fertility performance, reproductive organ toxicity markers, from testis or epididymis weight to a variety of histopathological findings, semen quality markers assessed in sperm suspensions and testosterone level have been considered. All those are interconnected: testosterone regulates spermatogenesis and its homeostasis may be

affected by damage to reproductive organs, toxic effects in the testis and epididymis potentially causes alterations of semen quality, which in turn may compromise male fertility.

Figure 2: Flow chart of the paper selection process for experimental animal studies according to the template proposed by PRISMA 2020.

Figure 3: Flow chart of the paper selection process for human sperm *in vitro* studies according to the template proposed by PRISMA 2020.

Figure 4: Forest plot of experimental animal studies on the rate of infertile males categorised by the RoB level of concern. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group.

Figure 5: Forest plot of experimental animal studies on the incidence of nonpregnant females, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 6: Forest plot of studies on the size of litters sired by experimental males, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 7: Forest plot of experimental animal studies on sperm count categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 8: Forest plot of experimental animal studies on the percentage of morphologically abnormal sperm, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 9: Forest plot of experimental animal studies on the percentage of dead or immotile sperm, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 10: Forest plot of experimental animal studies on sperm DNA/chromatin alterations, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 11: Forest plot of *in vitro* studies on the percentage of dead or immotile human sperm, categorised as “low or some concern” or “high concern” for RoB. The bottom lines

report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 12: Forest plot of human sperm *in vitro* studies on DNA/chromatin alterations, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 13: Forest plot of experimental animal studies on testis or epididymis weight, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 14: Forest plot of experimental animal studies on testis histomorphometry, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 15: Forest plot of experimental animal studies on testis histopathology (Johnsen score), categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 16: Forest plot of experimental animal studies on testicular cell death, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 17: Forest plot of experimental animal studies on testicular sperm production, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Figure 18: Forest plot of experimental animal studies on testosterone level, categorised as “low or some concern” or “high concern” for RoB. The bottom lines report the results and statistics of the meta-analysis for all included studies. Asterisks mark studies in which data from multiple exposure groups were combined to match a single comparator group. Progressive numbers after a reference indicate different studies reported in the same paper.

Author statement

Individual study results included in the systematic review and meta-analyses are reported in Supplementary Files 4a and 4b. Extraction forms are available upon request.

Journal Pre-proofs

Figure 1

Click here to access/download;Figure;Cordelli-Male_fertility-Fig1-endpoint cartoon.tif

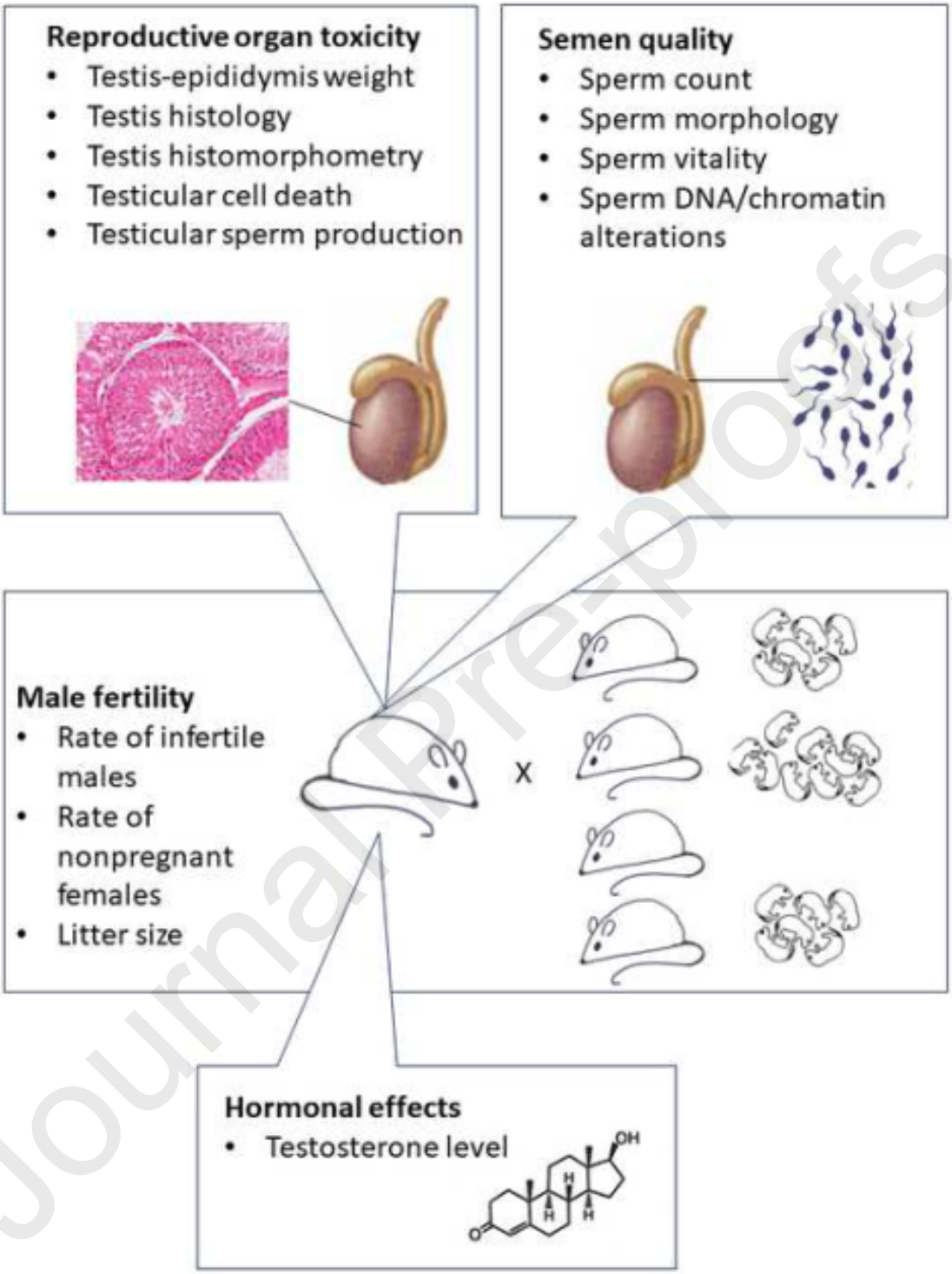


Figure 2

Figure 2

Click here to access/download;Figure;Cordelli-Male_fertility-Fig2-PRISMA_2020-in vivo.docx

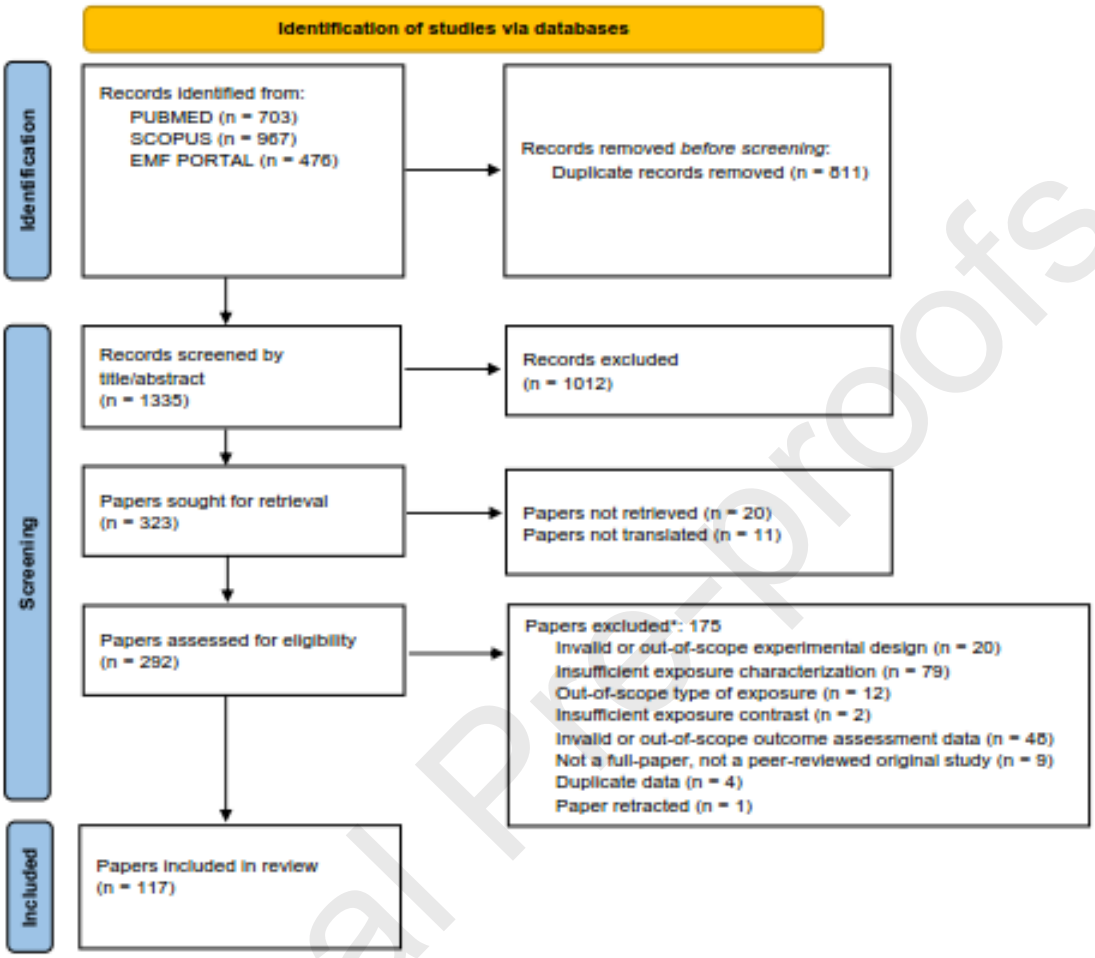


Figure 2

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Journal Pre-proofs

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Journal Pre-proofs

Figure 3

Click here to access/download;Figure;Cordelli-Male_fertility-Fig3-PRISMA_2020-in vitro.docx

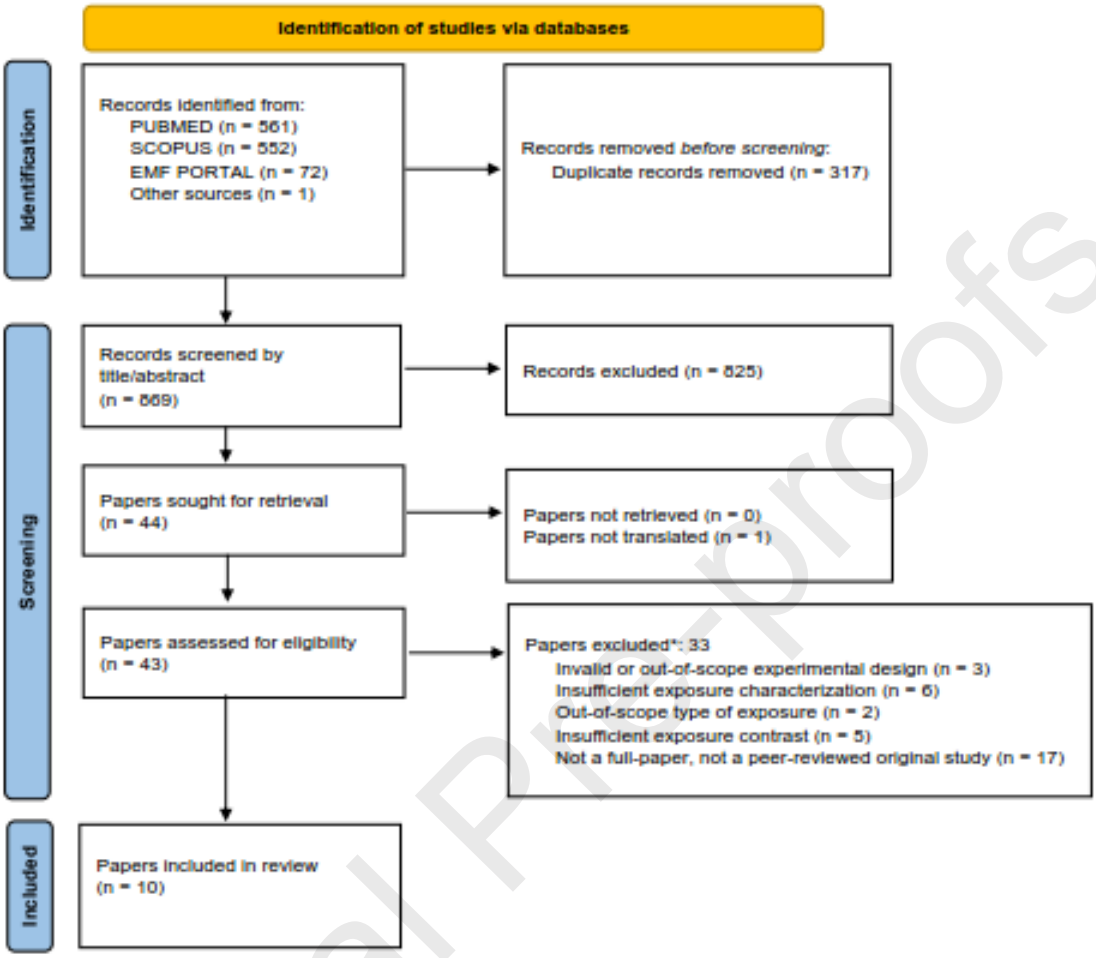


Figure 3

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Journal Pre-proofs

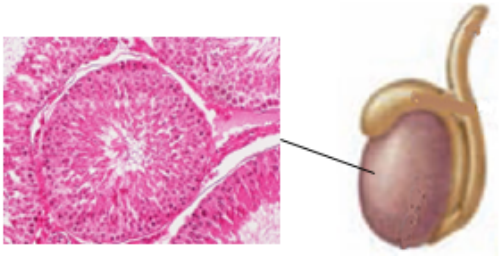
Figure 3

From: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71

Journal Pre-proofs

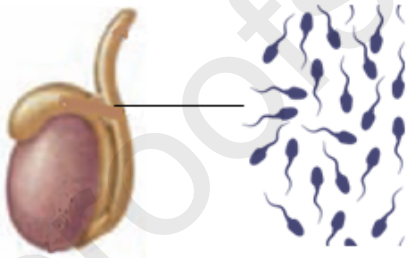
Reproductive organ toxicity

- Testis-epididymis weight
- Testis histology
- Testis histomorphometry
- Testicular cell death
- Testicular sperm production



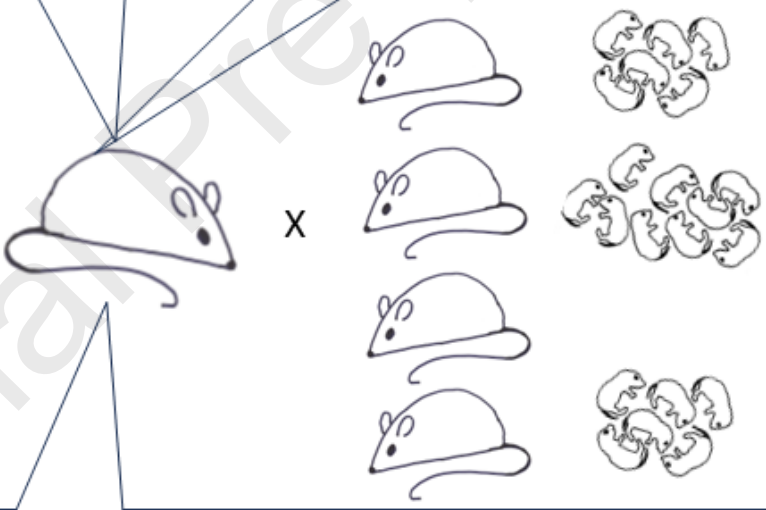
Semen quality

- Sperm count
- Sperm morphology
- Sperm vitality
- Sperm DNA/chromatin alterations



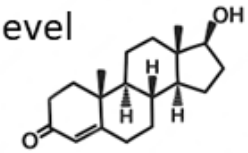
Male fertility

- Rate of infertile males
- Rate of nonpregnant females
- Litter size



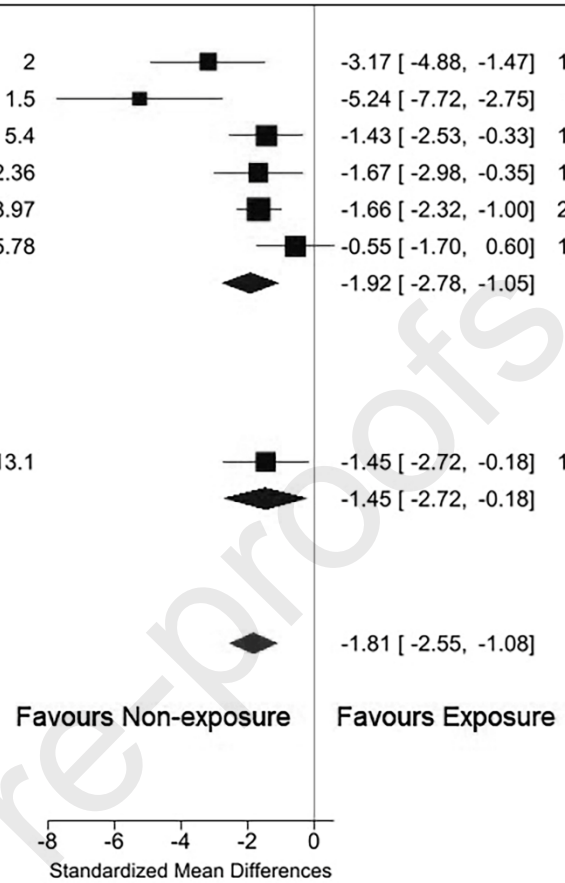
Hormonal effects

- Testosterone level



Journal Pre-proofs

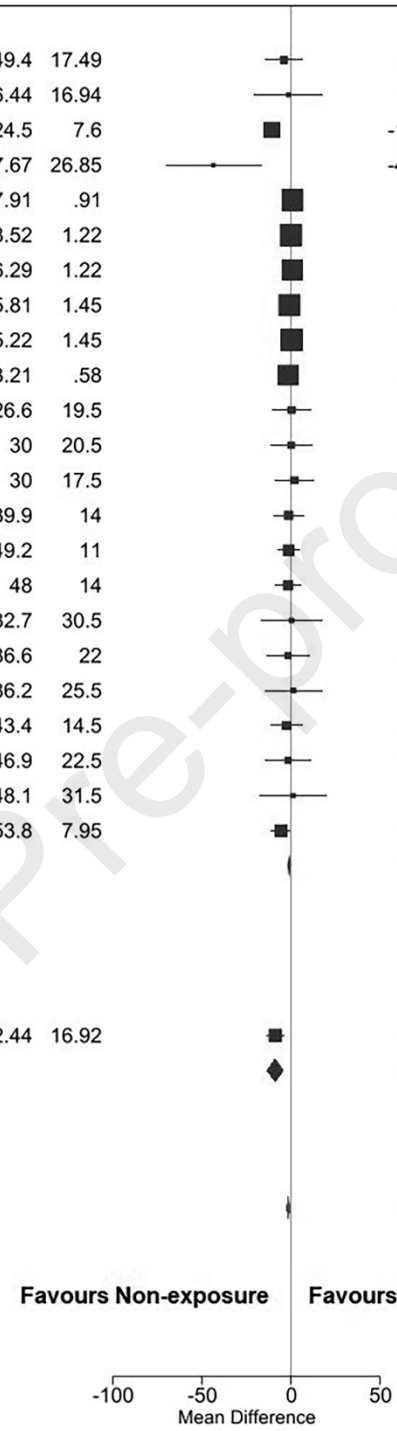
Study	Non-Exposed N	Non-Exposed Mean	Non-Exposed SD	Exposed N	Exposed Mean	Exposed SD	Cohen's d with 95% CI	Weight (%)
Low or Some Concern								
Houston 2019-1	6	1.51	.98	6	6.51	2	-3.17 [-4.88, -1.47]	10.84
Houston 2019-2	5	1.83	1.25	6	9.13	1.5	-5.24 [-7.72, -2.75]	6.51
Houston 2019-3	8	2.15	1.36	8	7.78	5.4	-1.43 [-2.53, -0.33]	16.45
Kumar 2014	6	6.73	4.91	6	22.42	12.36	-1.67 [-2.98, -0.35]	14.19
Liu 2015	24	11.51	2.6	24	22.47	8.97	-1.66 [-2.32, -1.00]	21.57
Yan 2022	6	3.08	2.65	6	5.56	5.78	-0.55 [-1.70, 0.60]	15.85
Heterogeneity: $\tau^2 = 0.71$, $I^2 = 65.96\%$, $H^2 = 2.94$							-1.92 [-2.78, -1.05]	
Test of $\theta_i = \theta_j$: Q(5) = 14.69, p = 0.01								
Test of $\theta = 0$: z = -4.34, p = 0.00								
High Concern								
Kumar 2013	6	1.5	2.01	6	15.1	13.1	-1.45 [-2.72, -0.18]	14.60
Heterogeneity: $\tau^2 = 0.00$, $I^2 = .\%$, $H^2 = .$							-1.45 [-2.72, -0.18]	
Test of $\theta_i = \theta_j$: Q(0) = -0.00, p = .								
Test of $\theta = 0$: z = -2.24, p = 0.03								
Overall							-1.81 [-2.55, -1.08]	
Heterogeneity: $\tau^2 = 0.54$, $I^2 = 59.45\%$, $H^2 = 2.47$								
Test of $\theta_i = \theta_j$: Q(6) = 14.79, p = 0.02								
Test of $\theta = 0$: z = -4.85, p = 0.00								
Test of group differences: $Q_b(1) = 0.35$, p = 0.55								



Random-effects DerSimonian–Laird model

Journal Pre-proofs

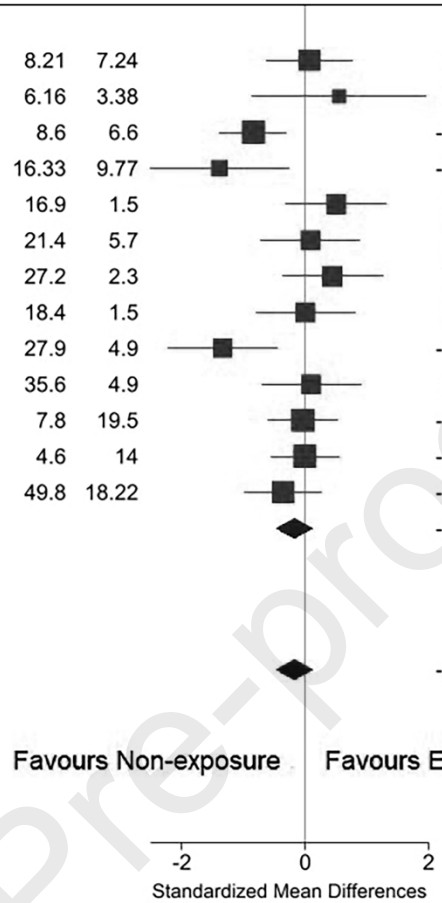
Study	Non Exposed N	Non Exposed Mean	Non Exposed SD	Exposed N	Exposed Mean	Exposed SD	Mean diff. with 95% CI	Weight (%)
Low or Some Concern								
Agarwal 2009-1	23	45.2	17.61	23	49.4	17.49	-4.20 [-14.34, 5.94]	1.19
Agarwal 2009-2	7	54.75	19.42	7	56.44	16.94	-1.69 [-20.78, 17.40]	0.36
Avendano 2012	29	13.6	5.6	29	24.5	7.6	-10.90 [-14.34, -7.46]	6.10
De Iuliis 2009*	4	14	4	24	57.67	26.85	-43.67 [-70.44, -16.90]	0.19
Falzone 2008-1	12	8.52	1.83	12	7.91	.91	0.61 [-0.55, 1.77]	11.62
Falzone 2008-2	12	8.22	.3	12	8.52	1.22	-0.30 [-1.01, 0.41]	12.52
Falzone 2008-3	12	16.89	2.44	12	16.29	1.22	0.60 [-0.94, 2.14]	10.65
Falzone 2008-4	12	4.64	.58	12	5.81	1.45	-1.17 [-2.05, -0.29]	12.20
Falzone 2008-5	12	5.22	.87	12	5.22	1.45	0.00 [-0.96, 0.96]	12.06
Falzone 2008-6	12	11.47	2.61	12	13.21	.58	-1.74 [-3.25, -0.23]	10.73
Nakatani-Enomoto 2016-1	25	26.6	19.5	25	26.6	19.5	0.00 [-10.81, 10.81]	1.06
Nakatani-Enomoto 2016-2	25	30	20.5	25	30	20.5	0.00 [-11.36, 11.36]	0.96
Nakatani-Enomoto 2016-3	25	31.8	20.5	25	30	17.5	1.80 [-8.77, 12.37]	1.10
Nakatani-Enomoto 2016-4	25	38.3	15	25	39.9	14	-1.60 [-9.64, 6.44]	1.79
Nakatani-Enomoto 2016-5	25	47.6	10	25	49.2	11	-1.60 [-7.43, 4.23]	3.04
Nakatani-Enomoto 2016-6	25	46.1	12	25	48	14	-1.90 [-9.13, 5.33]	2.15
Nakatani-Enomoto 2016-7	25	32.7	30.5	25	32.7	30.5	0.00 [-16.91, 16.91]	0.45
Nakatani-Enomoto 2016-8	25	34.7	21.5	25	36.6	22	-1.90 [-13.96, 10.16]	0.86
Nakatani-Enomoto 2016-9	25	37.4	31.5	25	36.2	25.5	1.20 [-14.69, 17.09]	0.51
Nakatani-Enomoto 2016-10	25	40.7	16.5	25	43.4	14.5	-2.70 [-11.31, 5.91]	1.59
Nakatani-Enomoto 2016-11	25	45	22.5	25	46.9	22.5	-1.90 [-14.37, 10.57]	0.81
Nakatani-Enomoto 2016-12	25	49.1	35	25	48.1	31.5	1.00 [-17.46, 19.46]	0.38
Veerachari 2012	20	47.85	8.91	20	53.8	7.95	-5.95 [-11.18, -0.72]	3.57
Heterogeneity: $\tau^2 = 2.11$, $I^2 = 64.73\%$, $H^2 = 2.84$							-1.37 [-2.46, -0.28]	
Test of $\theta_i = \theta_j$: Q(22) = 62.37, p = 0.00								
Test of $\theta = 0$: z = -2.47, p = 0.01								
High Concern								
Wang 2015	97	63.36	16.93	97	72.44	16.92	-9.08 [-13.84, -4.32]	4.08
Heterogeneity: $\tau^2 = 0.00$, $I^2 = .\%$, $H^2 = .$							-9.08 [-13.84, -4.32]	
Test of $\theta_i = \theta_j$: Q(0) = -0.00, p = .								
Test of $\theta = 0$: z = -3.74, p = 0.00								
Overall							-1.78 [-2.94, -0.62]	
Heterogeneity: $\tau^2 = 2.67$, $I^2 = 69.15\%$, $H^2 = 3.24$								
Test of $\theta_i = \theta_j$: Q(23) = 74.55, p = 0.00								
Test of $\theta = 0$: z = -3.01, p = 0.00								
Test of group differences: $Q_b(1) = 9.55$, p = 0.00								



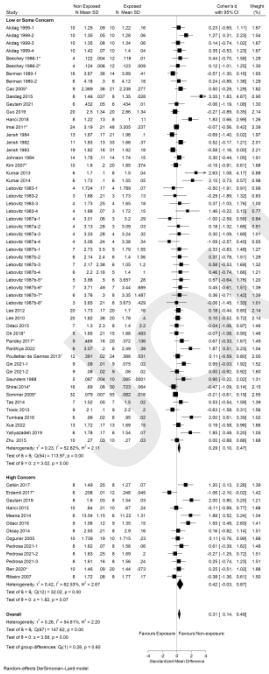
Random-effects DerSimonian–Laird model

Journal Pre-proofs

Study	Non-Exposed N	Non-Exposed Mean	Non-Exposed SD	Exposed N	Exposed Mean	Exposed SD	Cohen's d with 95% CI	Weight (%)
Low or Some Concern								
Agarwal 2009-1	16	8.66	6.45	16	8.21	7.24	0.07 [-0.63, 0.76]	8.48
Agarwal 2009-2	4	7.56	1.24	4	6.16	3.38	0.55 [-0.86, 1.96]	3.52
Avendano 2012	29	3.3	6	29	8.6	6.6	-0.84 [-1.38, -0.30]	10.33
De Iuliis 2009*	4	3.5	2	24	16.33	9.77	-1.39 [-2.51, -0.27]	4.93
Falzone 2010-1	12	18	2.7	12	16.9	1.5	0.50 [-0.31, 1.32]	7.26
Falzone 2010-2	12	21.8	3.4	12	21.4	5.7	0.09 [-0.72, 0.89]	7.38
Falzone 2010-3	12	28.3	2.7	12	27.2	2.3	0.44 [-0.37, 1.25]	7.29
Falzone 2010-4	12	18.4	1.9	12	18.4	1.5	0.00 [-0.80, 0.80]	7.38
Falzone 2010-5	12	22.6	2.7	12	27.9	4.9	-1.34 [-2.23, -0.45]	6.61
Falzone 2010-6	12	36	3.4	12	35.6	4.9	0.09 [-0.71, 0.90]	7.37
Nakatani-Enomoto 2016-1	25	7.1	19	25	7.8	19.5	-0.04 [-0.59, 0.52]	10.11
Nakatani-Enomoto 2016-2	25	4.5	15.5	25	4.6	14	-0.01 [-0.56, 0.55]	10.11
Veerachari 2012	20	43.55	16.76	20	49.8	18.22	-0.36 [-0.98, 0.27]	9.25
Heterogeneity: $\tau^2 = 0.15$, $I^2 = 52.72\%$, $H^2 = 2.12$							-0.17 [-0.48, 0.13]	
Test of $\theta_i = \theta_j$: Q(12) = 25.38, p = 0.01								
Test of $\theta = 0$: z = -1.13, p = 0.26								
Overall							-0.17 [-0.48, 0.13]	
Heterogeneity: $\tau^2 = 0.15$, $I^2 = 52.72\%$, $H^2 = 2.12$								
Test of $\theta_i = \theta_j$: Q(12) = 25.38, p = 0.01								
Test of $\theta = 0$: z = -1.13, p = 0.26								
Test of group differences: $Q_b(0) = 0.00$, p = .								

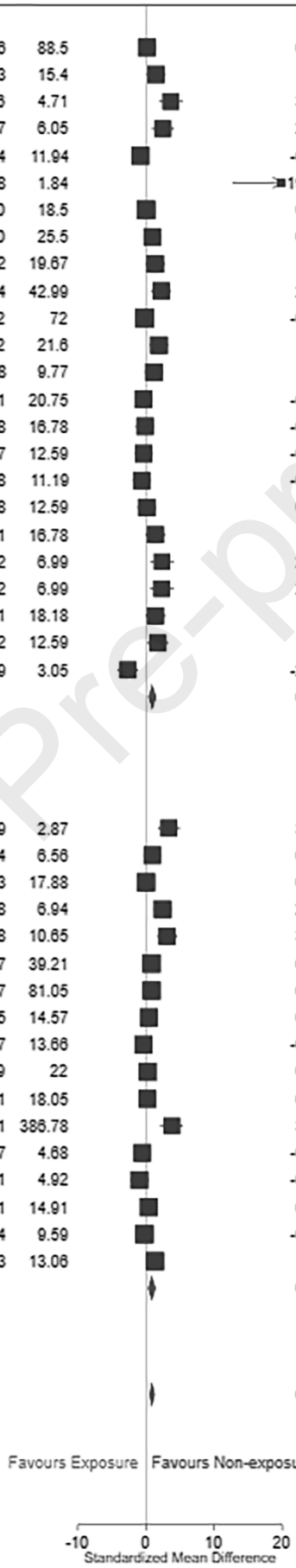


Random-effects DerSimonian–Laird model



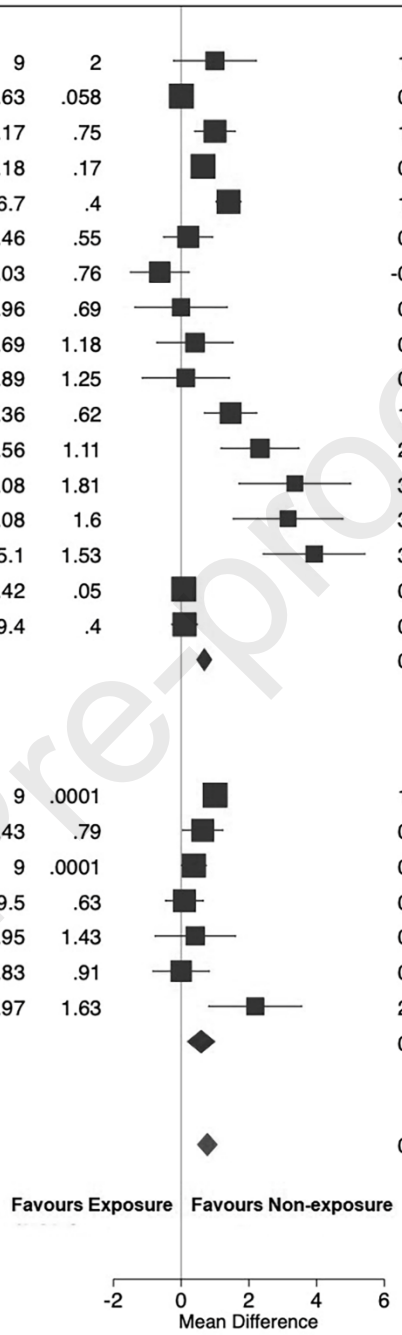
Journal Pre-proofs

Study	Non Exposed N	Non Exposed Mean	Non Exposed SD	Exposed N	Exposed Mean	Exposed SD	Cohen's d with 95% CI	Weight (%)
Low or Some Concern								
Almasiova 2021-1	6	236.77	97.98	6	225.16	88.5	0.12 [-1.01, 1.26]	2.56
Andraskova 2021-2	6	219.7	23.8	6	190.3	15.4	1.47 [0.19, 2.74]	2.43
Dasdag 2015	8	394.7	5.55	8	376	4.71	3.63 [2.04, 5.23]	2.14
Gautam 2021	6	288.89	6.05	6	274.07	6.05	2.45 [0.95, 3.95]	2.23
Guo 2019	8	253.5	11.94	8	263.4	11.94	-0.83 [-1.85, 0.19]	2.66
Hanci 2018	8	304.13	2.55	8	260.08	1.84	19.81 [12.88, 26.74]	0.31
Kim 2007*	10	250	20	20	250	18.5	0.00 [-0.76, 0.76]	2.87
Kumar 2013	6	199.7	16.6	6	180	25.5	0.92 [-0.27, 2.11]	2.51
Kumar 2014	6	181.49	15.28	6	157.62	19.67	1.36 [0.10, 2.61]	2.45
L'Abbate 1982	7	205	11.65	7	135.14	42.99	2.22 [0.89, 3.55]	2.38
Lee 2010	20	321	34	20	332	72	-0.20 [-0.82, 0.43]	2.97
Odaci 2015	7	295.2	11.2	8	262	21.6	1.89 [0.67, 3.11]	2.48
Pardhiya 2022	6	128.68	10.12	6	117.28	9.77	1.15 [-0.07, 2.37]	2.48
Saygin 2011	6	284.91	25.73	6	293.91	20.75	-0.39 [-1.53, 0.76]	2.55
Tang 2022-1	5	219.58	11.19	5	221.68	16.78	-0.15 [-1.39, 1.09]	2.46
Tang 2022-2	5	224.48	11.19	5	228.67	12.59	-0.35 [-1.60, 0.90]	2.45
Tang 2022-3	5	213.99	11.19	5	220.98	11.19	-0.62 [-1.89, 0.64]	2.44
Tang 2022-4	5	218.18	15.38	5	216.78	12.59	0.10 [-1.14, 1.34]	2.46
Tang 2022-5	5	206.99	11.19	5	187.41	16.78	1.37 [-0.00, 2.75]	2.34
Tang 2022-6	5	198.6	8.39	5	180.42	6.99	2.35 [0.74, 3.97]	2.13
Tang 2022-7	5	201.4	9.79	5	181.82	6.99	2.30 [0.70, 3.90]	2.14
Tang 2022-8	5	208.39	13.99	5	186.01	18.18	1.38 [0.00, 2.76]	2.34
Tang 2022-9	5	206.99	15.38	5	183.22	12.59	1.69 [0.25, 3.14]	2.28
Tas 2014	7	283.7	3.84	7	292.9	3.05	-2.65 [-4.09, -1.22]	2.28
Heterogeneity: $\tau^2 = 1.66$, $I^2 = 81.46\%$, $H^2 = 5.39$							0.90 [0.32, 1.49]	
Test of $\theta_i = \theta_j$: Q(23) = 124.05, p = 0.00								
Test of $\theta = 0$: z = 3.02, p = 0.00								
High Concern								
Cetkin 2017	8	292.53	2.64	8	283.29	2.87	3.35 [1.83, 4.87]	2.21
Delafarivar 2020	6	212.5	8.75	6	205.4	6.56	0.92 [-0.27, 2.11]	2.51
Gur 2021	8	292.63	17.88	8	292.63	17.88	0.00 [-0.98, 0.98]	2.69
Hanci 2013	10	110.03	6.06	10	94.28	6.94	2.42 [1.26, 3.57]	2.54
Odaci 2016	9	287.9	9.01	9	257.38	10.65	3.09 [1.72, 4.46]	2.34
Ozguner 2005	10	261.6	30.99	10	233.7	39.21	0.79 [-0.12, 1.70]	2.75
Pedrosa 2021-1	8	338.61	21.35	8	291.07	81.05	0.80 [-0.22, 1.82]	2.66
Pedrosa 2021-2	8	331.21	16.9	8	324.5	14.57	0.43 [-0.57, 1.42]	2.68
Pedrosa 2021-3	8	335.97	18.88	8	341.97	13.66	-0.36 [-1.35, 0.62]	2.69
Ribeiro 2007	8	383	11	8	379	22	0.23 [-0.75, 1.21]	2.69
Shahin 2019	6	207.52	25.56	6	203.01	18.05	0.20 [-0.93, 1.34]	2.56
Shahin 2018a*	5	2683.76	114.67	15	1384.61	386.78	3.76 [2.22, 5.31]	2.19
Simaiova 2019-1	6	30.36	3.9	6	32.77	4.68	-0.56 [-1.71, 0.59]	2.54
Simaiova 2019-2	6	30.46	4.54	6	35.1	4.92	-0.98 [-2.18, 0.22]	2.50
Yu 2020-1	7	66.04	9.59	12	60.71	14.91	0.40 [-0.54, 1.34]	2.73
Yu 2020-2	7	60.71	7.46	12	62.84	9.59	-0.24 [-1.17, 0.70]	2.73
Yu 2020-3	7	68.84	13.06	12	51.63	13.06	1.32 [0.30, 2.34]	2.66
Heterogeneity: $\tau^2 = 1.09$, $I^2 = 78.18\%$, $H^2 = 4.58$							0.84 [0.27, 1.40]	
Test of $\theta_i = \theta_j$: Q(16) = 73.33, p = 0.00								
Test of $\theta = 0$: z = 2.90, p = 0.00								
Overall							0.87 [0.46, 1.27]	
Heterogeneity: $\tau^2 = 1.35$, $I^2 = 79.77\%$, $H^2 = 4.94$								
Test of $\theta_i = \theta_j$: Q(40) = 197.77, p = 0.00								
Test of $\theta = 0$: z = 4.17, p = 0.00								
Test of group differences: $Q_b(1) = 0.03$, p = 0.87								



Random-effects DerSimonian–Laird model

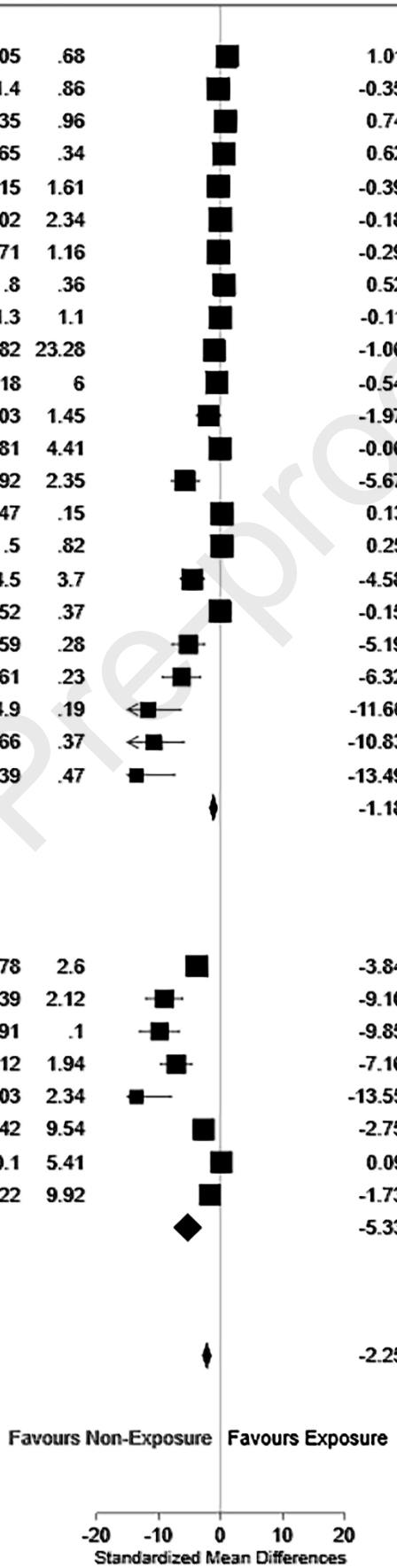
Study	Non Exposed N	Non Exposed Mean	Non Exposed SD	Exposed N	Exposed Mean	Exposed SD	Mean diff. with 95% CI	Weight (%)
Low or Some Concern								
Bilgici 2018	11	10	.5	11	9	2	1.00 [-0.22, 2.22]	3.05
Dasdag 2015	8	9.64	.052	8	9.63	.058	0.01 [-0.04, 0.06]	5.96
Hanci 2018	8	9.17	.41	8	8.17	.75	1.00 [0.41, 1.59]	4.87
Kim 2007*	10	9.83	.08	20	9.18	.17	0.65 [0.54, 0.76]	5.93
Odaci 2015	7	8.1	.3	8	6.7	.4	1.40 [1.04, 1.76]	5.51
Oh 2018*	3	9.67	.55	9	9.46	.55	0.21 [-0.51, 0.93]	4.48
Tang 2022-1	5	8.4	.62	5	9.03	.76	-0.63 [-1.49, 0.23]	4.05
Tang 2022-2	5	8.96	1.39	5	8.96	.69	0.00 [-1.36, 1.36]	2.72
Tang 2022-3	5	9.1	.49	5	8.69	1.18	0.41 [-0.71, 1.53]	3.30
Tang 2022-4	5	9.03	.76	5	8.89	1.25	0.14 [-1.14, 1.42]	2.90
Tang 2022-5	5	8.82	.62	5	7.36	.62	1.46 [0.69, 2.23]	4.33
Tang 2022-6	5	8.89	.69	5	6.56	1.11	2.33 [1.18, 3.48]	3.24
Tang 2022-7	5	9.44	.49	5	6.08	1.81	3.36 [1.72, 5.00]	2.18
Tang 2022-8	5	9.24	.91	5	6.08	1.6	3.16 [1.55, 4.77]	2.23
Tang 2022-9	5	9.03	.76	5	5.1	1.53	3.93 [2.43, 5.43]	2.44
Tas 2014	7	9.5	.04	7	9.42	.05	0.08 [0.03, 0.13]	5.97
Trosic 2013	9	9.5	.4	9	9.4	.4	0.10 [-0.27, 0.47]	5.49
Heterogeneity: $\tau^2 = 0.11$, $I^2 = 93.50\%$, $H^2 = 15.37$							0.69 [0.45, 0.92]	
Test of $\theta_i = \theta_j$: Q(16) = 245.98, p = 0.00								
High Concern								
Atasoy 2013	5	10	.01	5	9	.0001	1.00 [0.99, 1.01]	5.97
Cetkin 2017	8	9.07	.32	8	8.43	.79	0.64 [0.05, 1.23]	4.88
Dasdag 2003	8	9.375	.518	8	9	.0001	0.38 [0.02, 0.73]	5.52
Ozguner 2005	10	9.6	.63	10	9.5	.63	0.10 [-0.45, 0.65]	4.99
Yu 2020-1	7	8.37	.91	12	7.95	1.43	0.42 [-0.76, 1.60]	3.14
Yu 2020-2	7	8.83	.85	12	8.83	.91	0.00 [-0.83, 0.83]	4.14
Yu 2020-3	7	8.16	1.11	12	5.97	1.63	2.19 [0.82, 3.56]	2.71
Heterogeneity: $\tau^2 = 0.21$, $I^2 = 81.63\%$, $H^2 = 5.44$							0.59 [0.17, 1.01]	
Test of $\theta_i = \theta_j$: Q(6) = 32.67, p = 0.00								
Overall							0.77 [0.47, 1.08]	
Heterogeneity: $\tau^2 = 0.40$, $I^2 = 99.16\%$, $H^2 = 118.35$								
Test of $\theta_i = \theta_j$: Q(23) = 2721.97, p = 0.00								
Test of group differences: $Q_b(1) = 0.15$, p = 0.70								



Random-effects DerSimonian–Laird model

Journal Pre-proofs

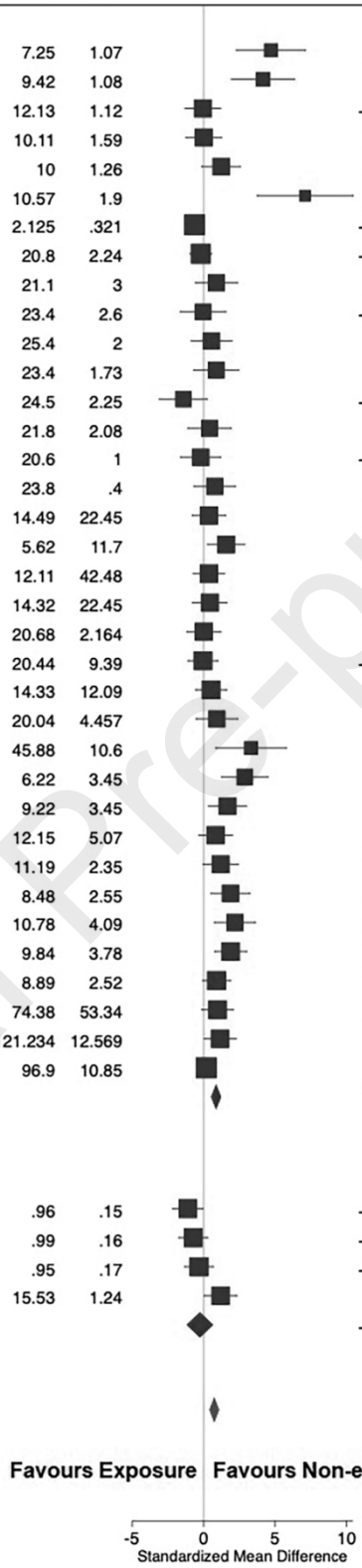
Study	Non-Exposed N	Non-Exposed Mean	Non-Exposed SD	Exposed N	Exposed Mean	Exposed SD	Cohen's d with 95% CI	Weight (%)
Low or Some Concern								
Cairnie 1981-1*	5	2.2	1.79	10	1.05	.68	1.01 [-0.13, 2.14]	3.79
Cairnie 1981-2*	5	1	1.57	10	1.4	.86	-0.35 [-1.44, 0.73]	3.82
Cairnie 1981-3*	5	2	.67	10	1.35	.96	0.74 [-0.37, 1.84]	3.81
Cairnie 1981-4*	5	1	.89	10	.65	.34	0.62 [-0.48, 1.71]	3.81
Cairnie 1981-5*	5	.6	.89	10	1.15	1.61	-0.39 [-1.47, 0.70]	3.82
Cairnie 1981-6*	12	1.62	1.18	60	2.02	2.34	-0.18 [-0.80, 0.44]	4.05
Cairnie 1981-7*	12	1.38	1.04	48	1.71	1.16	-0.29 [-0.92, 0.34]	4.04
Cairnie 1981-8	4	1.48	1.82	4	.8	.36	0.52 [-0.89, 1.93]	3.61
Dasdag 2008	7	1.2	.4	14	1.3	1.1	-0.11 [-1.01, 0.80]	3.92
Er 2022-1	6	7.95	5.42	6	25.82	23.28	-1.06 [-2.27, 0.15]	3.74
Er 2022-2	6	7.72	6.8	6	11.18	6	-0.54 [-1.69, 0.61]	3.78
Guo 2019	3	4.89	2.6	3	9.03	1.45	-1.97 [-3.92, -0.02]	3.21
Gur 2021	8	8.59	3.37	8	8.81	4.41	-0.06 [-1.04, 0.92]	3.88
Hanci 2018	8	11.84	.89	8	21.92	2.35	-5.67 [-7.87, -3.48]	3.03
Lee 2012	20	.49	.16	20	.47	.15	0.13 [-0.49, 0.75]	4.05
Lee 2010	20	.73	1.01	20	.5	.82	0.25 [-0.37, 0.87]	4.05
Odaci 2015	7	11	1.7	8	24.5	3.7	-4.58 [-6.50, -2.65]	3.23
Tang 2022-1	5	1.47	.28	5	1.52	.37	-0.15 [-1.39, 1.09]	3.72
Tang 2022-2	5	1.26	.23	5	2.59	.28	-5.19 [-7.78, -2.60]	2.73
Tang 2022-3	5	1.47	.42	5	3.61	.23	-6.32 [-9.35, -3.29]	2.42
Tang 2022-4	5	1.47	.37	5	4.9	.19	-11.66 [-16.92, -6.40]	1.31
Tang 2022-5	5	1.63	.14	5	4.66	.37	-10.83 [-15.74, -5.93]	1.44
Tang 2022-6	5	1.4	.23	5	6.39	.47	-13.49 [-19.53, -7.45]	1.08
Heterogeneity: $\tau^2 = 1.74$, $I^2 = 85.04\%$, $H^2 = 6.69$							-1.18 [-1.82, -0.54]	
Test of $\theta_i = \theta_j$: Q(22) = 147.08, p = 0.00								
Test of $\theta = 0$: z = -3.63, p = 0.00								
High Concern								
Gao 2016*	10	18.71	1.76	20	27.78	2.6	-3.84 [-5.07, -2.61]	3.73
Hanci 2013	10	3.42	.96	10	18.39	2.12	-9.10 [-12.05, -6.14]	2.48
Ma 2015	10	.21	.01	10	.91	.1	-9.85 [-13.03, -6.67]	2.33
Odaci 2016	9	3.35	.87	9	14.12	1.94	-7.16 [-9.68, -4.65]	2.79
Shahin 2019	6	2.34	1.08	6	27.03	2.34	-13.55 [-19.09, -8.01]	1.22
Shahin 2018a*	5	24.98	2.91	15	48.42	9.54	-2.75 [-4.07, -1.43]	3.67
Simaiova 2019-1	6	10.6	5.19	6	10.1	5.41	0.09 [-1.04, 1.23]	3.79
Simaiova 2019-2	6	8	4.22	6	21.22	9.92	-1.73 [-3.06, -0.41]	3.66
Heterogeneity: $\tau^2 = 9.30$, $I^2 = 92.45\%$, $H^2 = 13.24$							-5.33 [-7.62, -3.04]	
Test of $\theta_i = \theta_j$: Q(7) = 92.67, p = 0.00								
Test of $\theta = 0$: z = -4.56, p = 0.00								
Overall							-2.25 [-2.98, -1.52]	
Heterogeneity: $\tau^2 = 3.32$, $I^2 = 90.24\%$, $H^2 = 10.25$								
Test of $\theta_i = \theta_j$: Q(30) = 307.46, p = 0.00								
Test of $\theta = 0$: z = -6.06, p = 0.00								
Test of group differences: $Q_b(1) = 11.65$, p = 0.00								



Random-effects DerSimonian–Laird model

Journal Pre-proofs

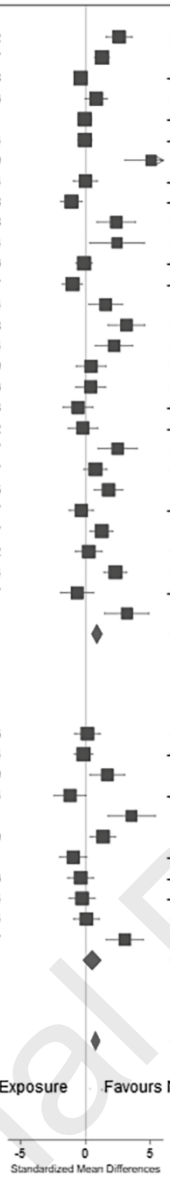
Study	Non Exposed N	Non Exposed Mean	Non Exposed SD	Exposed N	Exposed Mean	Exposed SD	Cohen's d with 95% CI	Weight (%)
Low or Some Concern								
Chen 2014 - 1	5	20	3.67	5	7.25	1.07	4.72 [2.31, 7.13]	1.31
Chen 2014 - 2	5	16.44	2.13	5	9.42	1.08	4.16 [1.95, 6.36]	1.47
Chen 2014 - 3	5	12.06	1.64	5	12.13	1.12	-0.05 [-1.29, 1.19]	2.59
Chen 2014 - 4	5	10.13	1.88	5	10.11	1.59	0.01 [-1.23, 1.25]	2.59
Chen 2014 - 5	5	12.56	2.67	5	10	1.26	1.23 [-0.12, 2.58]	2.43
Chen 2014 - 6	5	20.81	.75	5	10.57	1.9	7.09 [3.74, 10.43]	0.81
Imai 2011 *	24	1.92	.33	48	2.125	.321	-0.63 [-1.13, -0.13]	3.68
Johnson 1984	14	20.4	1.87	14	20.8	2.24	-0.19 [-0.94, 0.55]	3.35
Lebovitz 1983 - 1	4	23.2	1.4	4	21.1	3	0.90 [-0.56, 2.35]	2.29
Lebovitz 1983 - 2	3	23.3	2.94	3	23.4	2.6	-0.04 [-1.64, 1.56]	2.10
Lebovitz 1983 - 3	4	26.8	3	4	25.4	2	0.55 [-0.86, 1.96]	2.35
Lebovitz 1983 - 4	4	25.3	2.4	3	23.4	1.73	0.88 [-0.69, 2.45]	2.14
Lebovitz 1987a - 1	4	22.1	1.2	3	24.5	2.25	-1.41 [-3.08, 0.26]	2.02
Lebovitz 1987a - 2	4	22.9	3	3	21.8	2.08	0.41 [-1.10, 1.92]	2.21
Lebovitz 1987a - 3	4	20.4	1	4	20.6	1	-0.20 [-1.59, 1.19]	2.38
Lebovitz 1987a - 4	4	24.5	1.2	4	23.8	.4	0.78 [-0.66, 2.22]	2.31
Lebovitz 1987b - 1	7	20.32	8.33	5	14.49	22.45	0.37 [-0.78, 1.53]	2.72
Lebovitz 1987b - 2	6	19.46	4.2	6	5.62	11.7	1.57 [0.28, 2.87]	2.51
Lebovitz 1987b - 3	7	23.14	9.87	6	12.11	42.48	0.37 [-0.73, 1.47]	2.80
Lebovitz 1987b - 4	6	21.08	5.52	5	14.32	22.45	0.44 [-0.77, 1.64]	2.65
Lebovitz 1987b - 5 *	5	20.72	3.6	6	20.68	2.164	0.01 [-1.17, 1.20]	2.67
Lebovitz 1987b - 6 *	7	20.06	7.77	7	20.44	9.39	-0.04 [-1.09, 1.00]	2.88
Lebovitz 1987b - 7 *	6	22.28	18.66	8	14.33	12.09	0.52 [-0.55, 1.60]	2.84
Lebovitz 1987b - 8 *	3	23.83	3	6	20.04	4.457	0.93 [-0.52, 2.38]	2.30
Pandey 2018	3	72.47	4.07	3	45.88	10.6	3.31 [0.85, 5.78]	1.27
Qin 2014 - 1	6	17.37	4.26	6	6.22	3.45	2.88 [1.26, 4.49]	2.08
Qin 2014 - 2	6	16.04	4.63	6	9.22	3.45	1.67 [0.36, 2.98]	2.49
Qin 2014 - 3	6	16.19	4.53	6	12.15	5.07	0.84 [-0.34, 2.02]	2.68
Qin 2014 - 4	6	15.63	4.73	6	11.19	2.35	1.19 [-0.04, 2.42]	2.61
Qin 2014 - 5	6	16.26	5.27	6	8.48	2.55	1.88 [0.52, 3.24]	2.42
Qin 2014 - 6	6	19.44	3.82	6	10.78	4.09	2.19 [0.76, 3.62]	2.32
Qin 2021 - 1	9	20.05	6.63	9	9.84	3.78	1.89 [0.78, 3.00]	2.79
Qin 2021 - 2	9	11.47	3.15	9	8.89	2.52	0.90 [-0.07, 1.87]	3.00
Saunders 1981a *	4	123.15	4.41	24	74.38	53.34	0.97 [-0.12, 2.06]	2.82
Saunders 1981b *	4	135.23	8.7	20	121.234	12.569	1.16 [0.03, 2.28]	2.77
Sommer 2009 *	32	98.9	11.1	94	96.9	10.85	0.18 [-0.22, 0.59]	3.79
Heterogeneity: $\tau^2 = 0.71$, $I^2 = 69.01\%$, $H^2 = 3.23$							0.87 [0.51, 1.22]	
Test of $\theta_i = \theta_j$: Q(35) = 112.93, p = 0.00								
High Concern								
Pedrosa 2021 - 1	8	.8	.14	8	.96	.15	-1.10 [-2.15, -0.05]	2.88
Pedrosa 2021 - 2	8	.89	.11	8	.99	.16	-0.73 [-1.74, 0.28]	2.94
Pedrosa 2021 - 3	8	.9	.14	8	.95	.17	-0.32 [-1.31, 0.67]	2.98
Xu 2020	7	16.86	.99	7	15.53	1.24	1.19 [0.05, 2.32]	2.75
Heterogeneity: $\tau^2 = 0.61$, $I^2 = 68.41\%$, $H^2 = 3.17$							-0.26 [-1.19, 0.67]	
Test of $\theta_i = \theta_j$: Q(3) = 9.50, p = 0.02								
Overall							0.74 [0.41, 1.08]	
Heterogeneity: $\tau^2 = 0.74$, $I^2 = 70.11\%$, $H^2 = 3.35$								
Test of $\theta_i = \theta_j$: Q(39) = 130.48, p = 0.00								
Test of group differences: $Q_b(1) = 4.98$, p = 0.03								



Random-effects DerSimonian–Laird model

Journal Pre-proofs

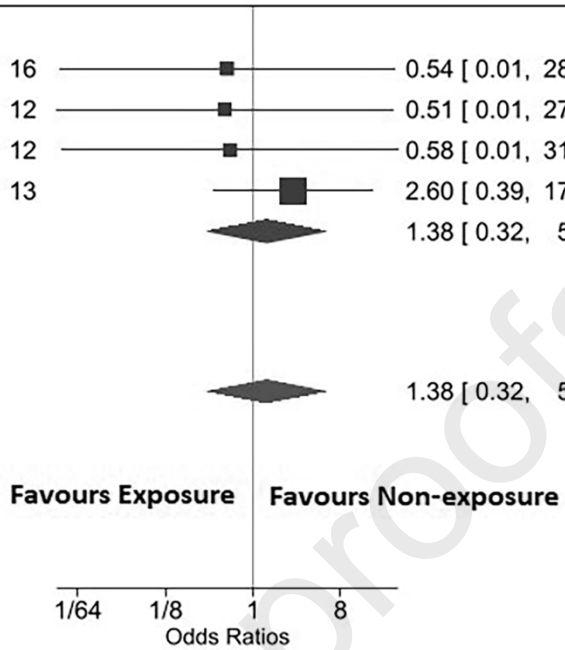
Study	Non-Exposed N	Non-Exposed Mean	Non-Exposed SD	Exposed N	Exposed Mean	Exposed SD	Cohen's d with 95% CI	Weight (%)
Low or Some Concern								
Azimzadeh 2019*	10	3557	696.98	20	2304	340.12	2.59 [1.58, 3.59]	2.60
Chen 2014	30	29.21	2.85	30	25.72	2.67	1.26 [0.71, 1.82]	3.00
Forgacs 2006	52	5.12	5.7	56	7.85	8.08	-0.39 [-0.77, -0.01]	3.11
Guo 2019	12	14.44	1.77	12	12.74	2.36	0.81 [-0.02, 1.65]	2.76
Jin 2013-1*	20	12.39	2.6	40	12.58	2.81	-0.07 [-0.61, 0.47]	3.01
Jin 2013-2*	20	13.45	2.6	40	13.56	2.24	-0.05 [-0.58, 0.49]	3.01
Jonwal 2018	8	4.02	.31	8	1.95	.49	5.05 [3.04, 7.05]	1.65
Khillare 1998	9	6.5	16.8	9	6.7	14.4	-0.01 [-0.94, 0.91]	2.68
Kim 2007*	10	1.1	.5	20	2.2	1.18	-1.09 [-1.89, -0.28]	2.79
Kumar 2013	6	4.1	1.4	6	1.4	.8	2.37 [0.89, 3.84]	2.12
Kumar 2011	3	5.41	2.17	3	1.36	.94	2.42 [0.32, 4.53]	1.57
Lee 2012	20	1.31	.694	20	1.39	.666	-0.12 [-0.74, 0.50]	2.95
Ozlem Nisbet 2012*	11	5.73	.46	22	6	.07	-1.01 [-1.77, -0.25]	2.83
Pardhiya 2022	6	5.478	1.47	6	3.583	.894	1.56 [0.27, 2.85]	2.31
Qin 2018*	6	48.69	2.54	12	35.42	4.78	3.15 [1.73, 4.57]	2.18
Qin 2014-1	6	41.96	8.43	6	21.24	10.34	2.20 [0.76, 3.63]	2.16
Qin 2014-2	6	32.58	5.94	6	29.84	6.9	0.43 [-0.72, 1.57]	2.46
Qin 2014-3	6	31.09	6.51	6	28.74	5.36	0.39 [-0.75, 1.54]	2.46
Qin 2014-4	6	23.66	6.13	6	28.74	10.73	-0.58 [-1.74, 0.57]	2.44
Qin 2014-5	6	25.54	6.51	6	27.18	9.2	-0.21 [-1.34, 0.93]	2.47
Qin 2014-6	6	41.57	7.28	6	16.86	12.07	2.48 [0.97, 3.98]	2.09
Qin 2012	6	11.41	2.71	36	9.32	2.77	0.76 [-0.12, 1.64]	2.72
Qin 2021-1	9	9.9	3.27	9	5.31	1.65	1.77 [0.68, 2.86]	2.51
Qin 2021-2	9	6.3	2.52	9	7.05	2.07	-0.33 [-1.26, 0.60]	2.67
Saygin 2016	12	.99	.45	12	.59	.07	1.24 [0.37, 2.12]	2.73
Sepehrimanesh 2014*	5	3.06	2.49	15	2.6	1.72	0.24 [-0.78, 1.25]	2.59
Shahin 2018a*	10	5.77	1.58	30	2.88	1.14	2.30 [1.42, 3.17]	2.72
Xue 2022	5	.0233	.0018	5	.0252	.0037	-0.65 [-1.93, 0.62]	2.33
Yahyazadeh 2019	6	.438	.042	6	.271	.061	3.19 [1.48, 4.89]	1.90
Heterogeneity: τ^2 = 1.12, I^2 = 85.23%, H^2 = 6.77							0.87 [0.43, 1.30]	
Test of $\theta_i = \theta_j$: Q(28) = 189.56, p = 0.00								
Test of θ = 0: z = 3.91, p = 0.00								
High Concern								
Cetkin 2017	8	2.57	.43	8	2.49	.65	0.15 [-0.84, 1.13]	2.62
Fahim 1975*	10	7.04	.95	40	7.24	1.34	-0.16 [-0.85, 0.54]	2.89
Kesari 2012	6	5.05	2.11	6	1.76	1.79	1.68 [0.37, 3.00]	2.28
Kismali 2009	6	1.71	.64	6	3.01	1.4	-1.19 [-2.42, 0.03]	2.37
Meena 2014	6	4.41	.52	6	2.75	.41	3.55 [1.73, 5.36]	1.81
Ozguner 2005	10	2.28	.73	10	1.53	.29	1.35 [0.38, 2.32]	2.63
Pedrosa 2021-1	8	1.1	.83	8	1.79	.61	-0.95 [-1.98, 0.09]	2.57
Pedrosa 2021-2	8	.89	.73	8	1.19	.86	-0.38 [-1.38, 0.61]	2.61
Pedrosa 2021-3	8	.83	.66	8	1.05	.94	-0.27 [-1.26, 0.71]	2.62
Ribeiro 2007	8	2.61	1.46	8	2.54	.64	0.06 [-0.92, 1.04]	2.62
Shahin 2019	8	3.3	.65	8	1.7	.37	3.03 [1.59, 4.46]	2.16
Heterogeneity: τ^2 = 1.18, I^2 = 80.43%, H^2 = 5.11							0.50 [-0.22, 1.23]	
Test of $\theta_i = \theta_j$: Q(10) = 51.10, p = 0.00								
Test of θ = 0: z = 1.35, p = 0.18								
Overall							0.76 [0.40, 1.13]	
Heterogeneity: τ^2 = 1.09, I^2 = 83.85%, H^2 = 6.19								
Test of $\theta_i = \theta_j$: Q(39) = 241.46, p = 0.00								
Test of θ = 0: z = 4.07, p = 0.00								
Test of group differences: $Q_b(1)$ = 0.72, p = 0.40								



Random-effects DerSimonian–Laird model

Journal Pre-proofs

Study	Exposed N c	Exposed N non-c	Non-Exposed N c	Non-Exposed N non-c	Odds ratio with 95% CI	Weight (%)
Low or Some Concern						
Shirai 2014*	0	30	0	16	0.54 [0.01, 28.53]	13.51
Shirai 2017*	0	24	0	12	0.51 [0.01, 27.27]	13.42
Takahashi 2010*	0	21	0	12	0.58 [0.01, 31.16]	13.40
Yu 2020	4	10	2	13	2.60 [0.39, 17.16]	59.67
Heterogeneity: $\tau^2 = 0.00$, $I^2 = 0.00\%$, $H^2 = 1.00$					1.38 [0.32, 5.94]	
Test of $\theta_i = \theta_j$: Q(3) = 1.07, p = 0.78						
Test of $\theta = 0$: z = 0.44, p = 0.66						
Overall					1.38 [0.32, 5.94]	
Heterogeneity: $\tau^2 = 0.00$, $I^2 = 0.00\%$, $H^2 = 1.00$						
Test of $\theta_i = \theta_j$: Q(3) = 1.07, p = 0.78						
Test of $\theta = 0$: z = 0.44, p = 0.66						
Test of group differences: $Q_b(0) = 0.00$, p = .						



Random-effects DerSimonian–Laird model
Sorted by: _meta_id

Journal Pre-proofs