Environment International 157 (2021) 106806

Contents lists available at ScienceDirect

Environment International

journal homepage: www.elsevier.com/locate/envint

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*

Francesca Pacchierotti [a,*], Lucia Ardoino [a], Barbara Benassi [a], Claudia Consales [a], Eugenia Cordelli [a], Patrizia Eleuteri [a], Carmela Marino [a], Maurizio Sciortino [b], Martin H. Brinkworth [c], Guangdi Chen [d], James P. McNamee [e], Andrew William Wood [f], Carlijn R. Hooijmans [g], Rob B.M. de Vries [g]

[a] *Division Health Protection Technologies, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Rome, Italy*
[b] *Section Technology Transfer to Developing Countries Relating to Climate Change, 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] *SYRCLE, Department for Health Evidence, Radboud Institute for Health Sciences, Radboudumc, Nijmegen, the Netherlands*

ARTICLE INFO

Handling Editor: Paul Whaley

Keywords:
Radiofrequency electromagnetic fields
Male infertility
Adverse pregnancy outcomes
Animal studies
Human sperm
Systematic review

ABSTRACT

Background: Radiofrequency Electromagnetic Fields (RF-EMF) at environmental level have been reported to induce adverse effects on the male reproductive system and developing embryos. However, despite the number of experiments conducted since the 1970s, the diversity of testing approaches and exposure conditions, inconsistencies among results, and dosimetric flaws have not yet permitted a solid assessment of the relationship between RF-EMF exposure and such effects, warranting a more systematic and methodologically rigorous approach to the evaluation of available data.

Objectives: This study aims at evaluating the effects of RF-EMF exposure on male fertility and pregnancy outcomes by a systematic review (SR) of experimental studies, conducted in compliance with international guidelines. The evidence will be organized into three streams: 1) Studies evaluating the impact of RF-EMF on the male reproductive system of experimental mammals; 2) studies evaluating the impact of RF-EMF on human sperm exposed *in vitro*; 3) studies evaluating the impact of RF-EMF on adverse pregnancy, birth outcomes and delayed effects in experimental mammals exposed *in utero*.

Study eligibility and criteria: Eligible studies will include peer-reviewed articles reporting of original results about effects of controlled exposures to RF-EMF in the frequency range 100 kHz–300 GHz on the selected outcomes without any language or year-of-publication restrictions. Eligible studies will be retrieved by calibrated search strings applied to three electronic databases, PubMed, Scopus and EMF Portal and by manual search of the list of references of included papers and published reviews.

Study appraisal and synthesis method: The internal validity of the studies will be evaluated using the Risk of Bias (RoB) Rating Tool developed by National Toxicology Program/Office of Health Assessment and Translation (NTP/OHAT) integrated with input from the SYRCLE RoB tool. Given sufficient commensurate data, *meta*-analyses will be performed, otherwise narrative syntheses will be produced. Finally, the certainty of the effects of RF-EMF exposure on male fertility and pregnancy and birth outcomes will be established following GRADE.

Funding: The study is financially supported by the World Health Organization.

* Corresponding author.
E-mail address: francesca.pacchierotti@enea.it (F. Pacchierotti).

https://doi.org/10.1016/j.envint.2021.106806
Received 25 February 2021; Received in revised form 9 July 2021; Accepted 27 July 2021
Available online 26 August 2021

Registration: OSF Registration DOI https://doi.org/10.17605/OSF.IO/7MUS3; PROSPERO CRD42021227729, CRD42021227746.

1. Introduction

1.1. Background

The technological applications of radiofrequency electromagnetic fields (RF-EMF; frequencies 100 kHz to 300 GHz) have been steadily increasing since the 1950s. RF-EMF are used in medicine (e.g. magnetic resonance imaging, diathermy, radiofrequency ablation), industry (e.g. heating and welding), domestic appliances (e.g. baby monitor, Wi-Fi), security and navigation (e.g. radar and RFID) and especially in telecommunications (e.g. radio and TV broadcasting, mobile telephony). These developments mean that large parts of the global population are now exposed to RF-EMF and more will be exposed in the future. Concern has been raised regarding public health consequences from RF-EMF exposure, and it is therefore crucial to perform a health risk assessment to support decision-makers and inform the general public.

The World Health Organization (WHO) has an ongoing project to assess potential health effects of exposure to RF-EMF in the general and working population. To prioritize potential adverse health outcomes from exposure to these fields, WHO conducted a broad international survey amongst RF experts in 2018 (Verbeek et al. 2021). Six major topics were identified (cancer, adverse reproductive outcomes, cognitive impairment, symptoms, oxidative stress, and heat related effects) for which the WHO has commissioned systematic reviews (SR) of observational and experimental studies to analyse and synthesize the available evidence. In the current paper, we present the protocol for a SR of exposure to RF fields and adverse reproductive outcomes in animal and *in vitro* experimental studies. A separate SR will be conducted of observational studies on RF-EMF exposure and adverse reproductive outcomes, in addition to systematic reviews of experimental and observational studies dealing with the other adverse outcomes, the protocols of which will be similarly submitted for publication in Environment International.

1.2. Description of the outcome

Both male infertility and adverse pregnancy and birth outcomes will be considered as adverse reproductive outcomes.

Infertility and subfertility, affecting a significant proportion of humanity, are considered by the WHO global public health issues (http://www.who.int/reproductivehealth/topics/infertility/perspective/en/). Infertility, defined as the inability to conceive after one year of unprotected intercourse, affects between 15 and 25% of couples worldwide. It has been estimated that the male contribution is totally or partly the cause of infertility in 20–70% of couples (Agarwal et al. 2015). A variety of conditions can affect the male reproductive system to different extents and they often coexist so that the precise etiology of male factor infertility remains undefined in 30–50% of patients. Spermatogenesis, the long and complex process leading to mature spermatozoa may be affected, along the course of life, by a variety of genetic, lifestyle and environmental factors (Gabrielsen and Tanrikut 2016; Xavier et al. 2021). Alteration of spermatogenesis can lead to defects of mature spermatozoa such as decreased concentration, motility and morphology. These endpoints are collectively analyzed, according to WHO guidelines, as sperm quality biomarkers and are used as indicators of the fertility potential of an individual.

Moreover, DNA and chromatin alterations have been suggested as possible causes of male infertility or transmissible genetic mutations that may be lethal during embryo/fetal development. A decrease in semen quality over the past decades has been reported (Levine et al. 2017) especially in the most industrialized countries, and even if the issue is controversial, the possibility that environmental exposures might, in the future, induce a decline in male fertility has not been discounted.

Prenatal development is a finely regulated process that is highly sensitive to chemical and physical stressors. Depending on the period of exposure, different outcomes may be expected: arrest of development and embryo/fetal death when exposure occurs in the early post-implantation phase, and reduced growth, preterm births, malformations or neonatal deaths when exposure occurs later. Spontaneous miscarriage is the most common complication of human pregnancy; it is estimated that at least 30% of all pregnancies and 10–15% of clinically recognized pregnancies end in miscarriage (Wilcox et al. 1988). Preterm births account for around 5% of births in high-income countries and 25% in low-middle income countries (Steer, 2005). According to the WHO, each year approximately 3.2 million children worldwide are born with a congenital anomaly (CA) and approximately 300,000 newborns with a diagnosis of birth defect die within the first 28 days of life. In Europe, CAs are the leading cause of perinatal mortality: the European Surveillance of Congenital Anomalies (EUROCAT) network estimated a perinatal mortality associated with CAs of 9.2 per 10,000 births in 2008–2012. The causes of many CAs are complex and multifactorial, but in most cases their etiology remains unknown. According to the WHO, approximately 5% of CAs are attributable to environmental exposures (Baldacci et al. 2018). More recently, increasing comprehension of the molecular mechanisms regulating cell differentiation and gene expression and lines of evidence from experimental models and human surveys have led to the knowledge that alterations of prenatal development may have consequences during adulthood, especially, but not exclusively, regarding behavioural traits.

1.3. Description of the exposure

Radiofrequency electromagnetic fields (RF-EMF) are defined as fields with frequencies from 100 kHz to 300 GHz. Human exposure to RF-EMF may occur in the general living environment as well as in the workplace from sources close to or far away from the body, resulting in localized near-field and whole-body far-field exposure conditions, respectively (near-field refers to distances from the RF-EMF source less than a few wavelengths, for example, approximately 1 m at 1 GHz).

Mobile phones, computers, laptops and tablets used with Wireless Local Area Network (WLAN) are the main source of near-field exposures, which are potentially hazardous for the male reproductive system and the embryo-fetal development *in-utero*. One relevant concern about these sources and the consequent exposures, is the fact that they are mainly indoor sources and accounting for reflections could be critical for the field evaluation. The most common sources of far-field exposure are radio and television antennas, mobile phone base stations and the base station antennas of any other wireless communication system (Digital Enhanced Cordless Phone (DECT), portable professional radio-transmitters), WLAN, or WiFi access points. Many more sources of exposure are present in the living environment, from baby monitors to remote control devices or to RF security systems. Sources of exposure in occupational settings may be RF polyvinyl chloride welding machines and radar systems.

In addition to the distance from the field, the main variables influencing the interaction of RF-EMF with the human body are the signal frequency (the higher the frequency, the lower the penetration depth), the exposure intensity (defined as the strength of the incident electric and magnetic fields) and the exposure duration. However, polarization

of the field emitted by the source, modulation of the signal and dielectric characteristics of tissues also play a role.

Exposure to RF-EMF results in induction of internal electric fields and absorbed energy. The most relevant quantities used to link the exposure to the possible adverse health effects are the internal electric field strength and the rate of RF energy absorption (Specific Absorption Rate, SAR, W/kg tissue weight) depending on the frequency. In the range 10 MHz – 6 GHz, the most relevant exposure metric is the SAR. The SAR depends on the RF signal characteristics of frequency, field strength and polarization, and on the characteristics of the absorbing tissue (dielectric parameters) and it results in a direct proportionality with the (squared) internal electric field.

Depending on the exposure level and time, energy absorption can result in a tissue temperature increase when thermoregulatory responses are not sufficient, or in changes in energy balance in the exposed target, which may induce heat-related biological effects.

Since measurements of SAR or internal electric field strength are usually difficult to perform, the levels for maximum safe human exposure to RF fields are defined in terms of external electric- and magnetic-field strength or power density (measurable quantities). These are evaluated using dosimetric methods that determine the levels of these external field strengths that would produce the maximum internal electric field or SAR within the body. The International Commission on Non-Ionizing Radiation Protection (ICNIRP) provides limits for both quantities, referring to the external field strength limits as "reference levels" and to the internal field limits as "basic restrictions".

For experimental studies, dosimetry and evaluation of the internal electric field (and the relative SAR) are especially relevant because the exposure generally takes place in near-field conditions, and are necessary to link any biological effect to a value of induced electric field and not to the used exposure system. The SAR can be numerically and experimentally calculated with different methods (numerical and experimental dosimetry). In experimental dosimetry, the SAR can be determined by means of different techniques and probes according to the different relationships between the SAR and the internal electric field or the temperature; usually the choice of the method depends on the characteristics of the sample and on the exposure scenario (exposure system). To determine the SAR distribution within inhomogeneous and morphologically irregular objects, it is necessary to resort to numerical dosimetry which makes use of different types of calculation codes.

At high frequencies, above 6 GHz, the absorption is limited to the surface of the exposed target and the absorbed Power Density (expressed in W/m^2) can be used as an appropriate tissue internal exposure metric.

1.4. Rationale for a systematic review

Over the years, several human biomonitoring studies and laboratory experiments have been conducted to investigate the possibility of an association between RF-EMF exposure and male fertility impairment. These studies were reviewed in papers that pointed to findings supporting an association, but at the same time acknowledged the heterogeneity and limitations of study designs and the variability of results (Deepinder et al. 2007; Agarwal et al. 2011; Kesari et al. 2018; Vornoli et al. 2019). The difficulties of designing the exposure set up, monitoring the exposure level and establishing the appropriate controls were among the reasons undermining a firm conclusion about an effect of RF-EMF on male fertility. In addition, the lack of molecular and cellular mechanisms supporting a causal relationship has been pointed out, and despite some attractive hypotheses (Houston et al. 2016; Santini et al. 2018; Jaffar et al. 2019) this knowledge gap has not been filled so far. Based on a bibliometric analysis of the published literature, recommendation towards more standardized models and improvement of the research quality and of the information exchange within the scientific community have been put forward (Bernabò et al., 2017).

In light of the variability of results, a more rigorous approach to literature review was undertaken applying systematic review and *meta*-analysis methodologies. Meta-analyses of human observational and *in vitro* experimental studies on the quality of sperm exposed to RF-EMF (Dama et al. 2013; Adams et al. 2014) inferred associations between exposure and variations of specific parameters, in particular sperm viability and motility, but the Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR) questioned this conclusion on the basis of dosimetric shortcomings of the studies (SCENIHR 2015).

A SR and *meta*-analysis of epidemiological, *in vitro* and animal studies on RF-EMF exposure and sperm quality (Liu et al. 2014) supported a detrimental effect of RF-EMF on sperm motility and viability/concentration based on *in vitro* and animal studies, whereas no significant mean difference was observed in the pooled analysis of human observational studies. However, limitations of the analysis were outlined, such as small number of included studies, potential publication bias for studies reporting positive results, large heterogeneity among different studies, and recall bias for human studies. The authors pointed out that the limited number of human studies, due to the difficulty of their organization and performance, did not allow to fully describe the actual data in populations, even though such studies would provide the best evidence to assess RF-EMF effects.

Finally, a more recent SR of the literature on 2.45 GHz RF-EMF exposure and male reproductive effects, which analyzed 1 retrospective cohort study, 4 *in vitro* studies on human semen and 18 studies in experimental animals, all reporting SAR values below the limits recommended by the ICNIRP, concluded that "exposure towards 2.45 GHz RF-EMR emitted by Wi-Fi transmitter is hazardous on the male reproductive system" (Jaffar et al. 2019).

Several experimental studies have also been carried out since the 1980s to investigate the impact of RF-EMF on adverse pregnancy outcomes and congenital disorders, but, until now, these studies have not been systematically reviewed. Teratogenic effects were reported in both non-mammalian and mammalian experimental models after RF-EMF exposure, but only at levels that produce internal temperature rises, which only occurred at exposures well above the recommended limits. No sound experimental evidence was found for non-thermal RF-EMF teratogenic effects (Heynick and Merritt 2003). A SR of the literature published between 1990 and 2010 on adverse pregnancy outcomes associated with physiotherapists' occupational exposure to RF-EMF concluded that the association was inconsistent and expressed the need for further, preferably prospective studies (Shah and Farrow 2014).

Limitations of study designs, inconsistencies among results and lack of a solid mechanistic interpretation of positive findings are among the reasons why national and international advisory boards, so far, have not supported an association between low-level (non-thermal) RF-EMF exposure and adverse effects on reproduction and development. In 2012, the independent Advisory Group on Non-Ionizing Radiation (AGNIR) reporting to the UK Health Protection Agency (HPA) (AGNIR 2012) concluded that: "A substantial number of studies have investigated the effects of RF fields on testicular function, principally in rats, and most report large, obvious effects. However, these results are largely uninterpretable due to inadequate dosimetry or other shortcomings in the studies, and thus are unsuitable for the purposes of health risk assessment." Regarding developmental effects, the conclusion was that: "Six well-conducted studies using rodents exposed to a variety of signals associated with mobile telecommunications systems indicate an absence of any consistent teratological effect following prenatal exposures that did not induce maternal hyperthermia". Similarly, in its most recent overall assessment of the evidence on potential health effects of exposure to electromagnetic fields, the SCENIHR expressed the opinion that "there is strong overall weight of evidence against an effect of low level RF fields on reproduction or development" (SCENIHR 2015). Finally, in 2020 the ICNIRP in the update of its guidelines for limiting exposure to RF-EMF in the frequency range 100 kHz–300 GHz (ICNIRP 2020) concluded that "no adverse effects of radiofrequency EMF exposure on fertility, reproduction or development relevant to human health have been substantiated".

Unfortunately, it is clear that in spite of the efforts undertaken to globally assess the body of evidence, an international consensus on the association between RF exposure and adverse reproductive outcomes has not been reached so far. Previously published systematic reviews of animal and *in vitro* studies dated back to 2014 (Dama et al. 2013; Adams et al. 2014; Liu et al. 2014) and could not take into account the most recent publications, or suffered from some methodological limitation such as the lack of a Risk of Bias analysis (Jaffar et al. 2019). Moreover, to our knowledge, no systematic review on adverse pregnancy outcomes has been published. Thus, a new systematic review and a *meta*-analysis of all the available experimental *in vitro* and animal studies on RF-EMF and adverse reproductive outcomes, carried out according to the most rigorous, internationally acknowledged and transparent methodology, are warranted.

2. Objectives

Through a SR and *meta*-analysis of animal studies and human sperm *in vitro* studies we aim to assess the experimental evidence about possible effects of RF-EMF exposure on male fertility and pregnancy and birth outcomes. A further aim is to possibly assess whether part of the effects is due to a RF-EMF-induced temperature increase. This will be done by comparing effects between exposed and sham-exposed groups, under conditions in which RF-EMF exposure does not induce an increase of temperature, or by comparing RF-EMF induced effects with effects induced in a comparator group in which an equal temperature increase was induced in the absence of RF-EMF exposure.

The final goal is to answer the following PECO 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*?
3) What is the effect of exposure to RF-EMF *in utero* on pregnancy outcomes, congenital disorders or delayed effects compared to either a sham control or an equal-temperature control in non-human mammals?

Many experimental studies investigate outcomes that are surrogate markers of the ultimate effect, especially in the case of fertility (e.g. testicular toxicity or sperm quality). To capture the variety of the literature in the field, multiple outcomes will concur to assess the impact of RF-EMF on fertility and pregnancy. They have been selected based on their association with the ultimate effect. In the synthesis of results, the strength of this association will be considered to attribute the relative importance to each stream of evidence.

3. Methods

The SR will be carried out following the recommendations for the conduct of systematic reviews in preclinical toxicology and environmental health research (Rooney et al. 2014; Hooijmans et al. 2014a; Hooijmans et al. 2018; Whaley et al. 2020) and the guidelines of the WHO (WHO 2014) and of the National Toxicology Program/Office of Health Assessment and Translation (NTP/OHAT) Handbook (NTP, 2015a). It will be reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021) (Supplementary File 1). Due to the complementarity of the PECO questions and the considerable overlapping methodologies, the three protocols are presented together in this paper to provide the reader with an overall comprehension of the on-going undertaking. In the following text the 3 protocols will be identified as Protocol 1 (P1), Protocol 2 (P2) and Protocol 3 (P3), each one corresponding, in the same order, to the 3 PECO questions listed above. Tools and criteria for data extraction, Risk of Bias (RoB) assessment and rating of the overall evidence will be adapted to the specificity of animal (P1 and P3) and *in vitro* (P2) studies.

3.1. Eligibility criteria

Studies will be included if they comply to a set of pre-defined criteria for each element of the specific PECO question (Table 1).

Each paper may contain several studies that investigated different outcomes for the same exposure condition, and each outcome may have been evaluated by different endpoints and methods of analysis. Some endpoints and some methods of analysis are deemed invalid because not predictive of the outcome or not sufficiently standardized. Data relative to these endpoints and methods will not be included in the systematic review. Studies investigating both valid and invalid endpoints, or endpoints measured with some invalid and some valid methods will be included, and only data measured on valid endpoints measured by valid methods will be included in the systematic review.

3.1.1. Types of populations

Protocol 1. The population of interest will comprise male subjects of any experimental *in vivo* non-human mammalian models.

Protocol 2. Human sperm exposed to RF-EMF *in vitro*.

Protocol 3. The population of interest will comprise subjects of any species of experimental non-human mammals exposed exclusively *in utero* and analyzed at the embryonic/fetal stage, at birth or during adulthood.

3.1.2. Types of exposures

We will include all identified studies that have applied electric, magnetic or electromagnetic fields in the frequency range of 100 kHz to 300 GHz and that have reported exposure using at least one of the situations listed below. These different situations are to accommodate different exposure metrics and exposure types reported in studies, allowing an optimal degree of inclusivity whilst at the same time ensuring confidence in there being a specified contrast between experimental and control conditions. These situations are as follows:

A) body/tissue/sample internal exposure metrics measured or calculated for the particular conditions of the experiment as,

- SAR [expressed in W/kg or equivalent units];
- SA [expressed in J/kg or equivalent units];
- induced electric field strength [expressed in V/m or equivalent units];
- internal magnetic field strength [expressed in A/m or equivalent units].

For exposure applied as pure or predominantly magnetic fields in the lower frequency range the external magnetic field strength at sample position is considered a sufficient surrogate for the tissue internal magnetic field as long as the penetration depth is high compared to the sample dimension.

B) body/tissue/sample internal exposure metrics describing superficial absorption at frequencies above 6 GHz measured or calculated for the conditions of the experiment as follows,

- incident power flux density [expressed in W/m^2 or equivalent units];
- incident energy density [expressed in J/m^2 or equivalent units];
- transmitted (absorbed) power flux density [expressed in W/m^2 or equivalent units];
- transmitted (absorbed) energy density [expressed in J/m^2 or equivalent units].

C) body/tissue/sample external exposure metrics

- external electric field strength [V/m] ($E > 1$ V/m or $E > \sqrt{10}$*background level in unshielded environment, otherwise no restriction);
- external magnetic field strength [mA/m] ($H > 2.7$ mA/m or $H > \sqrt{10}$*background level in unshielded environment, otherwise no restriction);
- incident power flux density [mW/m^2] ($PD > 2.5$ mW/m^2 or $PD > 10$*background level in unshielded environment, otherwise no restriction).

We will only include studies reporting external metrics as under C if (i) either of these exposure metrics was measured or calculated at the location of the exposed body/tissue/sample in the approximate far-field of the field source, and (ii) the exposure level is at least a factor of 10 (power flux density) or $\sqrt{10}$ (field strength) above background level. In the case where no specific background exposure level in the laboratory is reported in the study, we will assume a value of 0.25 mW/m^2 (corresponding to 0.3 V/m and 0.9 mA/m, respectively) as the background exposure level. This results in an inclusion threshold of $PD = 2.5$ mW/m^2, $E = 1$ V/m, or $H = 2.7$ mA/m).

D) mobile phones or other RF-generating devices as source of exposure without reporting of metrics under A, B or C. We will consider these studies separately because there is likely a greater variation and/or uncertainty in exposure levels owing to these being, in most cases, inferred rather than directly measured. Two possible exposure scenarios will be considered:

1. With output controlled by appropriate software or hardware operated close to the tissue/sample. We will include studies that applied exposure with an output power controlled by hardware or software, provided that the output power and the distance to the sample are reported, enabling inference of the exposure.
2. In GSM mode with an active call operated close to the body/tissue/sample. An exposure applied as the field generated by a mobile phone in GSM mode with an active call operated at distances equal or less than 3 cm from the body/tissue/sample can be expected to generate a temporal peak SAR in the range of 0.01 to 100 W/kg, that means at least a factor 100 above the average background level. We will include these studies only if the active call was maintained throughout the experiment and the comparison was a similar phone switched off. Therefore, only mobile phones in GSM mode with an active call operated close to the body/tissue/sample are considered to ensure sufficient exposure contrast. It is noted that in operating modes other than GSM (e.g., UMTS, LTE), the power control of mobile phones is substantially more efficient than in GSM mode, resulting in the possibility of transmit power levels which are too low to ensure sufficient exposure contrast above background level. Therefore studies with mobile phones in mode other than GSM will be included provided that a sufficient contrast with respect to the comparator is assured by measured or calculated metrics.

We will exclude studies that:

- have applied exposure signals with more than 10% of the total signal energy outside the considered frequency range 100 kHz – 300 GHz, or the paper does not report sufficient detail to determine the spectral content and hence the percent signal energy outside the considered frequency range;
- have applied exposure with a mobile phone in cases where at least one of the following conditions apply: i) the mobile phone was not

Table 1
Eligibility criteria.

PECO	*Protocol N°*	*Inclusion criteria*	*Exclusion criteria*
Population	1	- Male subjects of experimental mammalian models	- Humans - Non-mammalian experimental models
	2	- Human sperm exposed *in vitro*	- Non-human sperm - Cell types other than sperm
	3	- Experimental mammals exposed exclusively *in utero*	- Non-mammalian species - Experimental mammals not exposed *in utero*
Exposure	1, 2, 3	- RF-EMF (frequency range 100 kHz − 300 GHz) at any exposure level	- Static or extremely low-frequency magnetic and/or electric fields - Optical radiation - 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 section 3.1.2 A, B or C. - Co-exposure to RF-EMF and other chemical or physical agents
Comparators	1, 2, 3	- Sham-exposed controls - Temperature controls	- Historical controls
Outcomes	1	- Fertilization rate and embryonic survival - Sperm quality (count, viability, motility and morphology) - Sperm oxidative stress, apoptosis, DNA/chromatin alterations - Reproductive organ toxicity (testis/epididymis weight and quantitative histological alterations, post-meiotic cell fraction decrease, testicular cell apoptosis) - Testosterone level in serum or reproductive organs	- Qualitative evidence of toxic effects on testis and epididymis - Endpoints not predictive for male fertility impairment - Endpoints measured with invalid methods
	2	- WHO markers of sperm quality (viability, motility, morphology) - Sperm oxidative stress, apoptosis, DNA/chromatin alterations	- Qualitative evidence of toxic effects on sperm - Sperm alterations not predictive for male fertility impairment - Sperm alterations measured with invalid methods
	3	- Adverse pregnancy outcomes - Birth defects, according to OECD TG 414 - Delayed effects (brain weight and neuropathology, behavioural effects, according to OECD TG 426, early-onset cancer, female fertility impairment)	- Qualitative evidence of alterations in fetuses, pups or adult progeny - Molecular alterations in the progeny that are not causally linked to human pathologies - Effects in adult progeny for which there is not a solid knowledge regarding their possible developmental origin - Delayed effects at cellular or tissue levels in organs other than brain or ovary

operated in GSM mode and the output power was not controlled by hardware or software specifically for the experiment; ii) no active call was established and maintained during the experiment, because of the potentially extremely small exposure contrast generated which we do not consider relevant.
- have evaluated biological effects of static and/or electric fields, optical radiation, or Magnetic Resonance Imaging (MRI);
- have applied co-exposure to RF-EMF and any other potentially beneficial or detrimental chemical or physical exposure.

3.1.3. Types of comparators

Protocols 1 and 3. Reference groups will comprise animals that are sham-exposed. This would mean that the sham-exposed animals were treated under conditions similar to those of exposed animals except for RF-EMF exposure, with particular reference to possible restraining conditions, anesthesia and stressing manipulations. In addition, when the study comprises more than one exposure level, the effects reported for each level of exposure will be also compared in order to evaluate the evidence for a dose–response. If the included studies used separate temperature control groups, we will use them as additional comparators, in order to investigate the specific impact of exposure-induced heating. Historical controls will be excluded.

Protocol 2. Reference groups will comprise human sperm samples, preferably from the same donor or pool of donors, which are handled in the same way as the exposed samples but are concurrently sham-exposed. In addition, when the study comprises more than one exposure level, the effects reported for each level of exposure will also be compared in order to evaluate the evidence for a dose–response. If the included studies used separate temperature control groups, we will use them as additional comparators, in order to investigate the specific impact of exposure-induced heating. Historical controls will be excluded.

3.1.4. Types of outcomes

We will include studies that evaluated effects on one or more of the outcomes listed below. In addition, only studies evaluating those outcomes by biomarkers that we deem valid as explained in Supplementary File 2, will be included.

Protocol 1. Male infertility-related outcomes measurable in sperm, testis and other reproductive organs, in serum or other body fluids, or alterations in the embryos conceived by exposed male animals. In particular, they will comprise: 1) fertilization rate and embryonic survival measured as pre- and post-implantation embryonic losses; 2) indicators of sperm quality, namely sperm count (count of sperm from epididymis or in semen), viability, motility, morphology; 3) oxidative stress, apoptosis, DNA/chromatin alterations in sperm; 4) reproductive organ toxicity (weight of testes and epididymes, quantitative alterations of testis or epididymis histology, decrease of testicular post-meiotic cell fraction; testicular cell apoptosis); 5) changes of testosterone level in serum or reproductive organs.

Protocol 2. 1) WHO markers of sperm quality such as decrease of viability, decrease of motility, changes in morphology; 2) oxidative stress, apoptosis, DNA/chromatin damage.

Protocol 3. Outcomes evaluated at the embryonic/fetal stage, at birth or during adulthood that can be related to adverse pregnancy outcomes or congenital disorders in humans. According to international guidelines (OECD 2007, 2018; Makris et al. 2009), they will be grouped into: 1) adverse pregnancy outcomes; 2) birth defects; 3) delayed effects, with special reference to behavioural traits.

For adverse pregnancy outcomes, the following parameters will be considered: number of implantations, number and percent of live and dead fetuses and resorptions; number and percent of pre-implantation losses (in comparison to the number of corpora lutea).

For birth defects, the following parameters will be considered: number and percent of live offspring; sex ratio; fetal body weight; total number and percent of fetuses at term with any external (gross abnormalities, body length, head length, head width), soft tissue (by gross or morphometric examination), or skeletal (by specific bone staining) malformations, as well as the types and incidences of individual anomalies and other relevant alterations; anogenital distance related to weight.

For delayed effects, behavioural traits such as motor activity, motor and sensory function, learning and memory will be considered if quantitatively assessed in addition to brain weight and neuropathology; furthermore, incidence of early-onset cancer or of female fertility impairment will be evaluated (male fertility impairment will be considered in Protocol 1).

3.1.5. Types of studies

3.1.5.1. Inclusion criteria. Only original, controlled experimental studies on non-human mammals exposed *in vivo* or on human sperm exposed *in vitro* will be considered.

3.1.5.2. Exclusion criteria. Non-experimental studies (e.g. human epidemiologic or other observational studies) will be excluded. In addition, papers reporting reviews, opinions, proceedings or abstracts to meetings will be excluded.

3.1.5.3. Years considered. No year-of-publication limitation will be applied.

3.1.5.4. Publication language. No language restriction will be applied. Articles in languages other than the ones spoken by the reviewers (English, Italian, Chinese) will be translated into English using Google Translator (https://translate.google.com/) and DeepL Translator (https://www.deepl.com/en/translator). However, considering that title and abstract of non-English articles published in peer-reviewed journals are in English, only English terms will be used to search the publication databases.

3.1.5.5. Publication types. Peer-reviewed publications.

3.1.6. Types of effect measures

While included studies have to provide an effect measure as described in the inclusion criteria, the type of effect measure will not influence the decision to include or exclude studies. For dichotomous outcomes we will use the Relative Risk (RR) as the measure of the effect, deriving it from the data reported if not provided by the authors. For continuous outcomes we will use Mean Differences (MD) as the effect size. When the same outcome is measured with different scales, we will use Standardized MDs as the effect size. If data are available for more than two exposure levels, we will calculate incremental RRs or MDs per unit of exposure increase.

3.2. Information source and search strategy

Three publication databases will be 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 effects. The list of retrieved publications will be integrated by a manual search of the lists of references of included papers and published reviews. An electronic library of publications will be created in the Endnote® format.

The NCBI PubMed and SCOPUS databases will be interrogated with search queries, specific for each of the three protocols, composed by elements identifying the exposure, the disease, the populations. These elements will be combined in the queries by the Boolean operators "AND/OR/NOT" (Supplementary Files 3–5). Search terms have been identified to retrieve all relevant peer-reviewed publications on RF-EMF

effects on male fertility, adverse pregnancy outcomes and congenital disorders, by reviewing PubMed Medical Subject Heading (MeSH) terms associated to relevant papers and testing these and other terms chosen by expert judgment through an iterative trial-and-error process. The aim is to reach a publication list as inclusive as possible. The removal of non-experimental and human studies (see eligibility criteria above) will be done manually rather than by the use of search filters because studies might have been incorrectly indexed in the databases.

The EMF Portal database will be searched 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 3–5).

The search outputs will then be aligned to exclude duplicates and the resulting list will be screened for eligibility criteria.

The searches will be re-run just before the final analyses to retrieve the most recent studies eligible for inclusion.

3.3. Study selection

Two reviewers will independently evaluate the relevance of the identified papers on the basis of titles and abstracts, to exclude records that are obviously not relevant or do not fulfil one or more of the inclusion criteria for the PECO elements. In case of disagreement between the reviewers, the paper will be automatically moved to the full text phase. Studies of which titles/abstracts do not provide enough information to decide will be automatically moved to the next phase to assess the full text. At this later phase, again two reviewers will independently evaluate the relevance of the identified papers, and disagreement between the reviewers will be resolved by discussion. If no consensus can be reached, a third reviewer will be consulted. Every effort will be made to retrieve the full text of a paper; in the event that full text of a study is not available, the reason will be clearly reported in a tracking report. Exclusion at this stage will also be recorded in the tracking document with a brief justification of exclusion rationale. This step will result in a list of included studies. A list of the studies excluded after reading the full text with reasons for exclusion will be attached as a Supplementary File to the open-access publication of the systematic review. If findings from a study are described in more than one article, these will be considered as one study only, although data retrieved from different papers might be used to obtain as much information as possible on the study. We will document the selection process in a study flow diagram according to the PRISMA reporting guidelines (Liberati et al. 2009).

3.4. Data extraction

A standard set of details will be extracted from the relevant publications and recorded on data extraction forms appositely designed and agreed among the team members. A list of data items to be extracted has been customized for each of the three systematic reviews starting from NTP/OHAT templates (NTP, 2015a) (Supplementary Files 6–8). The lists will include information on study identification and study characteristics such as, study design, description of population, exposure, comparator(s) and outcome(s), statistical methods applied to analyse results.

Regarding study results, for each analyzed endpoint effect size data will be extracted. For dichotomous variables, values necessary to calculate the Relative Risk (RR) from the 2×2 table will be extracted (e.g. number of events and number of animals/replicates in either the experimental or control group); for continuous outcomes, number of animals and means with standard deviations/errors in each experimental group will be extracted.

For all the eligible studies, one reviewer will extract and record the relevant features and a second reviewer will check all the extracted study information against the accompanying article(s) for completeness and accuracy as a quality control measure. If disagreement occurs between the reviewers, this will be resolved by discussion; if no consensus can be reached a third reviewer will be involved. Review authors will not undertake either data extraction or risk of bias assessments (see below) of any studies in which they have been involved.

Two documents have been prepared, one regarding the endpoints and techniques applied for assessing the outcomes, one regarding exposure setting and dosimetry. The first one (Supplementary File 2) is a guideline to help the reviewers in the extraction of results and risk of bias assessment of the studies, according to the validity and reliability of the measured endpoints and applied techniques. The second one (Supplementary File 9) is a tool aimed at reaching a harmonized awareness of the critical exposure issues to be considered.

The studies with their characteristics and the outcome results will be listed in Excel tables to be able to derive the most appropriate comparisons. If data necessary for analysis are missing from the articles, they will be calculated from figures using digital rulers or from other data that are reported. Whenever essential data are missing or there are inconsistencies among the reported data, the authors will be contacted to provide further details by sending an e-mail followed by one reminder in case of no-reply.

3.5. Risk of bias assessment

For each of the three systematic reviews, risk of bias will be 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. 2014b). Since the NTP/OHAT guidelines are specifically designed for human and animal studies, adaptations to *in vitro* studies have been developed for the systematic review on RF-EMF effects on human sperm *in vitro*, following the indication of Bodewein et al. (2019), Golbach et al. (2016), Romeo et al. (2021).

Six bias domains will be 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 will guide the reviewers in the assessment of the internal quality of data.

Two investigators will independently analyse included papers for RoB assessment, and possible disagreements will be resolved by discussion with a third investigator. Based on OHAT guidelines, one of four RoB scores will be assigned to each bias question: definitely low RoB, probably low RoB, probably high RoB, definitely high RoB, with a statement of justification based on the study text. In case of poor reporting of essential methodological details in the included studies "probably high RoB" will be assigned. RoB will be evaluated at the outcome level, meaning that one individual study subject to different outcome-based experiments will receive multiple RoB evaluations. Two spreadsheets, one specific for animal studies and one specific for *in vitro* studies, will be used to record the article quotes and the scores assigned by each reviewer. Explanatory guidelines on how to assign scores on the basis of a predefined set of criteria have been prepared (Supplementary file 10).

As a last step, the scores for the different questions will be integrated to obtain the study overall RoB estimate. Due to the still limited experience on the application of RoB tools to experimental animal studies, it has been chosen not to select key questions among all those considered, but to assign the overall study RoB based on all of them. A study will be labelled "high concern" in cases where one question is answered with "definitely high RoB". A study will be labelled "low concern" in cases where none of the questions are answered with "probably high RoB" or "definitely high RoB". All other studies will be labelled as "some concern". This 3-level scale corresponds to the 3-tier system of study classification proposed by the OHAT approach.

3.6. Synthesis of results

The impact of RF-EMF on male fertility in experimental mammals

will be assessed based on 5 different outcomes: 1) fertilization rate and embryonic survival, 2) WHO sperm quality markers, 3) other sperm integrity biomarkers, 4) reproductive organ toxicity, 5) testosterone level. The impact of RF-EMF on male fertility in human sperm exposed *in vitro* will be assessed based on 2 different outcomes: 1) WHO sperm quality markers, and 2) other sperm integrity biomarkers. The impact of RF-EMF on pregnancy and birth of experimental mammals will be assessed based on 3 different outcomes: 1) adverse pregnancy outcomes, 2) birth defects and 3) delayed effects.

Data relative to several different endpoints, expressed as effect size measures, will contribute to each outcome (as summarized in Table 2).

Among the 5 outcomes evaluated to assess RF-EMF impact on male fertility, the highest importance will be given to fertilization rate and embryonic survival, and sperm quality measured by WHO parameters. The first one because of directly measuring the fertility potential of the experimental subjects; the second one because the analyzed target is the mature male gamete, direct effector of fertilization and the endpoints evaluated are the same applied in andrological clinics for the assessment of human male fertility. Similarly, to assess the effects on RF-EMF on human sperm exposed *in vitro*, higher relevance will be given to the first than to the second outcome. To assess the impact of RF-EMF on pregnancy and birth, the highest relevance will be given to the second outcome because of the social burden of perinatal deaths and congenital malformations in humans, and because the connection between prenatal exposure and adverse birth effects is closer than for delayed effects.

For each of the three PECO questions we will pool data from studies using a random effects *meta*-analysis. If this is not possible a narrative synthesis of the findings will be conducted from the included studies supported by tables and plots. In the *meta*-analyses, for dichotomous outcomes contingency tables will be built and relative risks (RRs) will be calculated. For continuous outcomes, the mean difference (MD) with 95% confidence intervals will be estimated. If similar continuous outcomes are measured on different scales Standardized Mean Differences (SMD) will be used. If data are available for more than two exposure levels, incremental RRs or MDs per unit of exposure increase will be calculated.

3.6.1. Assessment of heterogeneity

Only studies that are considered sufficiently similar in the PECO elements will be combined. For each PECO, the different outcomes will be considered independent streams of evidence, which will be, eventually, integrated to assess RF-EMF impact on the ultimate effects, infertility and adverse effects on pregnancy and birth. Studies evaluating different endpoints within the same outcome will be combined. For exposure, we consider studies similar enough to be combined irrespectively of near-field or far-field exposure condition, modulation and RF-EMF frequency applied.

The *meta*-analysis will be based on a random effects model, because the underlying effect size is expected to differ between studies due to the explorative nature and diversity in animal studies. Statistical heterogeneity of results will be assessed by the I^2 and the τ^2 statistics and reported. I^2 represents the inconsistency between the study results and

Table 2
Body of evidence structure based on major experimental outcomes and endpoints to guide grouping and sub-grouping of evidence.

Outcome	*Endpoints*
Protocol 1	
Decrease of fertilization rate and embryonic survival	- *In vitro* fertilization rate - *In vivo* mating/fertilization rate - Pre- and post-implantation embryonic losses
Alterations of WHO sperm quality parameters	- Sperm count - Sperm viability - Sperm motility - Sperm morphology
Alterations of other sperm integrity biomarkers	- Oxidative stress - Apoptosis - DNA/chromatin alterations
Reproductive organ toxicity	- Weight of testes and epididymis - Quantitative alterations of testis or epididymis histology - Decrease of testicular post-meiotic cell fraction - Testicular cell apoptosis
Alterations of reproductive hormones	- Testosterone level in serum or reproductive organs
Protocol 2	
Alterations of WHO sperm quality parameters	- Sperm viability - Sperm motility - Sperm morphology
Alterations of other sperm integrity biomarkers	- Oxidative stress - Apoptosis - DNA/chromatin alterations
Protocol 3	
Adverse pregnancy outcomes	- Number of implantations - Number and percent of live and dead fetuses and resorptions - Number and percent of pre-implantation losses
Birth defects	- Number and percent of live offspring - Sex ratio - Fetal body weight - Total number and percent of fetuses at term with any external, soft tissue, or skeletal malformations - Anogenital distance
Delayed effects	- Brain weight and neuropathology - Behavioural traits - Early onset cancer - Female infertility*

* Male infertility will be evaluated in protocol 1.

quantifies the proportion of observed dispersion that is real, i.e. due to between-study differences and not due to random error. I^2 reflects the extent of overlap of the confidence intervals of the study effects. I^2 represents the inconsistency on a scale between 0 and 100, therefore it can be compared with suggested limits for low or high inconsistency. However, sometimes it may be misleading, because it depends on sample size of the included studies; with very large (highly precise) studies, even tiny differences in effect size may result in a high I^2, while with small (imprecise) studies, very different exposure effects can yield a low value of I^2. The τ (the square root of τ^2) is the standard deviation of the between-study variation on the scale of the original outcome. The τ^2 is the direct estimate of the between-study variation and therefore it is useful in calculations, for example, for the prediction interval. In addition to the I^2 and the τ^2 statistics, an 80% prediction interval will be estimated from the distribution of effect estimates, which is the interval of effect estimates comprising the true effect size for 80% of analyzed studies (IntHout et al., 2016). When *meta*-analysis is not possible, the similarity between the studies in the direction of the effects will be judged.

3.7. Additional analyses

3.7.1. Subgroup analyses

To explore the sources of heterogeneity we will consider the possibility of subgroup analyses within the groups previously identified by expert judgment to comprise studies sufficiently similar for exposure and outcome. We will conduct the subgroup analysis if there are at least 3 studies per subgroup and we will assess whether the pooled effect sizes are statistically different. For Protocols 1 and 3, subgroups will be based on:

- the animal species (mouse vs rat);
- the exposed stage of development (P1: pre- vs post-pubertal; P3: early vs late pregnancy);
- the specific endpoints contributing to define the same outcome (e.g., sperm motility vs sperm morphology or percent of live offspring vs percent of offspring with malformations);
- the frequency, duration, modulation and level of exposure (e.g., acute vs protracted exposure or low vs high exposure level);
- the induction or not of substantial temperature increase in the target.

For Protocol 2 subgroups will be based on:

- the source of sperm (fertile vs infertile or subfertile individuals);
- the specific endpoints contributing to define the same outcome (e.g., sperm motility vs sperm morphology);
- the frequency, duration, modulation and level of exposure (e.g., acute vs protracted exposure or low vs high exposure level);
- the induction or not of substantial temperature increase in the target.

3.7.2. Sensitivity analyses

We will assess the effect of the risk of bias by comparing the results of the studies that are at different levels of concern for specific questions and specific types of outcome and assess if these are statistically different. We will also assess the sensitivity of assumptions that we have made to assess exposure levels and categories.

3.7.3. Publication bias

A covariate that may be influential and is highly relevant to random-effects *meta*-analyses is study size. A relationship between exposure effect and study size (or study precision) induces an asymmetric funnel plot. If the funnel plot, upon visual inspection shows that small studies with non-harmful effects are missing, this would be an indication of publication bias. If ten or more studies are included in the same *meta*-analysis, an Egger's test will be applied to evaluate potential small study bias, otherwise a qualitative evaluation will be made (Egger et al. 1997).

3.8. Certainty of evidence assessment

In order to judge the certainty in the evidence of the effects observed in the SRs and draw reliable conclusions, a framework based on the GRADE principles will be used (https://www.gradeworkinggroup.org).

We will start the rating of the certainty of the evidence for animal studies at high certainty evidence just like in human experimental studies (Hooijmans et al. 2018). Then, depending on which criteria for which domains are met, we will downgrade the certainty of the evidence to moderate, low, or very low. The same strategy will be used for *in vitro* studies.

One upgrading factor may apply for animal studies, namely consistency between species.

The reasons for downgrading that will be considered are: limitations in studies, indirectness, inconsistency, imprecision and publication bias. Indirectness for both animal and *in vitro* studies will be assessed in a two-step process: first assess how well the animal/*in vitro* PECO has been addressed and then how well the study addressed the human PECO.

An elaboration on how to score each criterion is created for both the animal and *in vitro* studies (Supplementary File 11).

Financial support

This project is partially funded by the World Health Organization (contract 2020/1026306-0). Additional in-kind funds are provided by ENEA, Health Canada and Swinburne University of Technology.

Author contributions

Starting from WHO problem formulation and protocol drafts, FP led the team and together with EC coordinated protocol development. RBMdV and CRH provided expertise on systematic review methodology and risk of bias assessment. Expertise on RF-EMF dosimetry and exposure set-ups was provided by LA, GC, CM, JPMN, AWW. Expertise on male infertility and adverse pregnancy outcome assessment, in experimental animal and *in vitro* model systems, was provided by FP, EC, PE and MHB. BB and CC provided expertise on molecular markers of male germ cell alterations and adverse pregnancy outcomes. MS provided specific expertise on statistical issues. All authors contributed to editing and finalizing the paper.

Declaration of Competing Interest

AWW directs a research group, which includes three technical associates who are telecommunications company employees. The group is also providing advice for a local government authority and a utility on electric and magnetic field exposure issues on a fee-for-service basis. 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. JPM 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, Eric van Rongen and Jos Verbeek, for the protocol drafts, fruitful discussion, helpful advice and manuscript revision.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2021.106806.

References

Adams, J.A., Galloway, T.S., Mondal, D., Esteves, S.C., Mathews, F., 2014. Effect of mobile telephones on sperm quality: a systematic review and meta-analysis. Environ. Int. 70, 106–112.

Agarwal, A., Mulgund, A., Hamada, A., Chyatte, M.R., 2015. A unique view on male infertility around the globe. Reprod. Biol. Endocrinol. 26 (13), 37.

Agarwal, A., Singh, A., Hamada, A., Kesari, K., 2011. Cell phones and male infertility: a review of recent innovations in technology and consequences. Int. Braz. J. Urol. 37 (4), 432–454.

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.

Baldacci, S., Gorini, F., Santoro, M., Pierini, A., Minichilli, F., Bianchi, F., 2018. Environmental and individual exposure and the risk of congenital anomalies: a review of recent epidemiological evidence. Epidemiol. Prev. 42, 1–34.

Bernabò, N., Ciccarelli, R., Greco, L., Ordinelli, A., Mattioli, M., Barboni, B., Lozano, S., 2017. Scientometric study of the effects of exposure to non-ionizing electromagnetic fields on fertility: a contribution to understanding the reasons of partial failure. PLoS One. 12 (12), e0187890.

Bodewein, L., Schmiedchen, K., Dechent, D., Stunder, D., Graefrath, D., Winter, L., Kraus, T., Driessen, S., 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.

Dama, M.S., Bhat, M.N., 2013. Mobile phones affect multiple sperm quality traits: a meta-analysis. F1000Research 2, 40.

Deepinder, F., Makker, K., Agarwal, A., 2007. Cell phones and male infertility: dissecting the relationship. Reprod. Biomed. Online 15 (3), 266–270.

Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315, 629–634.

Gabrielsen, J.S., Tanrikut, C., 2016. Chronic exposures and male fertility: the impacts of environment, diet, and drug use on spermatogenesis. Andrology. 4, 648–661.

Golbach, L.A., Portelli, L.A., Savelkoul, H.F.J., Terwel, S.R., Kuster, N., de Vries, R.B.M., Verburg-van Kemenade, B.M.L., 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.

Heynick, L.N., Merritt, J.H., 2003. Radiofrequency fields and teratogenesis. Bioelectromagnetics 24 (S6), S174–S186.

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

Hooijmans, C.R., IntHout, J., Ritskes-Hoitinga, M., Rovers, M.M., 2014a. Meta-analyses of animal studies: an introduction of a valuable instrument to further improve healthcare. ILAR J. 55 (3), 418–426.

Hooijmans, C.R., Rovers, M.M., de Vries, R.B., Leenaars, M., Ritskes-Hoitinga, M., Langendam, M.W., 2014b. SYRCLE's risk of bias tool for animal studies. BMC Med. Res. Methodol. 14, 43.

Houston, B.J., Nixon, B., King, B.V., De Iuliis, G.N., Aitken, R.J., 2016. The effects of radiofrequency electromagnetic radiation on sperm function. Reproduction 152, R263–R276.

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.

IntHout, J., Ioannidis, J.P.A., Rovers, M.M., Goeman, J.J., 2016. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open 6 (7), e010247. https://doi.org/10.1136/bmjopen-2015-010247.

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

Kesari, K.K., Agarwal, A., Henkel, R., 2018. Radiations and male fertility. Reprod. Biol. Endocrinol. 16, 118.

Levine, H., Jørgensen, N., Martino-Andrade, A., Mendiola, J., Weksler-Derri, D., Mindlis, I., Pinotti, R., Swan, S.H., 2017. Temporal trends in sperm count: a systematic review and meta-regression analysis. Hum. Reprod. Update. 23, 646–659.

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med. 2009 (6), e1000100.

Liu, K., Li, Y., Zhang, G., Liu, J., Cao, J., Ao, L., Zhang, S., 2014. Association between mobile phone use and semen quality: a systemic review and meta-analysis. Andrology 2 (4), 491–501.

Makris, S.L., Solomon, H.M., Clark, R., Shiota, K., Barbellion, S., Buschmann, J., Ema, M., Fujiwara, M., Grote, K., Hazelden, K.P., Hew, K.W., Horimoto, M., Ooshima, Y., Parkinson, M., Wise, L.D., 2009. Terminology of developmental abnormalities in common laboratory mammals (Version 2). Birth Defects Res. B Dev. Reprod. Toxicol. 86 (4), 227–327.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Akl, E.A., Brennan, S.E., Chou, R., Glanville, J., Grimshaw, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Stewart, L.A., Thomas, J., Tricco, A.C., Welch, V.A., Whiting, P., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. Syst. Rev. 10, 89.

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.

Organization for Economic Cooperation and Development (OECD), 2007. TG 426. Developmental Neurotoxicity Study. OECD Guideline for Testing of Chemicals.

Organization for Economic Cooperation and Development (OECD), 2018. TG 414. Prenatal developmental toxicity study. OECD Guideline for Testing of Chemicals.

Romeo, S., Zeni, O., Sannino, A., Lagorio, S., Biffoni, M., Scarfì, M.R., 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.

Rooney, A.A., Boyles, A.L., Wolfe, M.S., Bucher, J.R., Thayer, K.A., 2014. Systematic review and evidence integration for literature-based environmental health science assessments. Environ. Health Perspect. 122 (7), 711–718.

Santini, S.J., Cordone, V., Falone, S., Mijit, M., Tatone, C., Amicarelli, F., Di Emidio, G., 2018. Role of mitochondria in the oxidative stress induced by electromagnetic fields: focus on reproductive systems. Oxid. Med. Cell Longev. 2018, 1–18.

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).

Shah, S.G.S., Farrow, A., 2014. Systematic literature review of adverse reproductive outcomes associated with physiotherapists' occupational exposures to non-ionising radiation. J. Occup. Health 56 (5), 323–331.

Steer, P., 2005. The epidemiology of preterm labour. BJOG 112, 1–3.

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

Vornoli, A., Falcioni, L., Mandrioli, D., Bua, L., Belpoggi, F., 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. https://doi.org/10.3390/ijerph16183379.

Xavier, M.J., Salas-Huetos, A., Oud, M.S., Aston, K.I., Veltman, J.A., 2021. Disease gene discovery in male infertility: past, present and future. Hum. Genet. 140 (1), 7–19.

Whaley, P., Aiassa, E., Beausoleil, C., Beronius, A., Bilotta, G., Boobis, A., de Vries, R., Hanberg, A., Hoffmann, S., Hunt, N., Kwiatkowski, C.F., Lam, J., Lipworth, S., Martin, O., Randall, N., Rhomberg, L., Rooney, A.A., Schünemann, H.J., Wikoff, D., Wolffe, T., Halsall, C., 2020. Recommendations for the conduct of systematic reviews in toxicology and environmental health research (COSTER). Environ. Int. 143, 105926. https://doi.org/10.1016/j.envint.2020.105926.

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

Wilcox, A.J., Weinberg, C.R., O'Connor, J.F., Baird, D.D., Schlatterer, J.P., Canfield, R.E., Armstrong, E.G., Nisula, B.C., 1988. Incidence of early loss of pregnancy. N. Engl. J. Med. 319, 189–194.