Environment International 148 (2021) 106386

Contents lists available at ScienceDirect

Environment International

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

Genotoxicity of radiofrequency electromagnetic fields: Protocol for a systematic review of *in vitro* studies

Stefania Romeo [a,*], Olga Zeni [a], Anna Sannino [a], Susanna Lagorio [b], Mauro Biffoni [b], Maria Rosaria Scarfì [a]

[a] Institute for Electromagnetic Sensing of the Environment (IREA), Italian National Research Council (CNR), 80124 Napoli, Italy
[b] Department of Oncology and Molecular Medicine, National Institute of Health, 00161 Roma, Italy

ARTICLE INFO

Handling Editor: Dr. Paul Whaley

Keywords:
Radiofrequency electromagnetic fields
Genotoxicity
In vitro
Systematic review protocol

ABSTRACT

Background: Exposure to radiofrequency electromagnetic fields (RF-EMF, 100 kHz – 300 GHz) emitted by wireless communication technologies is pervasive and ubiquitous. Concern has been raised about possible adverse effects to human health. In 2011 the International Agency for Research on Cancer has classified RF-EMF as possibly carcinogenic to humans, highlighting that the evidence is weak and far from conclusive. Updated systematic reviews of the scientific literature on this topic are lacking, especially for mechanistic studies.

Objectives: To develop a protocol for a systematic review of experimental studies investigating genotoxic effects induced by RF-EMF in *in vitro* cellular models. Genotoxicity is one of the key-biological indicators of carcinogenicity, and the most common characteristics of established carcinogens. The predefined procedures for conducting the systematic review are outlined below.

Methods: We will follow the guidelines developed by the National Toxicology Program-Office of Health Assessment and Translation (NTP-OHAT), adapted to the evaluation of *in vitro* studies.

Eligibility criteria: We will include experimental *in vitro* studies addressing the relationship between controlled exposures to RF-EMF and genotoxicity in mammalian cells only. Eligibility for inclusion will be further restricted to peer reviewed articles reporting findings from primary studies.

Information sources: We will search the scientific literature databases NCBI PubMed, Web of Science, and EMF-Portal. No filter on publication date will be applied. Only studies published in English will be considered. The reference lists of the included papers and available reviews will be screened for unidentified relevant papers. References will be managed through Endnote X9 software.

Data extraction and synthesis of results: Data from included papers will be extracted according to predefined forms. Heterogeneity within the available evidence will determine the type of evidence synthesis that is appropriate. Findings will be summarized in tables, graphical displays and in a narrative synthesis of the available evidences. A *meta*-analysis will be carried out if subgroups of studies homogeneous in terms of exposure characteristics, endpoint, and cell types will be identified.

Risk of bias: The internal validity of included studies will be assessed using the NTP-OHAT Risk of Bias Rating Tool for animal studies, adapted to *in vitro* studies. This stage of the process will be managed through the Health Assessment Workspace Collaborative (HAWC).

Evidence appraisal: To rate confidence in the body of evidence, we will use the OHAT GRADE-based approach for animal studies.

Framework and funding: This protocol concerns one of the evidence streams considered in a larger systematic review of the scientific literature on the potential carcinogenicity of RF-EMF, performed by scientists from several Italian public research agencies. The project is supported by the Italian Workers' Compensation Authority (INAIL) in the framework of the CRA with the Istituto Superiore di Sanità "BRiC 2018/06 – Scientific evidence on the carcinogenicity of radiofrequency electromagnetic fields".

* Corresponding author at: Institute for Electromagnetic Sensing of the Environment (IREA), Italian National Research Council (CNR), 80124 Napoli, Italy.
E-mail addresses: romeo.s@irea.cnr.it (S. Romeo), zeni.o@irea.cnr.it (O. Zeni), sannino.a@irea.cnr.it (A. Sannino), susanna.lagorio@iss.it (S. Lagorio), mauro.biffoni@iss.it (M. Biffoni), scarfi.mr@irea.cnr.it (M.R. Scarfì).

https://doi.org/10.1016/j.envint.2021.106386
Received 2 July 2020; Received in revised form 30 December 2020; Accepted 6 January 2021
Available online 21 January 2021

1. Introduction

1.1. Background and rationale

Exposure to radiofrequency electromagnetic fields (RF-EMF, 100 kHz–300 GHz) emitted by wireless communication technologies (mobile phones, base stations, wireless local area networks, etc.) has become pervasive and ubiquitous, raising concern about possible adverse effects to human health.

The only established effects of RF-EMF are the stimulation of excitable tissues in the intermediate frequency range (100 kHz–10 MHz), and the increase of temperature due to the absorption of electromagnetic energy by the body tissues (above 10 MHz). This information is the basis for setting of limits of exposure to RF-EMF (ICNIRP, 2020).

Inconclusive and in many cases conflicted are the results regarding possible health hazards due to long-term exposure to low levels (i.e. below the exposure limits) of RF-EMF.

The International Agency for Research on Cancer (IARC) classified RF-EMF as possibly carcinogenic for humans (Group 2B). The IARC Working Group performed a critical review of the relevant literature (human, animal and mechanistic studies) published up to mid- 2011. The overall evaluation was driven by the evidence provided by observational human studies and experimental studies in animals, both classified as *limited*, meaning that a causal association between the exposure and cancer development was not substantiated (IARC, 2013). The IARC panel examined mechanistic studies (which include *in vitro* studies) as supporting evidence, but did not formally consider them in the overall evaluation.

Over the years, several groups of experts have assessed the alleged health effects of RF-EMF by integrating different evidence streams (mainly epidemiological and animal, but also *in vitro* studies). A critical overview of the reviews published in 2009–2011 is provided in (Verschaeve, 2012), and a compilation of the conclusions of reports published between 2008 and 2014 is reported in (Vijayalaxmi and Scarfi, 2014). Similarly to the IARC evaluation, the evidence for adverse health effects of RF-EMF at current exposure levels is considered as weak or inadequate by the large majority of these expert panels. Most reports claim that long-term effects cannot be confidently excluded, and the need of additional research is always highlighted.

However, none of the cited reports was conducted applying the methodologies of systematic reviews, an approach considered as a requirement by the World Health Organization (WHO) in the framework of the updated assessment of health hazards from exposure to RF-EMF (WHO, 2020).

Guidelines for systematic reviews in healthcare are available and largely used for long (Higgins et al., 2019). There is increasing interest in applying this method to environmental health-related questions (Rooney et al., 2014; Whaley et al., 2020). The GRADE approach is flexible enough to cover a wide range of health-related topics, and has recently been extended to environmental health issues (Morgan et al., 2019). However, there are no published guidelines for systematic reviews of mechanistic studies with an *in vitro* exposure setting (Rooney et al., 2016). Similarly, there is no broadly accepted method for assessing carcinogenicity of exogenous agents. The IARC approach to cancer hazard assessment has been recently modified to include the evidence provided by mechanistic studies in the overall evaluation (Samet et al., 2019). The IARC appraisal of this evidence stream relies on a set a 10 key-characteristics of human carcinogens (Smith and Guyton, 2020). One of such characteristics is genotoxicity, i.e. the capability to induce DNA damage, mutation, or both. A large proportion of agents classified by the IARC as group 1 carcinogens is indeed genotoxic (Krewski et al., 2019; Smith et al., 2016).

In vitro investigations account for the majority of studies on the biological effects of RF-EMF available in the literature, as they have a key role in advancing knowledge on possible adverse effects of the exposure and underlying mechanisms (Simko et al., 2016). Several comprehensive reviews (Manna and Ghosh, 2016; Meltz, 2003) and *meta*-analyses (Halgamuge et al., 2020; Vijayalaxmi and Prihoda, 2008; 2012; 2019) of studies addressing genetic damage in mammalian cells exposed to RF-EMF have been carried out. However, although those papers considered a large set of studies, the lack of transparently reported methods, pre-defined inclusion criteria, and formal assessment of susceptibility to bias, makes them non compliant with the structured approach of systematic reviews (Whaley et al., 2020).

It is worth noting that WHO has commissioned systematic reviews of various evidence streams on health hazards from RF-EMF exposure, but none of them addresses experimental studies of genotoxic effects in cellular systems *in vitro* (Verbeek et al., 2021).

1.2. Objective

This paper describes a protocol for a systematic review of the scientific literature on RF-EMF and genotoxicity in *in vitro* experimental models. The overall aim of the planned systematic review is to assess the confidence and level of evidence for genotoxic effects induced by RF-EMF in mammalian cells.

The scientific question, formulated as a PECO statement, is outlined in Table 1. Of note, we use the term "outcome" with reference to genotoxicity (as the measurable construct variable), and define "endpoints" the analyzed biomarkers of this outcome.

This protocol was developed in the framework of a systematic review of the scientific literature on the potential carcinogenicity of RF-EMF which envisages the integration of the epidemiological, *in vivo* and *in vitro* evidence streams.

2. Methods

The systematic review will be performed according to the guidelines developed by the National Toxicology Program-Office of Health Assessment and Translation (NTP-OHAT, 2019). Since these guidelines concern human and *in vivo* studies, we have adapted them to *in vitro* studies. Hints on search strategies and risk of bias assessment were taken from the systematic reviews by (Golbach et al., 2016) and (Bodewein et al., 2019).

The review process will be partially managed through the Health Assessment Workspace Collaborative (HAWC), an open-source content management system (Shapiro et al., 2018).

The current protocol conforms to the PRISMA-P (Preferred Reporting Items for Systematic review and Meta-analysis Protocols) guidelines (Moher et al., 2015; Shamseer et al., 2015), provided as Supplementary Material 1. Possible amendments to this protocol, along with the change date and the rationale, will be documented and acknowledged in the systematic review report (Shamseer et al., 2015).

Table 1
PECO statement.

Population	*In vitro* models of healthy or cancerous mammalian cells (of human or animal origin), either immortalized or freshly collected via drawing/explantation.
Exposure	Controlled *in vitro* exposure to radiofrequency radiation (100 kHz-300 GHz), based on suitable exposure metrics. *Exposure details*: Frequency bands: 100 kHz to < 10 MHz ; 10 MHz to ≤ 6 GHz; > 6 GHz to ≤ 300 GHz; Metrics: induced electric field (E_{ind}, V/m) in the 100 kHz-10 MHz range, Specific Absorption Rate (SAR, W/kg) in the 10 MHz – 6 GHz range, power density (PD) of the incident field (W/m^2) in the 6 GHz – 300 GHz range; Signal characteristics: continuous waves (CW); pulsed (PW); Duration (hours).
Comparator	Either incubator (negative) or sham-exposed (sham) control samples.
Outcome	Genotoxicity, intended as capability of inducing DNA damage and/or mutations, assessed as: *Primary endpoints*: chromosomal aberrations, micronuclei, aneuploidy, spindle disturbances, sister chromatid exchanges, mutations. *Secondary endpoints*: Single and double DNA strand breaks, chromatin condensation, and 8-hydroxy-2′-deoxyguanosine adducts.

2.1. Inclusion and exclusion criteria

We will include experimental *in vitro* studies assessing the capacity of RF-EMF to induce genotoxic effects in mammalian cells (Table 1), with no restrictions on species (humans or animal), biological model (freshly collected cells, or immortalized cells), cell nature (healthy or cancerous), or cell lineage. We will not include studies on genotoxic effects of RF-EMF in non-mammalian cells in order to reduce as much as possible the indirectness of the evidence stream assessed.

We will not apply restrictions on the frequency band in the range 100 kHz to 300 GHz, or on exposure duration. We will exclude studies not providing information on the characteristics of the RF signal (continuous or pulsed waves, CW/PW), as well as those not reporting a quantitative measure of exposure level/dose expressed in the appropriate unit [induced electric field, E_{ind} in V/m (100 kHz–10 MHz); SAR in W/kg (10 MHz–6 GHz); or absorbed power density in W/m^2 (6 GHz–300 GHz)]. Studies in which RF-EMF exposure of the sample is obtained using a commercial source (e.g., a mobile telephone) in contact with or at a certain distance from the sample container, will be excluded if a dosimetry analysis is lacking, because, in such situation, the control of electromagnetic and environmental conditions cannot be assured, resulting in uninterpretable findings and unreplicable experimental conditions (Zeni and Scarfi, 2012).

With reference to the study design, admissibility is restricted to studies including unexposed samples, consisting of either incubator (negative) controls, or sham-exposed controls. The sham-control is a sample placed in an exposure system identical to that used to administer the treatment, except for the emission of RF-EMF, to guarantee the very same environmental conditions to all experimental groups.

Based on the type of DNA damage (irreversible *vs* repairable), we will distinguish between genotoxicity-related endpoints of primary interest (i.e., biomarkers of irreversibile damage, including chromosomal aberrations, micronuclei, aneuploidy, spindle disturbances, sister chromatid exchanges, or mutations), and of secondary interest (i.e., biomarkers of repairable damage, including single and double DNA strand breaks, chromatin condensation, and 8-hydroxy-2′-deoxyguanosine adducts).

For studies that evaluate genotoxicity in relation to both RF-exposure alone, and to co-exposure to RF fields and other agents, only findings concerning RF-exposure alone will be considered, because we want to focus on potential genotoxic effects of RF-EMF themselves.

We will restrict inclusion to peer-reviewed journal articles reporting findings from primary studies, and published in English.

Meeting abstracts, conference proceedings, and commentaries will be excluded, whereas reviews will be used to check for missing articles.

2.2. Information sources

Our primary information sources will be NCBI PubMed, Web of Science (WOS), and EMF-Portal. EMF-Portal (www.emf-portal.org), a thematically specialized literature database on biological and health related effects of EMF which, due to its content specificity and documentated high coverage of the reasearch topic (Bodewein et al., 2019; Driessen et al., 2017), is expected to have a better performance compared to the other two information sources. The time coverage of the review will start at the inception date of each database (e.g. 1946 for PubMed), and will end on 31 December 2020, defined as the article *in print* publication date.

Although it might affect the comprehensiveness of the literature search, we will not search for grey literature, because this would imply the need to perform a preliminary peer-review of the reports, which is out of the reach of the current project.

2.3. Search strategy

The search strategies developed for PubMed, WOS and EMF-Portal are reported in Supplementary Material 2. The PubMed search string has been developed by considering the Medical Subject Heading (MeSH) terms, literature tags for relevant and appropriate terms, as well as the standardized search strings for the key topic areas developed by NTP-Office of the Reports on Carcinogens (NTP-ORoC, 2016). The NOT operator has been included in the search to exclude references related to the use of RF-EMF for diagnostic and therapeutic purposes, and for catalysis of chemical reactions. The search string has been then adapted and calibrated to WOS.

In EMF-Portal, the following items have been toggled: "Experimental studies" and "Reviews, summaries, surveys" as topics, "Radiofrequency (≥10 MHz)", and "Mobile Communications" as frequency ranges, "Complete time span" as time coverage. The keywords "cell", "genotoxicity" and "DNA", connected by OR operator, have been also added to refine the search.

2.3.1. Calibration of the PubMed and WOS queries

To calibrate the PubMed and WOS search strategies, we performed a bibliographic search without time restrictions. We then compared the search outputs to the content of a library of "seed studies" investigating biological effects of RF-EMF on *in vitro* models. The library was created over the years for research purposes, and comprises all *in vitro* studies quoted in recent authoritative expert panel reviews (IARC, 2013; AGNIR, 2012; ANSES, 2013; SCENIHR, 2015). Such a "gold standard" database includes 176 papers addressing RF-EMF and genotoxicity endpoints.

The PubMed search was performed on May 5th 2020, and resulted in a total of 2528 records. The bibliographic records were saved in PubMed format and imported into an Endnote library, where they were matched to the records of the personal database. The linkage resulted in 153 successfully matched records. Of the 23 unidentified "gold standard" records, 13 were not indexed in PubMed, and 7 were not sources of primary data (reviews, commentaries, monographies). Therefore, only in 3 cases our search string failed in retrieving relevant papers available in PubMed, showing a very high sensitivity (proportion of relevant records identified by the search = 98%), despite the low precision [(number of relevant records/total retrieved)*100 = 6%] (Sampson and McGowan, 2011; Waffenschmidt et al., 2017).

The WOS search was performed on September 18th 2020, and resulted in 4048 records (4046 after duplicates removal). The bibliographic records were saved in ISI format, imported into an Endnote library, and compared to the "gold standard" obtaining a total of 150 successfully matched records (of the 176 gold standard records, 10 were not primary sources of data, 9 were not indexed in WOS, and 7 were not retrieved by the search). In this case, we obtained a 95% sensitivity, with a lower precision (3.7%) than the PubMed search.

2.4. Paper selection

All bibliographic records will be imported into the reference management software Endnote® X9, and the appropriate functions will be used to remove duplicates, and classify the papers by relevance, inclusion/exclusion status, reason for exclusion, and major features. Title and abstracts of each unique record will be screened for potential relevance. The full-text of all potentially relevant papers will be retrieved and assessed for compliance with the predefined inclusion/exclusion criteria. The results of abstract/title screening and of the full-text analysis will be recorded in a dedicated custom field of each article record. The results of the paper selection process will be graphically displayed in a flow-chart, and the list of papers excluded at the stage of full-text examination will be provided in a separate table, with indication of at least one reason for exclusion. Identified relevant reviews will be stored in a separate group, and will be used as a secondary sources of additional relevant papers missing from the searches through the main information sources.

The paper selection process will be performed in duplicate by two independent reviewers (MRS and SR), and possible disagreement will be

resolved by discussion with a third reviewer (OZ).

For papers reporting on multiple experiments, i.e., analyses of different endpoints or exposure conditions (in terms of frequency, waveform, exposure level or duration), data for all eligible experiments will be identified and extracted (see § 2.6).

2.5. Data extraction

The same investigators in charge of the paper selection will extract the relevant information regarding the experiments, based on the forms tested and refined during the pilot study (see § 2.7.1., and Supplementary Material 3).

The items below will be extracted and recorded in one or more of the systematic review databases (see Supplementary Material 3), as appropriate.

- *Paper ID* (common to all databases).
- *Bibliographic information*: first author, year of publication, journal, title, funding sources, declaration of conflict of interests, author contact information ("Bibliographic_info" database; one record per paper).
- *Experiment features (*"Experiment features" database*;* one record per experiment)
  - *Population*: cell type (with intrinsic information on cell status, healthy or cancerous, and species, humans or animals), source of cell cultures (primary cells or immortalized cell line).
  - *Exposure:* frequency (three subgroups: 100 kHz to $<$ 10 MHz; 10 MHz to $\leq$ 6 GHz; $>$6 GHz to $\leq$ 300 GHz); waveform (continuous or pulsed wave with signal type); exposure metric (E_{ind}, SAR, incident (S_{inc}) or absorbed (S_{ab}) power density); exposure level (three subgroups: below the exposure limits ($E_{ind} < 1.35x10^{-4}f$ V/m, or SAR $\leq$ 1 W/kg or $S_{ab} < 20$ W/m^2 or $S_{inc} < 10$ W/m^2), around the exposure limits ($E_{ind} = 1.35x10^{-4}f$ V/m, or 1 W/kg $<$ SAR $\leq$ 2 W/kg or $S_{ab} = 20$ W/m^2 or $S_{inc} = 10$ W/m^2), and above the exposure limits ($E_{ind} > 1.35x10^{-4}f$ V/m, SAR $>$ 2 W/kg or $S_{ab} > 20$ W/m^2 or $S_{inc} > 10$ W/m^2) (ICNIRP, 2020; Simko et al., 2016)); exposure duration (three subgroups: acute ($\leq$1h), long ($>$1 and $\leq$ 24 h), and chronic ($>$24 h) exposures (Simko et al., 2016)); exposure modality (continuous or intermittent); dosimetry analysis (yes/no); temperature control (yes/no); blinding to exposure allocation (yes/no); appropriate dosimetry (yes/no); field homogeneity (yes/no).
  - *Type of unexposed control*: negative; sham.
  - *Endpoint*: type of endpoint; assay procedure (test, timing of analysis post-exposure); appropriate biological method (yes/no;); positive control (yes/no); blinding of the analysis (yes/no);
- *Experiment results (*"Experiment results" database; one record per experiment):
  - Number of independent experiments or donors with replicates; number of events analysed.
  - *Statistical analysis:* statistical methods used to compare the occurrence of events between exposed and unexposed samples, or to describe the shape of the exposure–response relation, and p-value.
  - *Results* (multiple records per experiment)*:* qualitative description of results; results for unexposed and exposed samples.

In addition, the outcome of the risk-of-bias assessment, performed according to the procedure described in § 2.6, will be recorded in the "RoB data" database (one record per paper).

In the event that data cannot be clearly extracted from the papers, the authors will be contacted, and the date and results of the query will be registered in a dedicated form.

2.6. Risk of bias in individual studies

Due to the lack of a standardized approach to the Risk of Bias (RoB) assessment for *in vitro* studies (Rooney et al., 2016; Stephens et al., 2016), the RoB assessment tool presented in this protocol was developed based on the OHAT approach to the assessment of the study internal validity (NTP-OHAT, 2015; 2019), the recommendations by (Rooney, 2015), and taking into account the peculiar characteristics of RF-EMF exposures. According to (Rooney, 2015), the RoB assessment for *in vitro* investigations can be borrowed from the one developed by OHAT for experimental animal studies (NTP-OHAT, 2015), with some modifications. In brief, we will assess the RoB of all included studies based on the procedure outlined in Table 2, with additional details provided below.

- *Selection bias*: under the domains "Randomization of the exposure levels" and "Allocation concealment", all studies using homogeneous cell suspensions might be considered at definitely low risk of bias. As a matter of fact, the majority of *in vitro* studies investigating the genotoxicity use homogeneous cell suspensions, even when the cells are extracted from tissues (e.g. blood or sperm).
- *Confounding*: according to OHAT, confounding bias is not a relevant key-item for experimental animal studies, whereas the influence of particular confounding or effect-modifying factors may be assessed under "other potential threats to internal validity" (NTP-OHAT, 2015). We considered these indications also applicable to experimental *in vitro* studies.
- *Performance bias*: under this bias domain, we will address the presence of sham and/or incubator controls, as well as blinding of the research personnel to study groups during exposure assignment/administration. We will consider at definitely low risk of performance bias, studies including sham controls and handling controls in parallel to RF-exposed ones. We will take as indirect evidence of inclusion of sham controls instances where, based on description and images of the exposure set up, it may be argued that the control samples, not explicitly termed "sham" in the study report, did comply with the characteristics of a sham control. If sham controls were not included, but multiple exposure levels (e.g. multiple SAR values) were administered, the study will be considered at "probably low risk of perfomance bias", since it can be assumed that the environmental conditions were homogeneous across study groups. We will assess other potential threats to the uniformity of experimental conditions across groups by considering whether: i) all experimental samples (regardless exposure) were treated in the same environment; ii) identical exposure set-ups were used for all experimental samples (regardless exposure); iii) environmental factors or cell culture conditions (i.e., humidity, CO_2, cell culture media and any other non-treatment related experimental condition) had been monitored and proved identical across all experimental samples; iv) background EMF fields were measured (at least once before the experiment), or monitored continously during the experiment, and did not vary across the experimental samples (i.e., RF-exposed and control samples).
- *Attrition/Exclusion bias*: for the *in vitro* experimental studies considered herein, we will address loss of samples (proportion, and distribution across study groups) under the attrition/exclusion bias domain.
- *Detection bias*:
  - Confidence in exposure characterization. We will consider the following items as basic information required to assess the accuracy of the exposure set-up: full characterization of the RF signal (in terms of frequency, waveform, and modulation scheme); exhaustive description of the signal source (RF generator, with power stability and noise level), description of the measurement instruments (e.g. power meters, to check for the stability of the signal amplitude throughout the exposure) and probes (e.g., temperature sensors); measurement/calculation results provided with uncertainty budget; monitoring of all relevant technical and biological parameters throughout the experiment. We will assess the availability and appropriateness of methods of dosimetry analyses

Table 2
Customization of the OHAT RoB tool to the research topic. Questions 3 and 4 of the OHAT RoB tool do not apply to *in vitro* studies, and have not been considered. However, the original numbering of OHAT RoB has been kept.

Bias domain	Questions	Instructions for rating
Selection	1. *Was administered dose or exposure level adequately randomized?*	Unless there is direct or indirect evidence that exposure levels were not adequately randomized, it can be assumed that *in vitro* studies working with homogeneous cell suspensions can be considered, by default, at **Definitely low risk of bias (++)** for this source of possible selection bias.
	2. *Was allocation to study groups adequately concealed?*	Unless there is direct or indirect evidence that allocation to study group was not adequately concealed, it can be assumed that *in vitro* studies working with homogeneous cell suspensions can be considered, by default, at **Definitely low risk of bias (++)** for this source of possible selection bias.
Performance	5. *Were experimental conditions identical across study groups?*	**Definitely low (++):** There is direct evidence that both incubator and sham controls were used AND that they were handled in parallel to RF-exposed samples. **Probably low (+):** (There is indirect evidence that both incubator and sham control were used AND all study samples were handled in parallel) OR (only sham control was used AND was handled in parallel to RF-exposed sample), OR (incubator control AND multiple exposure levels were used AND all study samples were handled in parallel). **Probably high (-):** There is indirect evidence that only incubator controls were used OR sham controls were not handled in parallel to RF-exposed sample OR there is insufficient information on whether study groups were handled in parallel (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that only incubator control was used OR sham control was not handled in parallel to RF-exposed sample.
	6. *Were the research personnel blinded to the study group during the study?*	**Definitely low (++):** There is direct evidence that the research personnel were adequately blinded to study groups, and it is unlikely that they could have broken the blinding during the study. **Probably low (+):** There is indirect evidence that the research personnel were adequately blinded to study groups, and it is unlikely that they could have broken the blinding during the study, OR it is deemed that lack of adequate blinding during the study would not appreciably bias results. **Probably high (-):** There is indirect evidence that the research personnel were not adequately blinded to study groups, OR there is insufficient information provided about blinding to study groups during the study (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that the research personnel were not adequately blinded to study groups.
Attrition/ Exclusion	7. *Were endpoint data complete without attrition or exclusion from analysis?*	**Definitely low (++):** There is direct evidence that outcome data were complete OR that loss of samples was adequately addressed and reasons were documented when samples were removed from a study OR missing data have been imputed using appropriate methods. **Probably low (+):** There is indirect evidence that loss of samples was adequately addressed and reasons were documented when samples were removed from a study, OR it is deemed that the proportion lost would not appreciably bias results. **Probably high (-):** There is indirect evidence that loss of samples was unacceptably large and not adequately addressed, OR there is insufficient information provided about loss of samples (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that loss of samples was unacceptably large and not adequately addressed.
Detection	8. *Can we be confident in the exposure characterization?*	**Definitely low (++)**: There is direct evidence that the exposure was independently characterized, AND that exposure was consistently administered (i.e., with the same method and time-frame) across treatment groups. (For example, studies with well performed dosimetry analysis AND field inside samples homogeneously distributed will be rated ++). **Probably low (+)**: There is indirect evidence that the exposure was independently characterized, AND that exposure was consistently administered (i.e., with the same method and time-frame) across treatment groups. **Probably high (-):** There is indirect evidence that the exposure was assessed using poorly validated methods, OR there is insufficient information provided about the validity of the exposure

(*continued on next page*)

Table 2 (*continued*)

Bias domain	Questions	Instructions for rating
		assessment method (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that the exposure was assessed using poorly validated methods (*e.g.*, dosimetry not appropriately performed OR not performed at all).
	9. *Can we be confident in the outcome assessment?*	**Definitely low (++)**: There is direct evidence that the outcome was assessed using well-established methods (i.e., validated assays for each specific endpoint, correctly implemented as proven by the inclusion of positive controls), AND the outcome was assessed at the same length of time after initial exposure in all study groups, AND the outcome assessors were adequately blinded to the study groups, and it is unlikely that they could have broken the blinding prior to assessing outcomes. **Probably low** (+): There is indirect evidence that the outcome was assessed using acceptable methods (i.e., deemed of acceptable validity, although not the gold standard) AND assessed at the same length of time after initial exposure in all study groups OR it is deemed that the outcome assessment methods used would not appreciably bias results, AND there is indirect evidence that the outcome assessors were adequately blinded to the study group, and it is unlikely that they could have broken the blinding prior to assessing outcomes, OR it is deemed that lack of adequate blinding of outcome assessors would not appreciably bias results, which is more likely to apply to objective outcome measures. **Probably high (-):** There is indirect evidence that the outcome assessment method is an insensitive instrument, OR the length of time after initial exposure differed by study group, OR it was possible for outcome assessors to infer the study group prior to assessing outcomes without sufficient quality control measures, OR there is insufficient information provided about the validity of outcome assessment method and/or about blinding of outcome assessors (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that the outcome assessment method is an insensitive instrument (e.g. lack of positive control), OR the length of time after initial exposure differed by study group, OR there is direct evidence that outcome assessors were not blinded to the study groups, including no blinding or incomplete blinding without quality control measures.
Selective Reporting	10. *Were all measured endpoint conditions reported?*	**Definitely low (++)**: There is direct evidence that all of the study's measured endpoints (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported. **Probably low** (+): There is indirect evidence that all of the study's measured endpoints (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported, OR analyses that had not been planned in advance are clearly indicated as such and it is deemed that the unplanned analyses were appropriate and selective reporting would not appreciably bias results. This would include outcomes reported with insufficient detail such as only reporting results that were statistically significant (or not). **Probably high (-):** There is indirect evidence that not all of the study's measured endpoints (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported, OR there is indirect evidence that unplanned analyses were included that may appreciably bias results, OR there is insufficient information provided about selective outcome reporting (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that not all of the study's measured endpoints (primary and secondary) outlined in the protocol, methods, abstract, and/or introduction (that are relevant for the evaluation) have been reported.
Other Bias	11. *Were there no other potential threats to internal validity (e.g., statistical methods were appropriate and researchers adhered to the study protocol)?*	
	11.a Were statistical methods appropriate?	**Definitely low (++)**: There is direct evidence that, for each experimental protocol, appropriate statistics AND an adequate number of independent experiments were performed. **Probably**

(*continued on next page*)

Table 2 (*continued*)

Bias domain	Questions	Instructions for rating
		low (+): There is indirect evidence that appropriate statistics AND an adequate number of independent experiments were performed. **Probably high (-):** There is indirect evidence that unappropriate statistics OR an adequate number of independent experiments were performed OR there is insufficient information provided about appropriateness of statistical methods and/or number of independent experiments (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that unappropriate statistics OR an insufficient number of independent experiments were performed OR no statistical analysis was performed.
	11.b Did the study design or analysis account for important confounding and modifying variables (including unintended co-exposures) in experimental studies?	**Definitely low (++):** There is direct evidence that sample temperature was adequately monitored during exposure AND possible increases promptly counteracted. **Probably low** (+): There is indirect evidence that sample temperature was adequately monitored during exposure AND possible increases promptly counteracted. **Probably high (-):** There is indirect evidence that sample temperature was not adequately monitored during exposure AND possible increases were not promptly counteracted OR there is insufficient information provided about temperature control (record "NR" as basis for answer). **Definitely high (− −):** There is direct evidence that sample temperature was not adequately monitored during exposure AND possible increases were not counteracted.

N.R.: Not Reported.

(i.e., estimates of E-field/SAR/PD induced in the samples by experimental and/or numerical techniques), as well as the documented homogeneity of the exposure within samples. For example, estimates of SAR from measurements of the electric field in absence of the sample will not be considered an appropriate dosimetry method because such procedure does not take into account that the sample significantly perturbs the electric field in the RF range. On the other hand, estimates of SAR from computation of electric field in the sample or by calorimetric measurements will be considered appropriate. A disuniformity degree of the electric field distribution within the sample around 30% will be considered a good quality standard for exposure (Kuster and Schonborn, 2000).

- Confidence in outcome assessment will be assessed with reference to the suitability of the endpoint-specific assay (e.g., the method by (Fenech, 2000; 2007) for the micronucleus assay; that by (von Recklinghausen et al., 2007) for the chromosomal aberration assay; Comet Assay for the assessment of single and double strand breaks (Singh et al., 1988)), and its correct implementation as documented using positive controls. The latter provide evidence of controlled experimental conditions, and assurance that the assay methodology is responding adequately to a well-known agent (Simko et al., 2016).

- *Selective reporting*. Under this domain, as far as deductible from the information available in the study reports, we will assess whether reporting of all endpoints relative to the analysed samples, and findings from the analysed exposure conditions, is complete and independent of the magnitude and direction of the results.
- *Other bias*. The appropriateness of the statistical methods will be assessed with reference to the appropriateness of the tests given the data (e.g., application of parametric tests to normally distributed data) (Ceppi et al., 2011). Temperature increase during exposure is the more relevant confounding factor in the assessment of biological effects of RF-EMF. Therefore, temperature inside the samples must be monitored continously during treatment (or, at least, in preliminary experiments aimed at characterizing the temperature profile), using adequate instruments (e.g., fiber optic thermometers, infrared cameras, or other tools that do not perturb the field). If the sample is heated during exposure (e.g., when using SAR values above the exposure limits), specific measures to counteract such heating must be adopted (e.g. circulation of cooling water). During RF exposure, subtle temperature variations (ΔT) may occur due to the finite heat capacity of any material system. Therefore, we will consider a $\Delta T = 1$ °C as the threshold for temperature variation above which the heating effect has to be counteracted. However, if thermal increases are not counteracted but the study design includes a temperature control (i.e. a sample subjected to the same temperature increase induced by different methods, such as thermostatic water/oil-bath, or DC current), this will be considered as an appropiate method to reduce confounding risk of bias related to temperature (Michaelson and Elson, 1986).

The internal validity of the included studies will be assessed in duplicate, by two investigators (SR and MRS) per study, and possible disagreement will be resolved by discussion with a third investigator (OZ). For papers co-authored by SR, MRS or OZ, the RoB assessment will be performed by SL and MB.

For each bias domain, the experiment-specific potential for bias will be rated as: definitely low (++); probably low (+); probably high (-); or definitely high (–) (NTP-OHAT, 2015). A specific form was designed to record the assessors' RoB rating along with supporting information (see § 2.6), and the HAWC platform will be used to manage the RoB assessment process and generate RoB heatmaps.

We will follow the OHAT optional 3-level tiering of the quality of individual studies, based on summary assessments of RoB (NTP-OHAT, 2019). As recommended by OHAT, we selected four key elements, as particularly relevant to our research topic: 1) Identity of experimental conditions across study groups; 2) Confidence in exposure assessment; 3) Confidence in outcome assessment; 4) Temperature control (under the "Other bias" domain).

The criteria used for the tiering are described in Table 3.

2.6.1. Pilot study of the data extraction RoB assessment tools

The results of a pilot testing of the RoB assessment and quality tiering performed on five, randomly selected, relevant papers are briefly summarized below (see Supplementary Material 2 for further details).

Data extraction from the graphics was carried out using the WebplotDigitizer application (https://automeris.io/WebPlotDigitizer/). Fig. 1 shows the RoB rating and the classification of the studies by quality categories.

Table 3
Criteria for classifying the studies by quality categories (Tiers 1, 2, 3), based on results of the RoB assessment.

Quality category	Classification criteria
Tier 1 (high quality)	"++" (definitely low) or "+" (probably low) risk of bias in all key domains AND "++" (definitely low) or "+" (probably low) risk of bias for ≥ 50% of the other domains
Tier 2 (moderate quality)	Study does not meet criteria for placement in the 1st or 3rd tier
Tier 3 (low quality)	"- -" (definitely high) or "-" (probably high) risk of bias in all key domains AND "- -" (definitely high) or "-" (probably high) risk of bias for ≥ 50% of the other domains

2.7. Data analysis and synthesis

Heterogeneity within the available evidence will determine the type of evidence synthesis that is appropriate.

The characteristics of the included studies, and findings from investigation subsets not amenable to quantitative synthesis, will be summarized in tables and graphical displays illustrating the direction of effects and in a narrative synthesis of the available evidences, as suggested by (NTP-OHAT, 2019).

A *meta*-analysis will be considered for subgroups of studies homogeneous in terms of endpoint, cell type, and exposure characteristics, with comparable measures of effect, and of sufficient size (tentatively at least 10 studies) to conduct meaningful analyses.

The aim of the *meta*-analysis will be to calculate a precision-weighted average measure of effect, and to assess the consistency of results across studies and between study groups differing by relevant experimental design features and susceptibility to bias. The effect size will be calculated based on the indications by (Vesterinen et al., 2014). When results of the experiments are reported as event data (binary outcome), odd ratios will be calculated. In case of continuous outcomes, the mean, standard deviation (SD) and sample size (n) will be extracted to calculate standardized mean differences (SMD) (Vesterinen et al., 2014). A specialized software, RevMan (Cochrane), will be used for the preliminary data transformation and the *meta*-analysis. Individual effect sizes will be pooled to obtain an overall effect size and 95% confidence interval using the fixed-effects inverse variance model (Deeks et al., 2019).

Statistical heterogeneity will be quantified by means of the I^2 statistics, with values of I^2 of 25%, 50%, 75% considered as an indication of low, moderate and high degrees of heterogeneity, respectively. Between studies variance will be assessed by the tau-squared statistics (τ^2).

The influence of each individual study in the *meta*-analysis will be investigated by omitting one study at a time, and re-calculating the summary estimates (leave-one-out method) (Vesterinen et al., 2014).

Sensitivity analyses will be conducted by excluding studies at high RoB.

If there will be sufficiently numerous and homogeneous (see above) subgroups of experiments investigating more than one level or duration of exposure, a *meta*-regression or a dose–response *meta*-analysis will be

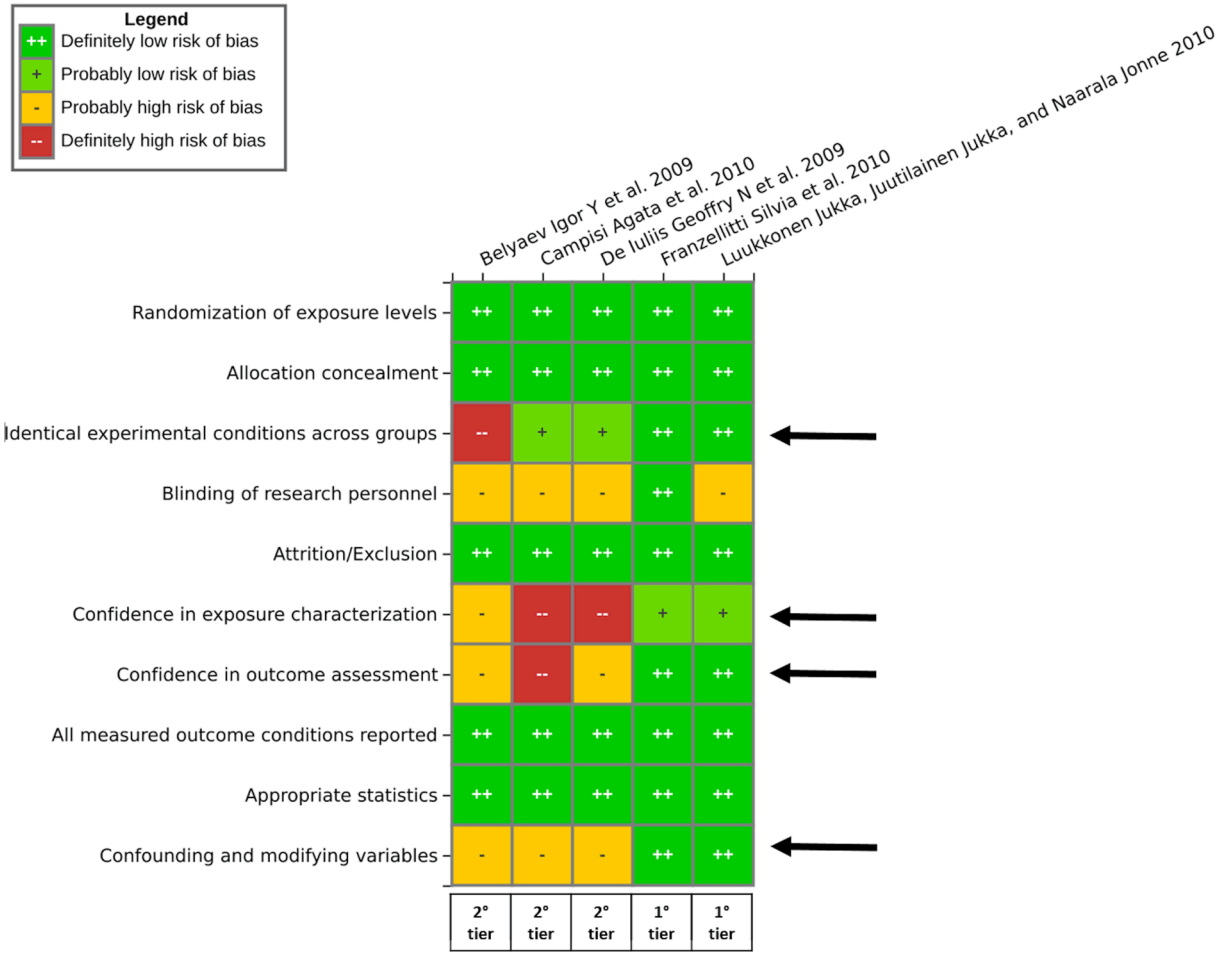

Fig. 1. Pilot risk of bias and quality assessment. Risk-of-bias ratings and placement in one out of the three quality categories (3-tier approach) for 5 of the included studies. Black arrows indicate key risk-of-bias items for quality assessment.

carried out to assess the heterogeneity across experiments in the shape of the exposure–response relation.

Publication bias will be assessed by funnel plots and the "trim and fill" correction (Deeks et al., 2019).

2.8. Evidence appraisal

We will assess the confidence in the body of evidence following the OHAT guidelines for animal studies (NTP-OHAT, 2019), which are based on the GRADE approach. Four descriptors are used to indicate the level of confidence in the body of evidence:

- High Confidence (++++) - The true effect is highly likely to be reflected in the apparent relationship.
- Moderate Confidence (+++) - The true effect may be reflected in the apparent relationship.
- Low Confidence (++) - The true effect may be different from the apparent relationship.
- Very Low Confidence (+) - The true effect is highly likely to be different from the apparent relationship.

The body of evidence is given an initial confidence rating based on the study design (set to high for experimental studies (NTP-OHAT, 2019)), which can be upgraded or downgraded depending on a number of factors, assessed across studies.

Factors for downgrading include: risk-of-bias (across studies); indirectness, inconsistency, imprecision, publication bias. Factors for upgrading include: magnitude of average effect, dose–response, and consistency of results across species/models.

The evidence appraisal will be conducted at the endpoint level, and then integrated across endpoints. To assess confidence in the body of evidence we will use the judgement described below, and the results of the appraisal will be summarized in an evidence profile table (see Supplementary Material 4). Risk of bias across studies will be assessed based on the aforementioned three-level tiering of study quality. Downgrading (one level) will be applied when most of the studies composing the body of evidence are classified in the tier-3 level.

We will downgrade (one level) for inconsistency in presence of a large variability in the direction or magnitude of the individual effect estimates for comparable measures of association, that cannot be explained by biological or methodological factors (i.e cell model, endpoint assessment method, funding source/conflict of interest, and risk of bias). The evidence will be downgraded 1 level in the presence of serious/very serious unexplained inconsistency (i.e., point estimates vary/vary widely; confidence intervals show minimal/no overlap; statistical heterogeneity has low p-value ($p \leq 0.1$); $I^2 > 50\%/75\%$).

Extrapolation of findings from isolated biological systems to living organisms is challenging, and *in vitro* mechanistic studies can only provide supportive evidence on potential cancer effects in humans (Guyatt et al., 2011). These inherent limitations of the body of evidence will be considered under the grading domains of indirectness. Herein, we will apply the following judgement rules:

1. *Relevance of the cell model to humans*: exposure-induced genotoxic effects in primary cells will be assigned greater confidence than similar findings detected in immortalized cells, and exposure-induced genotoxic effects in human cells will be assigned greater confidence than similar findings detected in non-human cells.
2. *Endpoint's predictivity of long-term DNA damage*: exposure-induced increases in biomarkers of irreversible DNA damage will be assigned greater confidence than similar findings for biomarkers of reparable damage (Krewski et al., 2019).

We will use 95% confidence intervals as the primary method to assess imprecision for ratio measures of effect; we will downgrade (one level) when the ratio of the upper to lower 95% confidence limits of a *meta*-risk estimate or of most studies is > 10. For comparisons of average values of continous variables across study groups, we will use the standard deviation (SD) of the standardized mean difference as a measure of imprecision, and will downgrade (one level) when the SD is larger than the mean. We will characterize publication bias as "undetected" (no downgrade) or "strongly suspected" (one level dowgrade).

Confidence in evidence will be upgraded (one level) in presence of large magnitude of effect. To determine whether the magnitude of the effect is large, we will consider, for each endpoint separetely, the background prevalence or rate for that effect, the species and dose range employed. Regarding dose–response, we note that the OHAT refers to either monotonic increases of the effect with increasing exposure level, or to non-monotonic gradients. Such a definition fits well with the subject of our systematic review because, due to the absence of widely accepted interaction mechanisms between RF-EMF and biosystems other than thermal effects, it is not clear which is the expected shape of the RF-genotoxicity relationship (Postow and Swicord, 1986). We will upgrade (one level) the confidence in the body of evidence in the presence of dose–response gradients consistently reported by studies investigating primary endpoints. Lastly, we will upgrade (one level) the confidence in evidence for consistency of results across species/models, exposure types (frequency range, SAR range, continuous/pulsed waveforms, continuous/intermittent exposure), and endpoints.

3. Concluding remarks

We presented the protocol for a systematic review of studies investigating possible genotoxic effects induced by RF-EMF in *in vitro* mammalian cell models. The protocol was developed in the framework of an Italian systematic review of studies on the potential carcinogenicity of RF-EMF which envisages the integration of the epidemiology, *in vivo* and *in vitro* evidence streams.

To the best of the authors' knowledge, this is the first protocol for a systematic review of mechanistic studies with an *in vitro* RF-EMF exposure regimen, with specific indications of criteria for papers inclusion and risk of bias assessment. This might also be the first protocol for a systematic review of *in vitro* studies only, contrary to other protocol papers for systematic reviews of both *in vivo* and *in vitro* studies (Matta et al., 2019).

There are no standardized approaches to the assessment of RoB for mechanistic studies with an *in vitro* exposure regimen. There are also no broadly accepted frameworks for reaching confidence ratings for use of mechanistic data in decision making, and thus there is a need for research efforts to gain experience and develop methods in this area (Rooney et al., 2016).

We followed the methodological approach suggested by guidelines and recommendations for hazards assessment mainly focussing on chemical agents. Consequently, we had to adapt the data extraction forms, and especially the RoB tool, to the peculiarities of RF-EMF, including exposure characteristics and dosimetry, and type of comparators, that have a large impact on the study quality (Simko et al., 2016).

We validated the search strategy by a record-linkage to a personal literature database including also references quoted in several recent authoritative assessment of health hazards from exposure to RF-EMF. Moreover, we pilot-tested the method and tools for data extraction, RoB assessment and study quality classification.

Possible deviations from this protocol will be reported in the systematic review paper with justifications for the changes.

Funding

This work was supported by the Italian Workers' Compensation Authority (INAIL), a public, no-profit entity, and performed in the framework of the CRA with the Istituto Superiore di Sanità (ISS) "BRiC 2018/06–Scientific evidence on the carcinogenicity of radiofrequency electromagnetic fields" (ISS PI Mauro Biffoni; INAIL representative

Carlo Grandi); grant code I85B19000120005. INAIL had no role in the protocol development.

SL withdrew from the BRiC 2018/06 project on 1 April 2020.

Author Contributions (based on CrediT – Contributor Roles Taxonomy)

Conceptualization (SR, OZ, MRS, SL); Data curation (SR, OZ, MRS); Funding acquisition (MRS, SL); Methodology (SR, OZ, MRS, AS, SL, MB); Writing - original draft, Writing - review & editing (SR, OZ, MRS, AS, SL, MB). Guarantor of the review (SR).

CRediT authorship contribution statement

Stefania Romeo: Conceptualization, Data curation, Methodology, Writing - original draft, Writing - review & editing. **Olga Zeni:** Conceptualization, Data curation, Methodology, Writing - original draft, Writing - review & editing. **Anna Sannino:** Methodology, Writing - original draft, Writing - review & editing. **Susanna Lagorio:** Conceptualization, Funding acquisition, Methodology, Writing - original draft, Writing - review & editing. **Mauro Biffoni:** Methodology, Writing - original draft, Writing - review & editing. **Maria Rosaria Scarfì:** Conceptualization, Data curation, Funding acquisition, Methodology, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envint.2021.106386.

References

AGNIR. Health effects from radiofrequency electromagnetic fields. London: Health Protection Agency; 2012.

ANSES. Radiofréquences et santé. Mise à jour de l'expertise. Maisons-Alfort: Agence Nationale de Sécurité Sanitaire de l'Alimentation de l'Environnement et du Travail 2013.

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 (300Hz to 1MHz). Environ. Res. 171, 247–259.

Ceppi, M., Gallo, F., Bonassi, S., 2011. Study design and statistical analysis of data in human population studies with the micronucleus assay. Mutagenesis 26, 247–252.

Deeks, J.J.; Higgins, J.P.T., Altm, D.G. Analysing data and undertaking meta-analyses. in: Higgins J., Thomas J., Chandler J., Cumpston M., Li T., Page M., Welch V., eds. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane; 2019.

Driessen, S.; Dechent, D.; Gollnick, F.; Gräfrath, D.; Schmid, G.; Cecil, S.; Hirtl, R.; Schneeweiß, P.; Klemcke, S.; Janke, G.; Trübswetter, A. Evaluierung des EMF-Portals und Ableitung von Erkenntnissen und Empfehlungen für dessen weitere Gestaltung. Abschlussbericht zum Vorhaben 2017.

Fenech, M., 2000. The in vitro micronucleus technique. Mutat. Res. 455, 81–95.

Fenech, M., 2007. Cytokinesis-block micronucleus cytome assay. Nat. Protoc. 2, 1084–1104.

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

Guyatt, G.H.; Oxman, A.D.; Kunz, R.; Woodcock, J.; Brozek, J.; Helfand, M.; Alonso-Coello, P.; Falck-Ytter, Y.; Jaeschke, R.; Vist, G.; Akl, E.A.; Post, P.N.; Norris, S.; Meerpohl, J.; Shukla, V.K.; Nasser, M.; Schunemann, H.J.; Group, G.W. GRADE guidelines: 8. Rating the quality of evidence–indirectness. J Clin Epidemiol 2011;64: 1303-1310.

Halgamuge, M.N., Skafidas, E., Davis, D., 2020. A meta-analysis of in vitro exposures to weak radiofrequency radiation exposure from mobile phones (1990–2015). Environ. Res. 184, 109227.

Higgins, J.P.T.; Thomas, J.; Chandler, J.; Cumpston, M.; Li, T.; Page, M.J.; Welch, V.A. Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane 2019.

IARC, 2013. Non ionizing radiation, Part 2: Radiofrequency electromagnetic fields. IARC Monographs on the evaluation of carcinogenic risks to humans. IARC, Lyon, France.

ICNIRP, 2020. Guidelines for limiting exposure to electromagnetic fields (100 kHz to 300 GHz). Health Phys. 118, 483–524.

Krewski, D., Bird, M., Al-Zoughool, M., Birkett, N., Billard, M., Milton, B., Rice, J.M., Grosse, Y., Cogliano, V.J., Hill, M.A., Baan, R.A., Little, J., Zielinski, J.M., 2019. Key characteristics of 86 agents known to cause cancer in humans. J. Toxicol. Environ. Health B Crit. Rev. 22, 244–263.

Kuster, N., Schonborn, F., 2000. Recommended minimal requirements and development guidelines for exposure setups of bio-experiments addressing the health risk concern of wireless communications. Bioelectromagnetics 21, 508–514.

Manna, D., Ghosh, R., 2016. Effect of radiofrequency radiation in cultured mammalian cells: A review. Electromagn. Biol. Med. 35, 265–301.

Matta, K., Ploteau, S., Coumoul, X., Koual, M., Le Bizec, B., Antignac, J.P., Cano-Sancho, G., 2019. Associations between exposure to organochlorine chemicals and endometriosis in experimental studies: A systematic review protocol. Environ. Int. 124, 400–407.

Meltz, M.L. Radiofrequency exposure and mammalian cell toxicity, genotoxicity, and transformation. Bioelectromagnetics 2003, Suppl 6:S196-213.

Michaelson, S., Elson, E., 1986. Interaction of nonmodulated and pulse modulated radio frequency fields with living matter: Experimental results. In: Polk, C., Postow, E. (Eds.), Handbook of biological effects of electromagnetic fields. CRC Press, Boca Raton.

Moher, D.; Shamseer, L.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A.; Group, P.-P. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev, ;4:1, 2015.

Morgan, R.L.; Beverly, B.; Ghersi, D.; Schunemann, H.J.; Rooney, A.A.; Whaley, P.; Zhu, Y.G.; Thayer, K.A.; Group, G.W. GRADE guidelines for environmental and occupational health: A new series of articles in Environment International. Environ Int 2019;128:11-12.

NTP-OHAT. OHAT Risk of Bias Rating Tool for Human and Animal Studies. Office of Health Assessment and Translation. 2015.

NTP-OHAT. Handbook for Conducting a Literature-Based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. National Toxicology Program - Office of Health Assessment and Translation; 2019.

NTP-ORoC. Standard Search Strings for Literature Database Searches: Appendix to the Draft Handbook for Preparing Report on Carcinogens Monographs. in: Carcinogens N.T.P.-O.o.t.R.o., ed; 2016.

Postow, E., Swicord, M.L., 1986. Modulated fields and window effects. In: Polk, C., Postow, E. (Eds.), Handbook of biological effects of electromagnetic fields. CRC Press, Roca Raton, FL.

Rooney, A.A. Extending a Risk-of-bias approach to address in vitro studies 2015. https://ofmpub.epa.gov/eims/eimscomm.getfile?p_download_id=526750.

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, 711–718.

Rooney, A.A., Cooper, G.S., Jahnke, G.D., Lam, J., Morgan, R.L., Boyles, A.L., Ratcliffe, J. M., Kraft, A.D., Schunemann, H.J., Schwingl, P., Walker, T.D., Thayer, K.A., Lunn, R. M., 2016. How credible are the study results? Evaluating and applying internal validity tools to literature-based assessments of environmental health hazards. Environ. Int. 92–93, 617–629.

Samet, J.M., Chiu, W.A., Cogliano, V., Jinot, J., Kriebel, D., Lunn, R.M., Beland, F.A., Bero, L., Browne, P., Fritschi, L., Kanno, J., Lachenmeier, D.W., Lan, Q., Lasfargues, G., Curieux, F.L., Peters, S., Shubat, P., Sone, H., White, M.C., Williamson, J., Yakubovskaya, M., Siemiatycki, J., White, P.A., Guyton, K.Z., Schubauer-Berigan, M.K., Hall, A.L., Grosse, Y., Bouvard, V., Benbrahim-Tallaa, L., Ghissassi, F.E., Lauby-Secretan, B., Armstrong, B., Saracci, R., Zavadil, J., Straif, K., Wild, C.P., 2019. The IARC Monographs: Updated procedures for modern and transparent evidence synthesis in cancer hazard identification. J. Natl. Cancer Inst.

Sampson, M., McGowan, J., 2011. Inquisitio validus Index Medicus: A simple method of validating MEDLINE systematic review searches. Res. Synth. Meth. 2, 103–109.

Scenihr. Potential health effects of exposure to electromagnetic fields (EMF) Luxembourg: European Commission; 2015.

Shamseer, L.; Moher, D.; Clarke, M.; Ghersi, D.; Liberati, A.; Petticrew, M.; Shekelle, P.; Stewart, L.A.; Group, P.-P. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. BMJ 2015;350: g7647.

Shapiro, A.J., Antoni, S., Guyton, K.Z., Lunn, R.M., Loomis, D., Rusyn, I., Jahnke, G.D., Schwingl, P.J., Mehta, S.S., Addington, J., Guha, N., 2018. Software Tools to Facilitate Systematic Review Used for Cancer Hazard Identification. Environ. Health Perspect. 126, 104501.

Simko, M., Remondini, D., Zeni, O., Scarfi, M.R., 2016. Quality Matters: Systematic Analysis of Endpoints Related to "Cellular Life" in Vitro Data of Radiofrequency Electromagnetic Field Exposure. Int. J. Environ. Res. Public Health 13.

Singh, N.P., McCoy, M.T., Tice, R.R., Schneider, E.L., 1988. A simple technique for quantitation of low levels of DNA damage in individual cells. Exp. Cell Res. 175, 184–191.

Smith, M.T.; Guyton, K.Z. Identifying carcinogens from 10 key characteristics. A new approach based on mechanisms. in: Wild C.P., Weiderpass E., B.W. S., eds. World Cancer Report: Cancer Research for Cancer Prevention. Lyon: IARC Press; 2020.

Smith, M.T., Guyton, K.Z., Gibbons, C.F., Fritz, J.M., Portier, C.J., Rusyn, I., DeMarini, D. M., Caldwell, J.C., Kavlock, R.J., Lambert, P.F., Hecht, S.S., Bucher, J.R., Stewart, B. W., Baan, R.A., Cogliano, V.J., Straif, K., 2016. Key Characteristics of Carcinogens as a Basis for Organizing Data on Mechanisms of Carcinogenesis. Environ. Health Perspect. 124, 713–721.

Stephens, M.L., Betts, K., Beck, N.B., Cogliano, V., Dickersin, K., Fitzpatrick, S., Freeman, J., Gray, G., Hartung, T., McPartland, J., Rooney, A.A., Scherer, R.W., Verloo, D., Hoffmann, S., 2016. The Emergence of Systematic Review in Toxicology. Toxicol. Sci. 152, 10–16.

Verschaeve, L. Evaluations of International Expert Group Reports on the Biological Effects of Radiofrequency Fields. in: Eksim A., ed. Wireless Communications and Networks - Recent Advances: IntechOpen; 2012.

Verbeek, Jos, Oftedal, Gunhild, Feychting, Maria, et al., 2021. Prioritizing health outcomes when assessing the effects of exposure to radiofrequency electromagnetic fields: A survey among experts. Environment International 146. https://doi.org/10.1016/j.envint.2020.106300.

Vesterinen, H.M., Sena, E.S., Egan, K.J., Hirst, T.C., Churolov, L., Currie, G.L., Antonic, A., Howells, D.W., Macleod, M.R., 2014. Meta-analysis of data from animal studies: a practical guide. J. Neurosci. Meth. 221, 92–102.

Vijayalaxmi; Prihoda, T.J. Genetic damage in mammalian somatic cells exposed to radiofrequency radiation: a meta-analysis of data from 63 publications (1990-2005). Radiat Res 2008, 169, 561-574.

Vijayalaxmi; Prihoda, T.J. Genetic damage in human cells exposed to non-ionizing radiofrequency fields: a meta-analysis of the data from 88 publications (1990-2011). Mutat Res 2012, 749:1-16.

Vijayalaxmi; Prihoda, T.J. Comprehensive Review of Quality of Publications and Meta-analysis of Genetic Damage in Mammalian Cells Exposed to Non-Ionizing Radiofrequency Fields. Radiat Res 2019;191, 20-30.

von Recklinghausen, U.; Johannes, C.; Riedel, L.; Obe, G. Aberration patterns and cell cycle progression following exposure of lymphocytes to the alkilating agent trenimon. in: Vijayalaxmi G.O.a., ed. Chromosomal Aberrations - Methods, results and Importance in human health. Heidelberg: Springer-Verlag; 2007.

Vijayalaxmi and Scarfi, M.R., 2014. International and national expert group evaluations: biological/health effects of radiofrequency fields. Int J Environ Res Public Health 11, 9376–9408.

Waffenschmidt, S., Hermanns, T., Gerber-Grote, A., Mostardt, S., 2017. No suitable precise or optimized epidemiologic search filters were available for bibliographic databases. J. Clin. Epidemiol. 82, 112–118.

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., Schunemann, 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.

WHO. Radio Frequency fields - Relaunch Call for Expressions of Interest for systematic reviews (2020). 2020. https://www.who.int/peh-emf/research/rf_ehc_page/en/index1.html.

Zeni, O.; Scarfi, M.R. Experimental requirements for in vitro studies aimed to evaluate the biological effects of radiofrequency radiation, microwave materials characterization. in: Costanzo S., ed. Microwave Materials Characterization. Rijeka, Croatia: InTech; 2012.