Journal Pre-proofs

Effects of radiofrequency electromagnetic field exposure on cancer in laboratory animal studies, a systematic review

Meike Mevissen, Angélique Ducray, Jerrold M. Ward, Annette Kopp-Schneider, James P. McNamee, Andrew W. Wood, Tania M. Rivero, Kurt Straif

PII: S0160-4120(25)00233-8
DOI: https://doi.org/10.1016/j.envint.2025.109482
Reference: EI 109482

To appear in: *Environment International*

Received Date: 15 February 2024
Revised Date: 25 March 2025
Accepted Date: 16 April 2025

Please cite this article as: M. Mevissen, A. Ducray, J.M. Ward, A. Kopp-Schneider, J.P. McNamee, A.W. Wood, T.M. Rivero, K. Straif, Effects of radiofrequency electromagnetic field exposure on cancer in laboratory animal studies, a systematic review, *Environment International* (2025), doi: https://doi.org/10.1016/j.envint.2025.109482

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

© 2025 The Author(s). Published by Elsevier Ltd.

Effects of radiofrequency electromagnetic field exposure on cancer in laboratory animal studies, a systematic review

Meike Mevissen[a,*,1,meike.mevissen@unibe.ch], Angélique Ducray[a,1], Jerrold M. Ward[b], Annette Kopp-Schneider[c], James P. McNamee[d], Andrew W. Wood[e], Tania M. Rivero[f], Kurt Straif[g,h]

[a]Veterinary Pharmacology & Toxicology, Department of Clinical Research and Veterinary Public Health (DCR-VPH), University of Bern, Bern, Switzerland

[b]Global VetPathology, Montgomery Village, Maryland, USA

[c]Division of Biostatistics, German Cancer Center, Heidelberg, Germany

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

[e]Department of Health Sciences and Statistics, Swinburne University of Technology, Hawthorn, Australia

[f]Medical Library, University Library, University of Bern, Bern, Switzerland

[g]ISGlobal, Barcelona, Spain

[h]Boston College, MA, USA

[*]Corresponding author.

[1]Shared first authorship.

Journal Pre-proofs

Abstract

Background

More than ten years ago, the World Health Organization's (WHO) International Agency for Research on Cancer (IARC) published a monograph concluding there was limited evidence in experimental animals for carcinogenicity of Radio Frequency Electromagnetic Field (RF EMF).

Objective

The objective of this review was to systematically evaluate the effects of RF EMF exposure on cancer in experimental animals.

Methods

Eligibility criteria: Based on pre-established Populations, Exposures, Comparators, Outcomes, and Study Type (PECOS) criteria, studies in experimental animals of the following study types were included: chronic cancer bioassays, initiation-(co-)promotion studies, and studies with tumor-prone animals.

Information sources: MEDLINE (PubMed), Science Citation Index Expanded and Emerging Sources Citation Index (Web of Science), and the EMF Portal.

Data abstraction and synthesis: Data are publicly available online as interactive visuals with downloadable metadata. *We* adapted the risk-of-bias (RoB*)* tool developed by Office of Health Assessment and Translation (OHAT) to include considerations pertinent to the evaluation of RF EMF exposure and cancer bioassays. Study sensitivity was assessed with a tool adopted from the Report on Carcinogens (RoC). We synthesized studies using a narrative approach. Effect size was calculated as the 1% Bayesian Average benchmark dose (BMD) of a respective study when dose-response or a trend was identified (see BMDAnalysisSupplementaryMaterial) (Supplement 1).

Evidence Assessment: Certainty of the evidence (CoE) was assessed using the Grading of Recommendations, Assessment, Developing and Evaluations (GRADE) approach, as refined by OHAT. Evidence from chronic cancer bioassays was considered the most directly applicable to evaluation of carcinogenicity.

Results

We included 52 studies with 20 chronic bioassays No studies were excluded based on risk of bias concerns. Studies were not considered suitable for meta-analysis due to heterogeneity in study design, species, strain, sex, exposure characteristics, and cancer outcome. No or minimal evidence of RF EMF exposure-related cancer outcomes was found in most systems or organs in any study (these included gastrointestinal/digestive, kidney, mammary gland, urinary, endocrine, musculoskeletal, reproductive, and auditory).

For lymphoma (18 studies), with 6 chronic bioassays (1,120 mice, 1,780 rats) inconsistency between two chronic bioassays was not plausibly explainable, and the CoE for lymphoma was rated ‘moderate’.

For brain tumors (20 studies), including 5 chronic bioassays (1,902 mice, 6,011 rats), an increase in glial cell-derived neoplasms was reported in two chronic bioassays in male rats. The CoE for an increased risk in glioma was judged as high. The BMD analysis was statistically significant for only one study and the BMD was 4.25 (95% CI 2.70, 10.24).

For neoplasms of the heart (4 chronic bioassays with 6 experiments), 3 studies were performed in rats (~2,165 animals), and 1 in mice (~720 animals). Based on 2 bioassays, statistically significant increases in malignant schwannomas was judged as high CoE for an increase in heart schwannomas in male rats. The BMDs from the two positive studies were 1.92 (95%CI 0.71, 4.15) and 0.177 (95%CI 0.125, 0.241), respectively.

Twelve studies reported neoplasms in the adrenal gland (5 chronic bioassays). The CoE for an increased risk in pheochromocytoma was judged as moderate. None of these findings were dose-dependent when compared to the sham controls.

Sixteen studies investigated tumors of the liver with 5 of these being chronic bioassays. The CoE was evaluated as moderate for hepatoblastomas.

For neoplasms of the lung (3 chronic bioassays), 8 studies were conducted in rats (~1,296 animals) and 23 studies in mice (~2,800 animals). In one chronic bioassay, a statistically significant positive trend was reported for bronchoalveolar adenoma or carcinoma (combined), which was rated as moderate CoE for an increase in lung neoplasms with some evidence from 2 initiation-(co-)promotion studies.

Discussion

Meta-analysis was considered inappropriate due to the heterogeneity in study methods. The GRADE/OHAT CoE framework has not been frequently applied to animal studies and experience

to date suggests refinements are needed. We deferred to standard methods in environmental health where CoE is framed in the context of strength of the evidence providing positive support for carcinogenicity. High CoE can be interpreted as the true effect is highly likely to be reflected in the apparent relationship. Moderate CoE indicates the true effect may be reflected in the apparent relationship. Cancer bioassays conducted in experimental animals are commonly used to identify potential human carcinogens. We note that the two tumor types with high CoE in animals in this systematic review are the same as those identified with limited evidence in humans by the IARC Working Group. However, even in cases where the animal evidence demonstrates high CoE, the extrapolation of risk from cancer bioassays to humans is particularly complex for RF EMF. Without a better understanding of the mechanism of the carcinogenicity of RF-EMF, the choice of exposure metric for risk extrapolation (whole body versus localized), intensity or cumulative exposure, whether or not a monotonic dose-response holds for carcinogenic effects, and whether SAR is the appropriate dose metric for adverse effects induced by RF-EMF may be critical.

Other

This review was partially funded by the WHO radioprotection programme.

The protocol for this review was registered in Prospero reg. no. CRD42021265563 and published in Environment International 2022 (Mevissen et al. 2022).

Keywords: Radiofrequency electromagnetic fields; Carcinogenicity; Bioassay; Animal studies; Carcinogenesis, toxicity, systematic review; Radiofrequency exposure

1. Introduction

1.1. Rationale

The use of radiofrequency electromagnetic fields (RF EMF; frequencies 100 kHz to 300 GHz) has steadily increased in a variety of fields including telecommunications, medicine, industry, domestic appliances, security, and navigation. Environmental and occupational exposure to RF EMF is expected to increase further, especially in telecommunications with the roll-out of 5G added new frequencies.

Public concern has been raised as to whether exposure to RF EMF associated with these new technologies might pose a health risk for humans. Experimental animal and epidemiological research have been conducted to investigate if RF EMF exposure is carcinogenic. In a first evaluation of the potential carcinogenicity, in 2011 IARC evaluated RF EMF as possibly carcinogenic to humans, based on limited evidence of carcinogenicity in humans and in animals (IARC 2013). Specifically, the latter is relevant here and defined in the Preamble as “The data suggest a carcinogenic effect but are limited for making a definitive evaluation because, e.g. (b) there are unresolved questions regarding the adequacy of the design, conduct or interpretation of the studies...” (IARC 2019). Since this evaluation, many new studies involving cancer in humans, cancer in experimental animals, and mechanistic data have been published and a re-evaluation has also been identified as a high priority by an international Advisory Group to the IARC Monographs program (IARC 2019). In order to assess potential adverse health outcomes of RF EMF, an updated assessment is crucial to support decision makers and to inform the general public.

1.1. Objective

This systematic review relates to carcinogenicity of RF EMF in animals, and it is part of the project launched by the World Health Organization (WHO) to assess potential adverse health effects of exposure to RF EMF in the general and the working population. The topic, 'animal cancer' is one of the six major topics identified in a broad international survey amongst RF experts to prioritize potential adverse health outcomes from exposure to these fields (Verbeek et al. 2021). In this review, we addressed the questions of whether RF EMF exposure leads to an increased incidence of cancer in experimental animals when compared to sham exposure/no exposure or lower exposure and whether there is an exposure-dependent and/or time-dependent relation between the exposure and the outcome (dose-response).

2. Methods

The methods used to conduct this analysis are briefly described below and presented in more detail in the study protocol (Mevissen et al. 2022), which has been updated to include amendments (Annex 1). The amendments were primarily to clarify analyses methods or presentation of information and were not considered by the authors to materially impact overall conclusions when compared to usage of the original study protocol. A tabular summary of the changes is included in Annex 1.

1.1. Eligibility criteria

A Population, Exposure, Comparator, Outcome, Study Type (PECOS) statement was developed to focus the research question(s), identify search terms, and develop inclusion/exclusion criteria used during screening. The PECOS statement was based on addressing the following questions:

1) Among laboratory animals, does exposure to RF EMF increase the risk of cancer compared to sham, no exposure or lower exposure in experimental studies, and to assess if there is an exposure-dependent and/or time-dependent relation between the exposure and the outcome?
2) Is there an exposure/dose response between the exposure and the outcome?

Eligibility criteria are summarized in Table 1, presented according to the PECOS elements. The protocol presents elaborations on eligibility criteria (Mevissen et al. 2022).

2.2. Information sources and search strategy

The literature search focused on exposure condition (microwaves, high frequency fields, radio waves, radio frequency fields, mobile phones, GSM, CDMA, Universal Mobile Telecommunications Service/UMTS), cancer, and animal studies. As described in the protocol (Mevissen et al. 2022), the search strategy was designed via text analysis of key studies identified by the team. Search terms consist of controlled vocabulary (e.g. MeSH) and free text for the following main concepts: electromagnetic fields, cancer, and animal studies. An animal studies search filter by The National Toxicology Program (NTP) is adapted and applied to the search strategies.

An information specialist (TR) used the inputs to develop search strategies using subject headings and free text terms approaches tailored for each of the databases presented below because each database has its own search architecture. Developing the search strings was iterative with refinements made by the team (MM, KAT, KS) during the early stages to confirm that key references were being retrieved. Animal studies were identified by using a NTP RoC search string that was adapted and applied to the current project (https://ntp.niehs.nih.gov/ntp/roc/handbook/rochandbookappendix_508.pdf). No language or publication date restrictions were applied. Full details of the search strategy for each database are presented in Annex 1 in the protocol (Mevissen et al. 2022).

The databases listed below were initially searched by an information specialist (TR) followed by team members (MM, KS, KAT) in October 2021 and last updated in July 2023,

- MEDLINE via PubMed (https://pubmed.ncbi.nlm.nih.gov/)
- Science Citation Index Expanded and Emerging Sources Citation Index via Web of Science (https://clarivate.com/webofscienceup/solutions/web-of-science/)
- EMF Portal (https://www.emf-portal.org/)

In addition to the database searches, reference lists of the eligible publications were checked to identify studies not captured by the database searches. Citation searches of the included studies were performed via the Citation Chaser tool https://estech.shinyapps.io/citationchaser/

2.2. Screening and Selection Processes

The records retrieved from the literature searches were imported to Endnote® 20 for deduplication. The remaining records were imported to DistillerSR (https://www.evidencepartners.com/products/distillersr-systematic-review-software). Title / Abstract (TIAB) and full-text screening was conducted by two independent reviewers at each level (MM, AD, KS for TIAB, and MM and KAT for full-text screening) and any screening conflicts were resolved by discussion between the primary screeners with consultation by a third reviewer (KS and AD), if needed. The records were screened for eligibility in sequence by title and abstract, and then full text review. For citations with no abstract, articles were initially screened based on the following: title relevance (title should indicate clear relevance), and page length (articles two pages in length or less are assumed to be conference reports, editorials, or letters). Eligibility status of non-English studies was assessed using the same approach with online translation tools or engagement with a native speaker used to facilitate screening.

Full-text records were sought for studies screened as meeting the PECOS criteria or 'unclear' based on the TIAB screening. Rationales for excluding studies were documented, e.g., study did not meet one or more of the PECOS criteria, full-text not available. References that were not able to be procured within 45 days of attempt were considered to be unavailable. For multiple publications using the same or overlapping data, all publications were included, with one selected for use as the primary study; the others were considered as secondary publications. For animal studies, the primary publication was typically chosen based on the longest duration of exposure, or with the outcome(s) most informative to the PECOS.

The screening decisions were then imported into EPA's version of Health Assessment Workspace Collaborative (HAWC, https://hawcproject.org/), a free and open-source web-based software application designed to manage and facilitate the process of conducting systematic literature assessments. Screening results are summarized in a study flow diagram (Page et al. 2021) and interactive HAWC literature trees.

2.2. Data Collection Process and Data Items

Data collection (summarizing study methods and results) for studies that met the PECOS criteria after full-text review was performed in HAWC using structured forms. Method information collected included study characteristics (e.g., study design), animal model, strain, sub-strain and sex, description of exposure (type and levels), methods for how cancer outcomes were assessed, and statistical methods. Data on cancer outcomes (e.g., number of animals with tumor, system, tissue or organ name, neoplastic or non-neoplastic, name of tumor based on histopathology / dose / group) were extracted for each endpoint. For dichotomous variables, values necessary to calculate the effect measures were extracted (e.g., number of events and number of animals/replicates in either the experimental or control group); for repeated measurements of

dichotomous outcomes, number of animals and means or medians with standard deviations/errors were extracted for each experimental group.

Data collection was done by two members of the team (AD and MM) and the entries were mutually verified for the outcomes displayed in figures or tables. Discrepancies during data collection were discussed by extractors and another team member (KAT, KS) was involved if no agreement could be reached. Data collection information can be downloaded from HAWC as Excel files at https://hawcproject.org/assessment/1170/ (see 'Downloads').

2.2. Risk of Bias and Sensitivity Assessment

Study evaluation focused on identifying potential sources of bias (factors that could systematically affect the magnitude or direction of an effect) or insensitivity (factors that limit the ability of a study to detect a true effect). Signaling/prompting questions and explanatory instructions can be found in the protocol (Mevissen et al. 2022); see Supplementary Data 3) as well as in the HAWC project for this assessment (https://hawcproject.org/rob/assessment/1170/; see the 'Risk of bias and sensitivity' module).

The risk of bias (RoB) assessment was conducted using a tool developed by the NTP Office of Health Assessment and Translation (OHAT) (National Toxicology Program 2015b; 2019a; National Toxicology Program 2019b). In brief, RoB judgements were reached for the following RoB domains: 1. selection bias, 2. performance bias, 3. detection bias, 4. attrition/exclusion bias, 5. selective reporting, and 6. other sources of bias. Domain judgements used a four-point scale of 'definitely high', 'probably high', 'probably low', or 'definitely low' RoB. Corresponding authors were asked for missing data, particularly when the information could address major study limitations identified during RoB assessment. In addition to RoB, study sensitivity was assessed separately using methods utilized by the NTP Report on Carcinogens (RoC) (National Toxicology Program 2015b). Sensitivity considerations are related to study design (namely animal model), statistical power, exposure duration/contrast, and observation period with respect to adequacy to detect a neoplasia. Sensitivity judgements were characterized as 'no, some, major or critical concerns'. The findings are used to conduct sensitivity analyses, e.g., helpful to assess potential inconsistency across studies.

At least two reviewers [AD, MM, KS, JMN, JW, AKS, AW] independently evaluated each study using structured web-forms housed within HAWC (https://hawcproject.org/assessment/1170/). The questions related to the dosimetry domain were always evaluated by at least one expert in the field (AW, JMN), including such issues as consistency of reported exposure parameters, exposure system design and the assessment of possibile temperature rise in experimental animals. The questions regarding ‘Sensitivity’ were evaluated by two experts (KS, MM). Differences between reviewers were resolved by discussions or by consulting additional reviewers.

2.2.2. Reporting Results of Bias and Sensitivity Assessments

Results are displayed using a 'heatmap' format with judgements for RoB and sensitivity presented separately. When accessed in HAWC, these heatmaps are interactive with 'click to see' capabilities to see the underlying rationales for judgements. Judgements (and underlying rationales) are also available for download in HAWC. We used a tier system for the evaluation of RoB and study sensitivity. Variation in study tiers can be used to probe patterns of findings across studies.

2.2. Effect Measures

Eligible studies have to provide an effect measure or sufficient data to calculate an effect estimate; the type of effect measure will not influence the decision to include or exclude studies.

For qualitative syntheses of results, the exposure leading to an increase in the outcome was assessed if a statistically significant difference between any of the exposed groups and the no-exposure control condition or a statistically significant dose-response relationship across the exposure levels. To assess if there was a statistically significant dose-response in the effects of the exposure per exposure level, we relied on published trend tests or calculated appropriate trend statistics.

Effect sizes were judged with respect to the direction of the outcome (e.g., increased, or decreased incidence in cancer for the respective tumor type) and the lowest dose (SAR level) where an increase in cancer incidence occurred. We also noted the exposure duration at which a statistically significant increase in the tumor incidence was observed. This assessment was primarily based on chronic cancer bioassays.

For the determination of the effect size, several considerations (such as the background incidence for a given neoplasm and the increase of the tumor incidence by exposure level are critical. Our original protocol indicated effects sizes would be presented as relative risk to help make methods consistent with the analysis of human evidence. However, deriving a relative risk for animal cancer bioassays has critical shortcomings most important being that variation in the background tumor rates across different animal species and strains makes it inappropriate to quantitatively extrapolate a relative risk from an animal bioassay to humans. For example, if one bioassay has no response in the control group (0% with tumors) and another has 5% and both have an additional risk of 5% at 1 SAR, then even though they have equal additional risk, the relative risk of 1 SAR versus control in the first study is infinite and the other is 2.0 suggesting a bigger effect in the first bioassay. However, this finding is driven entirely by the control response and not the effect of the RF-EMF. Therefore, we adjusted our effect size presentation to be similar to the metric used in cancer risk assessment (Annex 1). Specifically, effect size was based on the benchmark dose (BMD) of a respective study where dose-response or a statistically significant trend was identified for an outcome.

A BMD is defined as a dose or concentration that produces a predetermined change in the response rate of an adverse effect. The analysis uses Benchmark Dose Tools (BMDS) online (see BMDAnalysisSupplementaryMaterial (Supplement 1). In each case, the analysis used the default settings provided by the US EPA in BMDS but the benchmark response (BMR) was changed to 0.01 (1% BMD is the exposure in SAR that is estimated to result in a 1% extra risk where extra risk is defined as $\frac{R(BMD)-R(C)}{1-R(C)}$ where R(BMD) is the tumor incidence response at the BMD and R(C) is the tumor incidence response in the control group). A BMR of 1% was used because 5% and 10% are above the response ranges for the data sets and this would constitute an extrapolation upward rather than an interpolation in the range of the data.

There are many different models that could be used to analyze the benchmark dose such as a Weibull model, logistic model or multistage model (see BMDAnalysisSupplementaryMaterial) (Supplement 1). Using the individual models can result in highly different BMD values. Most applications of the BMD either choose one of these models to present their results or present some sort of average value across multiple models to avoid dependence upon a single model. In this analysis, the Bayesian Average as calculated in BMDS was used for the primary comparison across studies. For completeness, the 1% BMD for the individual model with the best fit as recommended by the EPA software is also presented in the appendix.

The output of the analysis is the Bayesian Average BMD 1% and the 95% lower and upper confidence bounds (BMDL, BMDU). We chose to present the BMD01 in the text and provide all the details for all of the analyses in the Supplement 1 to allow the interested reader to see what alternatives would be. The BMD is not provided in cases where the modeling failed and/or was questionable. In general, BMD modelling failed when there was no apparent dose-response either

due to a minimal or zero added risk or a pattern of dose-response that did not conform with any model.

2.2. Data Display and Evidence Synthesis

Studies were first grouped according to the broad organ system, i.e., auditory, brain, cardiovascular, endocrine, gastrointestinal, heart, immune, kidney, liver, lymphoma, mammary, musculoskeletal, pituitary, reproductive, respiratory, skin, urinary). For each organ system, exposure response arrays were developed to qualitatively show the direction of effect at a given exposure or dose-level (null, statistically significant increase, statistically significant decrease, statistically significant dose-response relationship or trend). Subsequent analyses were prioritized to focus on specific organ systems where at least one study was reporting a statistically significant harmful effect of RF EMF exposure. For outcome groups where no study showed an increase in cancer related outcomes (such as malignant neoplasms, preneoplastic lesions or a combination of benign and malignant lesions compared to no or sham exposure), we did not further analyse the data but provided a description of the studies. This was the case for gastrointestinal/digestive, kidney, mammary gland, urinary, endocrine, musculoskeletal, reproductive, and auditory system. In HAWC, visuals related to these studies are available (Table 13).

For outcome groups where at least one exposure group in one study or experiment showed an increase in cancer outcomes compared to no exposure or if there was a significant trend in outcome over exposure groups, we assessed that there was an effect of the exposure, regardless of null outcomes in other studies. As stated in the study protocol, relevant endpoints for cancer outcomes include tumor incidence and prevalence, tumor type, different stages of carcinogenesis including preneoplastic lesions, number of benign and malignant tumors per animal, and survival of the animals (IARC 2019). These endpoints depend on the experimental model (chronic bioassay over about 2 years, initiation-(co-)promotion and tumor co-carcinogenesis studies, tumor-prone animal models, implanted tumor cells), species and strain as well as sex and tumor type (IARC 2019).

Among study types, the long-term carcinogenicity studies (chronic bioassay) are the most informative because they are considered to provide the most direct evidence with respect to extrapolating to humans (Annex 1). initiation-(co-)promotion studies, co-carcinogenesis studies, studies in tumor-prone animals, and implantation (cancer cells) studies can provide supportive information. In cancer hazard assessment, these study designs are typically considered less generalizable to humans (e.g., the IARC Monographs program and the U.S. Report on Carcinogens). For example, initiation-(co-)promotion studies are targeted to an organ system, and the outcome depends on the initiator and/or promotor used in a specific animal species and strain, dose, and the length of the experiment. These studies are also shorter in length than a chronic bioassay (several months versus 2 years). Hence, not all initiation- (co-)promotion studies targeting the same organ are expected to have a similar outcome as other initiation-(co-)promotion studies and compared to a chronic cancer bioassay. Therefore, initiation- (co-)promotion studies with no effect compared to sham-exposed controls are not typically interpreted as negating chronic cancer bioassay studies with an increase in cancer. Studies in tumor-sensitive animals can also help screen potential carcinogens for specific tumor sites. Initiation-(co-)promotion- studies or studies with tumor-prone mice provide information about joint effects and can include single agent study arms of long-term duration. These study types can provide supportive evidence of an effect of the exposure of interest on a predetermined cancer type in specific organs or systems using either chemicals, or certain forms of radiation as initiators, promotors or co-carcinogens or using knock-in or knock-out mice. For reasons mentioned above, the tumor concordance analysis between experimental animals and humans also used this approach of relying on chronic cancer bioassays (Baan et al. 2019; Krewski et al. 2019). Therefore, we implemented the following amendment to the protocol: for cancer hazard identification, evidence from chronic cancer bioassays (around 2 years) in animals is necessary for concluding on 'high certainty' CoE because it is the most direct evidence for identifying

potential human carcinogens (IARC 2019; National Toxicology Program 2015a). If evidence was only based on results from other study types, e.g., initiation-promotion studies, initiation-(co-)promotion studies, at maximum moderate CoE could be reached.

2.2.2. Data presentation

We then analysed for which tumor types the risk was increased, and we assessed the certainty of the evidence. This was the case for the following outcomes: lymphoma, brain, heart, liver, and adrenal gland. For lymphoma, we grouped all tumors as lymphoma. For brain, we grouped outcomes according to site of origin as derived from glial or other cells - according to the 2021 WHO classification of central nervous system (CNS) tumors (Louis et al. 2021).

For lymphoma, brain, heart, adrenal gland and liver cancer outcomes, figures presenting exposure arrays, presenting the outcome by dose (SAR level), and the tumor incidence per treatment group or controls, were developed in HAWC to evaluate the CoE considerations (e.g., RoB across studies, consistency, dose-response, magnitude of effect). In these figures studies were categorized according to the specific outcome type or site of origin of the carcinogenic outcomes (e.g., lymphoma, glial cell derived brain neoplasms, malignant neoplasms other than glial cell derived). Outcomes were further organized as neoplastic lesions, namely benign or malignant tumors, a combination of both when appropriate, and pre-neoplastic lesions. Furthermore, the respective outcomes were organized by species and sex, and both were ordered by study type, namely chronic bioassay, initiation-(co-)promotion studies, and studies in tumor-prone animals. Separate conclusions were assessed for different phone bands [Global System for Mobile communication (GSM) versus Code-Division Multiple Access (CDMA)] or different regions of the electromagnetic field (i.e., near-field versus far-field) or the pattern of exposure, namely continuous versus pulsed exposure. CoE conclusions are expressed at the organ system level with elaboration on which specific outcomes were most informing the conclusion.

2.2. Certainty Assessment

We used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) certainty of the evidence (CoE) framework (Morgan et al. 2016) and systematic review methodology developed by the National Toxicology Program (NTP) for use in developing non-cancer and cancer assessments of environmental agents (National Toxicology Program 2015a; b; 2019a). The term certainty of the evidence is equivalent with the term confidence of the evidence used by OHAT. In addition to the systematic review resources listed above, methods used by the NTP Report on Carcinogens (RoC), IARC (IARC 2019), and the Organization for Economic Co-operation and Development (OECD) (OECD 2009; 2012; 2014; 2018) on the analysis of animal cancer bioassays were utilized to inform expert judgements that underpinned evidence synthesis and CoE conclusions.

The GRADE CoE framework as described in the OHAT Handbook (2019) (National Toxicology Program 2019a) was used, with adaption to the subject matter on RF EMF exposure and cancer in experimental animals as described in the study protocol in Annex 4 (Mevissen et al. 2022).

As shown in Figure 1, the initial rating of the CoE experimental studies in laboratory animal studies started at high CoE. Subsequently, downgrading or upgrading factors were considered and may change the initial CoE. Five factors for a body of evidence were used to determine if the initial certainty rating should be downgraded: (1) RoB, (2) indirectness, (3) unexplained inconsistency, (4) imprecision, and (5) publication bias. Five properties for a body of evidence

were used to determine if the initial certainty rating should be upgraded: (1) large magnitude of effect, (2) dose response, (3) residual confounding (e.g., studies report an effect and residual confounding is toward the null), (4) consistency across study designs and experimental model systems, and (5) other, such as specificity of the association in cases where the effect is rare. For cancer hazard identification, evidence from chronic cancer bioassays (around 2 years) in animals is prioritized over other study types, e.g., initiation-(co-)promotion studies, particularly for concluding on 'High CoE'. The OHAT certainty assessment findings were presented in tabular format in a narrative way, including conclusions regarding direction and magnitude of effect(s).

The OHAT method is based on GRADE, which has been developed for evidence synthesis of human evidence, and the application to animal data is still in the early stages (National Toxicology Program 2019b). Experience to date suggest that method refinement and clarification is needed (Hooijmans et al. 2018). In scenarios where the application of GRADE/OHAT was unclear we deferred to methods of evidence synthesis used in environmental health developed by organizations where the approaches are widely accepted and have undergone peer review (EPA 2022; National Toxicology Program 2015a). More specifically, this occurred when trying to characterize scenarios where we evaluated the experiments as well-conducted (Tier 1) but too heterogeneous with respect to study design to be combined in a meta-analysis. In our narrative synthesis, we acknowledge the inconsistent findings did not feel it appropriate to downgrade because the inconsistency was very plausibility explainable by heterogeneity in study design (e.g., species, age at exposure, nature and/or exposure characteristics, e.g., CDMA versus GSM). When all studies are Tier 1 with respect to RoB, any inconsistent findings are quite plausibly due to differences in study design, and there are no concerns for study sensitivity in the null studies, it is challenging to determine which conclusion to advance when framing CoE, i.e., evidence for carcinogenicity or null. In these cases, we deferred to methods of evidence synthesis widely used in environmental health where CoE is framed in the context of the studies reporting positive evidence of carcinogenicity (EPA 2005; National Toxicology Program 2019b).

3. Results

3.1. Study selection

The database searches were first performed on October 14th, 2021, with no language or publication date restrictions, and were last updated in January 2023. The flow of studies during the screening process is summarized in Figure 1, using a recommended format (Page et al. 2021).

A total of 4,350 records were retrieved from database resources prior to deduplication and 1,255 duplicates were detected. A total of 3,095 records were screened at the TIAB level with 3,002 records excluded and 93 records sought for full text retrieval. During full text screening, a total of 41 were excluded: 39 for not meeting one or more PECOS criteria, one for not having original data, and one reported the same data as presented in an included study. After full text screening, 52 studies were determined eligible. An interactive visualization summarizing the study flow and providing access to citation content is presented in Figure 2 and available for exploration in HAWC (Visualization | Literature review | EMF (Animal Toxicology) (2022) | HAWC (hawcproject.org).

2.2. Study Characteristics of Included Studies

Of the 52 animal cancer studies, there were 10 two-year chronic cancer bioassays (Anderson et al. 2004; Chou et al. 1992; de Seze et al. 2020; Falcioni et al. 2018; Jin et al. 2011; La Regina et al. 2003; National Toxicology Program 2018a; b; Smith et al. 2007; Tillmann et al. 2007), 24 initiation-(co-)promotion studies (Adey et al. 2000; Adey et al. 1999; Anane et al. 2003;

Bartsch et al. 2002; Chagnaud et al. 1999; Heikkinen et al. 2006; Heikkinen et al. 2003; Heikkinen et al. 2001; Hruby et al. 2008; Huang et al. 2005; Imaida et al. 2001; Imaida et al. 1998a; Imaida et al. 1998b; Lerchl et al. 2015; Mason et al. 2001; Shirai et al. 2007; Shirai et al. 2005; Szmigielski et al. 1982; Szudzinski et al. 1982; Tillmann et al. 2010; Wu et al. 1994; Yu et al. 2006; Zook and Simmens 2001; 2006) and 18 experiments in tumor-prone animals (Bellossi et al. 2000; Frei et al. 1998a; Frei et al. 1998b; Higashikubo et al. 1999; Jauchem et al. 2001; Lee et al. 2011; Oberto et al. 2007; Ouadah et al. 2018; Paulraj and Behari 2011; Preskorn et al. 1978; Repacholi et al. 1997; Salford et al. 1997; Salford et al. 1993; Saran et al. 2007; Sommer et al. 2007; Sommer et al. 2004; Toler et al. 1997; Utteridge et al. 2002). Of these, 26 studies were performed on rats, and of these 8 were in chronic bioassays, 14 initiation-(co)-promotion and 4 in tumor prone animals. In 23 studies animals were adults when exposed and in 3 studies juvenile animals were exposed.

Exposure time varied from 6 hours and 5 days in (Saran et al. 2007) to 7 days/week, 9 hours and 10 minutes/day (10 minutes ON and 10 minutes OFF) for 108 weeks (National Toxicology Program 2018a) in study. On average animals of chronic cancer bioassays were exposed for 104 weeks. Exposure frequency was 1900 MHz in 1 (GSM), 3 (CDMA), 3 (UMTS) studies, 1747 MHz in 2 studies, 1439 GHz TDMA in 2 studies, 898 to 929 MHz in 11 (GSM), 5 (RF), 1 (MW), 1 (pulsed), 1 time-divisions multiple access (TDMA); 835 to 847.74 MHz in 2 phase-modulated continuous wave (PMCW), 2 (pulsed), 5 (CDMA), 1 (TDMA); 130 GHz in 1 study; 2.45 GHz in 7 studies; 1.6 to 1.8 GHz in 4 studies; 1.95 GHz in 1 (CDMA) study; 94 MHz in 1 study; 60 GHz in 1 study; 1 kHz in 2 studies.

The most commonly assessed sites across studies were the brain, liver, adrenal gland, skin, thyroid, pituitary, mammary, uterus, ovary, lung, pancreas, heart, and kidney. Cancer outcomes of the brain were reported in 20 studies, of the liver in 16 studies, of the adrenal gland in 12 studies, of the heart in 4 studies, of the skin in 10 studies, of the auditory system in 1 study, of the endocrine system in 12 studies, of the gastrointestinal system in 4 studies, of the immune system in 13 studies, of the kidney in 10 studies, of the mammary system in 11 studies, of the musculoskeletal system in 5 studies, of the pituitary gland in 8 studies, of the reproductive system in 17 studies, of the respiratory system in 13 studies, and of the urinary system in 10 studies.

Eligible studies were dated from 1978 to 2020.

2.2. Risk of Bias and Sensitivity Analysis Across All Studies

2.2.2. Risk of Bias in studies

RoB judgements for the 52 included studies are presented in Figures 3A (for chronic bioassay studies), 3B (for initiation-(co)-promotion studies), 3C (for studies with tumor-prone mice). 45 studies were considered *definitely* or *probably* low RoB in most domains. Notable exceptions were seven studies (Bartsch et al. 2002; Bellossi et al. 2000; Chagnaud et al. 1999; de Seze et al. 2020; Salford et al. 1993; Zook and Simmens 2001) which had significant limitations in at least half of the domains with all of these besides one (de Seze et al. 2020) rated as 'probably high' RoB. Among the other studies, when limitations were identified, they were most commonly present in domains for randomization, blinding (during the study), and results presentation. Considerations on possible temperature rise in relation to RF EMF exposure metrics are given in Annex 2.

With respect to study sensitivity (Figure 4), six studies with major or critical sensitivity concerns overlapped with studies noted above that had significant RoB limitations (Bellossi et al. 2000; Chagnaud et al. 1999; de Seze et al. 2020) or were considered probably high RoB in at least three domains (Paulraj and Behari 2011; Salford et al. 1997; Szmigielski et al. 1982). Across the entire evidence base, the sensitivity domains, where concerns were most often noted, included endpoint specificity, and sensitivity and exposure contrast. According to the sensitivity

considerations outlined in the protocol, initiation-(co-)promotion studies on skin or breast cancer model using an initiator and/or promotor, e.g., 7,12-dimethylbenz(a)anthracene (DMBA), have a reduced sensitivity because cancers other than skin or mammary cancer, respectively, cannot be found (Anane et al. 2003; Bartsch et al. 2002; Imaida et al. 2001).

2.2. Prioritizing Outcomes for Analysis

No or minimal evidence of RF EMF exposure-related cancer outcomes was found in most systems organs when compared to the respective controls or dose-response effects in any study regardless of the study type (these included gastrointestinal/digestive, kidney, mammary gland, urinary, endocrine, musculoskeletal, reproductive, and auditory). A brief description of the evidence for each of these systems is summarized in supplemental Table S1 long with access to supplemental data figures and RoB/sensitivity judgements. These systems and organs were not advanced to full certainty assessment.

The most detailed analyses were prioritized for outcomes where at least one study reported exposure-related effects: lymphoma, brain, heart (malignant schwannomas and related lesions), adrenal gland, and liver.

2.2. Lymphoma

2.2.2. Characteristics of included studies for lymphoma-related carcinogenicity

Seventeen publications reported assessment of lymphomas (Anderson et al. 2004; Chou et al. 1992; Frei et al. 1998a; Heikkinen et al. 2006; Heikkinen et al. 2001; La Regina et al. 2003; Lee et al. 2011; Lerchl et al. 2015; National Toxicology Program 2018a; b; Oberto et al. 2007; Repacholi et al. 1997; Sommer et al. 2007; Sommer et al. 2004; Tillmann et al. 2007; Tillmann et al. 2010; Utteridge et al. 2002) (Figure 5). Among these publications, there were 12 studies in mice (~4,500 animals) and five in rats (~2,000 animals). Six publications included chronic bioassays (Anderson et al. 2004; Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018a; b; Tillmann et al. 2007), 4 included initiation-(co)-promotion studies (Heikkinen et al. 2006; Heikkinen et al. 2001; Lerchl et al. 2015; Tillmann et al. 2010), and 7 studies were done in tumor-prone mice (Frei et al. 1998a; Lee et al. 2011; Oberto et al. 2007; Repacholi et al. 1997; Sommer et al. 2007; Sommer et al. 2004; Utteridge et al. 2002). Sommer et al. did not report findings in a numerical way and could not be reported in the table (Sommer et al. 2007). Figure 5 is a thumbnail image of an interactive graphic in HAWC. The graphic presents the included studies and endpoints assessed across studies (Panel A) and summarizes study characteristics for the included studies (Panel B). Additional details on the studies can be accessed through the interactive dashboard as well as download of the underlying data in tabular format using the ‘Actions’ option.

This is a thumbnail image of an interactive dashboard available on HAWC that is filterable by citation, health system, species, strain, and frequency. **Panel A** Numbers in the heatmap represent the number of studies that investigated a health system within a study design. If a study evaluated multiple health outcomes or presented several experiments, it is shown here multiple times. **Panel B** provides additional experimental detail about the studies, including exposure characteristics. The table in the Panel B is only a partial representation of the full table available in HAWC.

2.2.2. Risk of bias and sensitivity considerations for lymphoma-related carcinogenicity

There were no significant concerns for RoB in studies assessing lymphomas from chronic bioassays (Figure 6) or initiation-(co)-promotion studies (Supplementary Figure S1A). The studies by Repacholi et al., and Sommer et al., had major concerns in domains for temperature monitoring, complete outcome data (Supplementary Figure S1B) (Repacholi et al. 1997), blinding during the study and presentation of the results (Sommer et al. 2004). Considerations on possible temperature rise in relation to RF exposure metrics are given in Annex 2. Three studies had major sensitivity concerns (Supplementary Figure S1C) related to the exposure contrast (Chou et al. 1992; Heikkinen et al. 2006) or endpoint sensitivity and specificity (Oberto et al. 2007). Among the chronic cancer bioassays that reported on lymphoma, the NTP studies, and the Tillmann study published in 2007, had only minor concerns in all domains and are therefore considered the most informative studies.

A high RoB was noted for four studies based on a lack of blinding among initiation-(co)-promotion studies (Supplementary Figure S1A) (Heikkinen et al. 2006; Heikkinen et al. 2001) and among studies with tumor-prone mice Supplementary Figure S1B (Frei et al. 1998a; Sommer et al. 2007), and the inadequate presentation of the results (Sommer et al. 2007).

2.2.2. Summary of the evidence for lymphoma-related carcinogenicity

****Figure 7 and Supplementary Figures S1D, S1E present exposure response arrays of lymphoma findings across the 17 studies where data were extracted to summarize exposure-specific level effects sorted by species and sex separated by dashed lines. Figure 7A displays the findings for chronic bioassays, Supplementary Figure S1D for initiation-(co-)promotion studies and Supplementary Figure S1E shows the findings across used SAR values for studies in tumor-prone mice. Figure 8 and Supplementary Figures S1F, S1G present the effect size of lymphoma findings. They present the exposure arrays to summarize exposure-related effects, and effect size, namely tumor incidence data to allow for analyses of factors such as exposure gradient, magnitude of effect, and a deeper consideration of consistency (or inconsistency) compared to what can be accomplished by focusing on statistical significance. In this case, the Bayesian Average 1% BMD was not calculated because there was no dose-response or trend for an increase or decrease in lymphomas.

Overall, there was considerable heterogeneity in the cancer models used to assess lymphoma which made it challenging to understand any apparent inconsistency in findings. Two studies provided data indicating an increase in the incidence of lymphoma with one being a well-conducted NTP chronic bioassay in female mice (National Toxicology Program

2018a), and the other an experiment using a transgenic mouse model with significant concern for how lymphoma incidence was assessed (Repacholi et al. 1997). No effect on lymphoma or a dose-response relationship was reported for male mice or rats regardless of the study type used.

2.2.2.2. Effects of RF EMF in chronic bioassays

2.2.2.2.2. Mice

All neoplasms found are considered neoplastic, and lymphoma can be considered a malignant disease. Benign lymphomas are referred to as either pseudo-lymphomas or benign lymphoid hyperplasias. None of benign lymphomas were reported, therefore, the following data describe malignant lymphomas.

One study provided evidence of effects on lymphomas in female mice but not in male mice, namely an increase in the incidence (National Toxicology Program 2018a). In the chronic bioassay performed by the NTP study (GSM-modulated RF EMF at 2.5, 5 or 10 W/kg 9 hours/day, seven days/week for 106 (males) or 108 (females) weeks), statistically significant increases in malignant lymphomas were found in the lowest and medium exposure groups in female B6C3F1/N mice. However, the trend test performed by the authors was not statistically significant, and the incidence of malignant lymphoma in the sham controls was significantly lower compared to the respective historic controls (National Toxicology Program 2018a). In the CDMA-modulated RF EMF experiment, significant increases in malignant lymphomas were observed in the lowest exposure group for female mice, but again, the number of lymphomas in the sham controls was lower than those in the respective historic controls (National Toxicology Program 2018a). No effect of RF EMF compared to the sham controls was reported for GSM-exposed female mice (National Toxicology Program 2018a).

No evidence for a statistically significant increase in lymphomas was found in a chronic bioassay (2 years) in male or female B6C3F1 mice exposed to 902 MHz (GSM) or 1747 MHz (DCS) for 2 hours per day, 5 days per week (Tillmann et al. 2007). In mice exposed to 1747 MHz (DCS), a statistically non-significant increase in lymphomas is reported, particularly at the lowest exposure group, and less in the two higher exposure groups.

2.2.2.2.2. Rats

In the NTP study performed in rats with the exposure starting on gestational day (GD) 5, no effect on lymphomas was found at any modulation or SAR level investigated in rats of both sexes (National Toxicology Program 2018b). Other chronic bioassays using unmodified rodent strains did not detect effects on lymphomas. Briefly, in male Sprague Dawley rats exposed at 2,450 MHz pulsed (8 Hz) for 25 months, no effect on lymphomas was found (Chou et al. 1992), but this study had critical sensitivity concerns. Anderson et al. reported no effects on lymphomas in female or male Fischer 344 rats after long-term exposure of 1.6 GHz at SAR levels (brain) of 1.6 and 0.16 W/kg (Anderson et al. 2004). Another study in Fischer 344 rats of either sex also reported no increase in malignant lymphomas at an exposure of 836 and 848 MHz at 1.3 W/kg (La Regina et al. 2003).

2.2.2.2. Effects of RF EMF in initiation-(co-)promotion studies

Four of the initiation-(co)-promotion studies studied effects of RF EMF on lymphomas (Heikkinen et al. 2006; Heikkinen et al. 2001; Lerchl et al. 2015; Tillmann et al. 2010).

2.2.2.2.2. Mice

No effects were observed in Heikkinen et al., 2001, who used an initiation-(co)-promotion model in female mice with additional exposure to X-ray (initiation, 4-6 MV, three weekly exposures of 1.333 Gy) followed by exposure to frequency-modulated radiofrequency at 902.5 MHz (Heikkinen et al. 2001). In another initiation-(co)-promotion study, no effect of UMTS exposure for 24 months was found in B6C3F1mice of both sexes (Tillmann et al. 2010). The study by Lerchl et al., designed as a replication study of Tillmann et al., 2010, indicated a significant increase in lymphomas in mice of both sexes at 0.4 W/kg (24% incidence in RF EMF group versus 9% in sham controls) with the effect not being dose-dependent (Lerchl et al. 2015).

2.2.2.2.2. Rats

Only one study was performed in rats. In this study by Heikkinen (Heikkinen et al. 2006), a co-carcinogenesis model was used, and no effect on lymphomas or related preneoplastic lesions was found in female rats exposed for 104 weeks to 900 MHz, 0.3 W/kg (Heikkinen et al. 2006).

2.2.2.2. Effects of RF EMF exposure on lymphoma in tumor-prone animals

2.2.2.2.2. Mice

Three studies were performed using the Eμ-pim1 transgenic mouse model system is susceptible to develop lymphomas. Across these three studies, mice were exposed to 900 MHz, GSM or CDMA, at different SAR levels for up to 18 months, 30 minutes twice a day in a restrained Ferris wheel exposure setup, except for the Repacholi study.

Only one of the studies in mouse models reported a statistically significant effect on lymphomas (Repacholi et al. 1997). With respect to the positive finding in Eμ-pim1 transgenic mice (Repacholi et al. 1997), there were important study design differences with other studies using this strain (Oberto et al. 2007; Utteridge et al. 2002), most notably in the data acquisition and analysis methods (Detection and Attrition domains) used by Repacholi et al. that reduces confidence in this study. In the Repacholi study, the authors presented a significant increase in the lymphoma incidence of all lymphomas and lymphoblastic lymphomas in female mice exposed to 1.4 W/kg (SAR ranging from 0.13-1.4 W/kg), 900 MHz GSM with a pulse repetition frequency of 2017 Hz and a pulse width of 0.6 milliseconds (Repacholi et al. 1997). However, there was also a difference in death before study termination (18 months) of 70/101 and 44/100, in treated and control animals, respectively. The numbers used in the statistical analysis stem from data of animals that died during the course of the study, whereas histopathological examinations at the end of the study of animals that survived, was not performed. This limits the interpretation of the data and is reflected in the RoB in the presentation of the data in the completeness of the outcome data domain, but the direction of bias (away or toward the null) is not obvious. These concerns do not apply to the assessment of lymphoma mortality, and if lymphomas in mice quickly results in moribund mice any bias would be small in magnitude for incident lymphomas.

In the study by Utteridge et al., which sought to replicate the Repacholi study, 4 SAR levels (0.25, 1, 2, 4 W/kg) were used with about the same number of animals per group and a similar exposure (Utteridge et al. 2002). No exposure-related differences in body weight or survival were noted. The outcome showed no increase in lymphomas. Despite the longer study duration of 108 weeks, a direct comparison between the studies by Repacholi and Utteridge is difficult to make since most mice in the latter study developed lymphomas only by study termination. A third study by Oberto et al. (Oberto et al. 2007) using the same exposure system as Utteridge et al., and an exposure for 18 months total to 900 MHz at 3

whole body SAR levels, observed statistically non-significant increases in lymphomas in female mice. In the Utteridge study, the background tumor incidence was higher in the Pim1 transgenic mice (74.2%), and therefore, it makes it less likely to detect any effect of RF EMF. In addition, the positive control (ethylnitrosurea (ENU) resulted in a lower tumor incidence when compared to the control animal, which indicates differences in the Pim1 strain used (Goldstein et al. 2003). In contrast to the Repacholi study, the sham controls in Oberto et al. (Oberto et al. 2007) had a lower lymphoma incidence (12% versus 22%) at study termination. In addition to the differences in data acquisition and analysis, we considered whether the unexplained inconsistency across these three studies using the same model system might stem from differences in the magnitude of the SAR levels. Repacholi et al. exposed the mice at an average SAR of approximately 1.4 W/kg (one SAR level/dose), whereas Utteridge et al., used 4 SAR/dose levels of a different magnitude, namely 0.25, 1, 2, and 4 W/kg, and Oberto et al., used 3 SAR/dose levels of 0.5, 1.5, and 4 W/kg (Oberto et al. 2007; Repacholi et al. 1997; Utteridge et al. 2002). However, the exposure level used in Repacholi et al. (1.4 W/kg) falls within the ranges tested by the other studies. Even though the SAR/dose levels were different, they were considered similar enough to assume that a difference in lymphoma development could have been found. In summary, the studies performed in tumor-prone mice on lymphomas give mixed results. The initial increase in the incidence of lymphomas in Eμ-pim1 transgenic mice by (Repacholi et al. 1997) was not seen in the two subsequent studies (Oberto et al. 2007; Utteridge et al. 2002) which did not have the data acquisition concerns, but the differences in the strain, leading to higher lymphoma incidences and other methodological differences make it difficult to compare the studies.

Other studies of rodent models used AKR/J or C3H/HeJ mice. No increase in lymphomas was reported by Lee and colleagues (Lee et al. 2011) in tumor-prone AKR/J mice with whole body exposures when compared to sham control animals. Two additional studies (Sommer et al. 2007; Sommer et al. 2004) using AKR/j mice that rapidly develop hematopoietic tumors, including lymphoblastic lymphomas, did not find significant differences in survival before study termination or the incidence of lymphomas (Sommer et al. 2007; Sommer et al. 2004). Data from Sommer et al. are not presented in Figures 7 and 8 because they are only presented as visuals in the publication (Sommer et al. 2007). AKR/j mice were also used by Lee et al. (Lee et al. 2011), with an RF EMF exposure for 42 weeks using CDMA and wideband CDMA (WCDMA). In this study CDMA or WCDMA exposure did not increase the incidence in lymphomas when compared to the respective controls (Lee et al. 2011). The study by Frei et al. did not indicate effects on lymphomas in C3H/HeJ mice exposed to 2450 MHz, at a SAR level of 0.3 W/kg when compared to sham-controls (Frei et al. 1998a).

2.2. Certainty Assessment for lymphoma-related carcinogenicity

The initial rating for the certainty for an increased incidence of lymphoma in all systems was set to high because all studies involved controlled exposure, exposures occurred prior to outcome, outcomes were measured in individual animals, and a concurrent sham-exposed group was used (Mevissen et al. 2022).

Among the 5 chronic cancer bioassays, key features of study design varied considerably (i.e., species, strain, sex, and exposure characteristics) (Table 2). A meta-analysis was not conducted due to variation in study design across the experiments which included 5 chronic bioassays, 4 initiation-(co-)promotion studies, and 7 studies in tumor-prone animals. The

Tillmann experiment with exposure to 902 MHz (GSM) had some concerns because of the unusually high (>70%) lymphoma incidence, which may have masked a possible increase in tumor incidence after Rf EMF exposure. However, the obtained differences could be explained by different exposure types, namely restrained versus non-restrained (ferris wheel versus freely moving animals, different frequencies used, and fragmentation of the dose (1 hour versus 2-times 30 min/day). The Tillmann experiment with exposure to 1747 MHz (DCS) reported higher incidences in exposed mice, particularly in the lowest exposure group (National Toxicology Program 2018a; Tillmann et al. 2007). The one positive study in tumor-prone mice reporting an increase in lymphoma had several RoB concerns (Repacholi et al. 1997). Further, the small number of studies precluded meaningful subgroup analyses. A narrative description of the evidence was considered based on the findings from the 6 chronic bioassays because this is the most direct study design for assessing carcinogenicity. No factors that potentially increase certainty were utilized. Thus, the overall certainty was rated moderate. No BMD was calculated for any of the studies due to lack of clear monotonic dose response.

2.2.2. Consideration of Factors potentially decreasing certainty

- Unexplained inconsistency: Two studies in mice provided some evidence for an effect (National Toxicology Program 2018a; Repacholi et al. 1997), while the others did not (Chou et al. 1992; Heikkinen et al. 2006; Oberto et al. 2007; Tillmann et al. 2007). The inconsistency of results between the chronic cancer bioassays (NTP 2018a; Tillmann et al 2007) might be explained by different exposure/sham-exposure with or without restraint, different frequencies, and fragmentation of the dose or no fragmentation.
- However, there is unexplained inconsistency within the positive study (NTP 2018a) given that there are only pairwise statistically significant increases, particularly at lower doses. While a positive dose–response would be an upgrading factor, we conclude that the pattern of increased tumors only at lower doses supports our conclusion of "unexplained inconsistencies". Differences in study methods between the Repacholi study (Repacholi et al. 1997) and the other studies using Eμ-pim1

transgenic mice (Oberto et al. 2007; Utteridge et al. 2002) that could potentially explain the different findings were discussed above.

- Indirectness: Downgrading for indirectness was not applied. The assessment of lymphoma in these studies was considered direct, with respect to the animal PECOS. Strong evidence on cancer in experimental animals is relevant to the identification of a carcinogenic hazard to humans (Baan et al. 2019). The NTP study with mice utilized a traditional rodent model, widely used in chronic cancer bioassays. The studies in tumor-prone model systems are considered appropriate for assessment of lymphoma in rodents (Kohnken et al. 2017). Regarding the PECOS for human relevance, there is no a priori compelling mechanistic or other biological evidence to suggest animal models are not reasonable for predicting potential effects in humans. Further, in additional analyses of the tumor site concordance between experimental animals and humans, Krewski et al. compared agent-specific concordance between (groups of) cancer sites with sufficient evidence in humans and in animals (the latter required replication of positive results at the same specific site in at least two animal experiments). The overlap for lymphoid and haematopoietic tissues was 58%. It has to be noted that this estimate is expected to underestimate concordance (Krewski et al. 2019).
- Publication bias: No downgrade was applied for publication bias.

2.2.2. Consideration of Factors potentially increasing certainty

- None identified.

2.2. Brain

2.2.2. Characteristics of included studies for brain-related carcinogenicity

Twenty publications reported an assessment of brain neoplasms after RF EMF exposure (Adey et al. 2000; Adey et al. 1999; Anderson et al. 2004; Falcioni et al. 2018; Heikkinen et al. 2006; Heikkinen et al. 2001; La Regina et al. 2003; Lerchl et al. 2015; National Toxicology Program 2018a; b; Ouadah et al. 2018; Salford et al. 1997; Salford et al. 1993; Saran et al. 2007; Shirai et al. 2007; Shirai et al. 2005; Sommer et al. 2004; Tillmann et al. 2010; Zook and Simmens 2001; 2006) (Figure 9). Six studies were conducted in mice (1,902 animals), 13 studies in rats (6,011 animals). Five studies were chronic bioassays (Anderson et al. 2004; Falcioni et al. 2018; La Regina et al. 2003; National Toxicology Program 2018a; b), 10 used models of cancer promotion/co-promotion (Adey et al. 2000; Adey et al. 1999; Heikkinen et al. 2006; Heikkinen et al. 2001; Lerchl et al. 2015; Shirai et al. 2007; Shirai et al. 2005; Tillmann et al. 2010; Zook and Simmens 2001; 2006), and 5 studies used tumor-prone animals (Ouadah et al. 2018; Salford et al. 1997; Salford et al. 1993; Saran et al. 2007; Sommer et al. 2004).

****Figure 9 is a thumbnail image of an interactive graphic in HAWC. The graphic presents the included studies and endpoints assessed across studies (Panel A) and summarizes study characteristics for the included studies (Panel B). Additional details on the

studies can be accessed through the interactive dashboard as well as download of the underlying data in tabular format using the 'Actions' option.

2.2.2. Risk of bias and sensitivity considerations for brain-related carcinogenicity

****Figure 10. Five studies (Salford et al. 1997; Salford et al. 1993; Shirai et al. 2007; Zook and Simmens 2001) had concerns for RoB in at least 2 domains (Supplementary Figures S2A, S2B). The RoB concerns was not considered significant enough to warrant a downgrade across the body of evidence. Four studies had major sensitivity concerns (Supplementary Figure S2C), especially regarding exposure contrast (Falcioni et al. 2018; Heikkinen et al. 2006; Ouadah et al. 2018; Salford et al. 1993). More specifically, the concerns were due to the large uncertainty in the SAR in both the high and the low RF EMF exposure groups, which overlap.

2.2.2. Summary of the evidence for brain-related carcinogenicity

Studies are presented in Figures 11 and 12 and Supplementary Figures S2, starting with malignant neoplasms derived from glial cells (Figure 11 for chronic bioassays; Supplementary Figures S2D and S2E for initiation-(co-)promotion), followed by other malignant neoplasms of the brain (Figure 12 for chronic bioassays; Supplementary Figure S2F for initiation-(co-)promotion), and benign neoplasms (Supplementary Figures S2G and S2H). Each figure is organized into a part A with exposure response arrays to summarize exposure-specific effects for the respective neoplasms, and evidence dose-response data in part B. The Bayesian Average 1% BMD was calculated when dose-response or a trend for an increase or decrease in the outcome was estimated from the data.

Journal Pre-proofs

2.2.2.2. Effects of RF EMF on glial cell-derived tumors in chronic bioassays

2.2.2.2.2. Rats

Glial cell-derived neoplasms in female rats were not increased with regard to RF EMF exposure when compared to sham controls in the chronic bioassays, but statistically significant positive findings (including statistically significant trends) were reported in two chronic bioassays both in male rats with minimal concern for RoB (Figure 10) and sensitivity (Figure S2C) (Anderson et al. 2004; National Toxicology Program 2018b). In male Sprague Dawley rats exposed to GSM-modulated RF EMF (National Toxicology Program 2018b), malignant glioma and glia cell hyperplasia were reported in all exposed groups, whereas no lesions were found in their respective sham controls. This effect was not dose-dependent. In male rats, for CDMA, a statistically significant increase in glioma was calculated at an exposure of 6 W/kg (National Toxicology Program 2018b). Increases in preneoplasic lesions did not reach statistical significance, namely glial cell hyperplasia occurred in 1.5 and 6 W/kg (CDMA) males and 3 and 6 W/kg (CDMA) females. A significant trend ($p<0.01$) for a dose-dependent increase in oligodendrogliomas was calculated by our team in male Fischer 344 rats in the study by Anderson et al. at SAR levels of 0.16 (1.1%) and 1.6 W/kg (2.2%) using the Cochran Armitage (exact) test with historical controls (Anderson et al. 2004) While not significant, the study by Falcioni had major concerns for sensitivity, but showed an increase in malignant tumors in glial cells in female Sprague Dawley rats at much lower SAR levels (0.003, 0.03, 0.1 W/kg) (Falcioni et al. 2018). One other chronic bioassay performed in rats with minimal concern for RoB and sensitivity did not report positive findings in male rats (La Regina et al. 2003).

2.2.2.2.2. Mice

One chronic bioassay was performed in female and male mice, and no effects on glial cell-derived tumors was found after RF-EMF exposure regardless of the sex or the exposure pattern used (National Toxicology Program 2018a).

2.2.2.2. Effects of RF EMF on malignant tumors other than glial cell-derived tumors in chronic bioassays

2.2.2.2.2. Rats and mice

Regardless of the study type, no significant increases in malignant neoplasms other than glial cell-derived neoplasms were found in mice and rats of both sexes compared to the sham controls (Figure 12) (Adey et al. 2000; Adey et al. 1999; Anderson et al. 2004; Chou et al. 1992; Falcioni et al. 2018; Heikkinen et al. 2006; Heikkinen et al. 2001; National Toxicology Program 2018b; Saran et al. 2007; Shirai et al. 2007; Shirai et al. 2005; Zook and Simmens 2001; 2006).

In female rats, male rats, rats of both sexes combined, and mice of both sexes combined, no significant increases in benign neoplasms were found compared to sham-exposed controls regardless of the study type (Falcioni et al. 2018; Heikkinen et al. 2006; Tillmann et al. 2010).

2.2.2.2. Effects of RF EMF on brain tumors in initiation tumor-(co-)promotion studies

2.2.2.2.2. Rats and mice

None of the 10 initiation-(co)-promotion studies, regardless of the species and sex resulted in RF EMF-related statistically significant differences in brain neoplasms (Figures S2D-S2E) (Adey et al. 2000; Adey et al. 1999; Heikkinen et al. 2006; Heikkinen et al. 2001;

Lerchl et al. 2015; Shirai et al. 2007; Shirai et al. 2005; Tillmann et al. 2010; Zook and Simmens 2001; 2006).

2.2.2.2. Effects of RF EMF on benign tumors in any study type

2.2.2.2.2. Rats and mice

No effects of RF EMF on benign tumors compared to the respective controls was found in any of the studies evaluated. In female rats, male rats, rats of both sexes combined, and mice of both sexes combined, no significant increases in benign neoplasms were found compared to sham-exposed controls regardless of the study type (Falcioni et al. 2018; Heikkinen et al. 2006; Tillmann et al. 2010).

In summary, two chronic bioassays in male rats an increase in glia cell-derived tumors with one being statistically significant by trend at 6 W/kg CDMA was observed (National Toxicology Program 2018b), and the other one being significant by trend test when combining the data with the respective historical controls (Anderson et al. 2004). A non-statistically significant increase in malignant glia cell tumors was reported in the study by Falcioni for female rats (Falcioni et al. 2018). With respect to other types of brain neoplasms, the no RF EMF-related statistically significant findings were found in the NTP study in mice (National Toxicology Program 2018a).

2.2.2. Certainty assessment for brain-related carcinogenicity

The initial rating for the certainty was high because it involved controlled exposure, exposures occurred prior to outcome, outcomes were measured in individual animals, and a concurrent sham-exposed group was used. The evidence was rated as high (no factors decreasing the certainty and/or dose-response increasing the certainty). A meta-analysis was not conducted due to heterogeneity in the study design across the experiments, which included 5 chronic cancer bioassays. Among the 5 chronic cancer bioassays, key features of study design varied considerably (i.e., species, strain, sex, and exposure characteristics) (Table 4). Further, the small number of studies precluded meaningful subgroup analyses. A narrative description of the evidence was considered based on the findings from the chronic cancer bioassays because this is the most direct study design for assessing carcinogenicity. The BMD provides data on the increase in the risk per dose unit (SAR in W/kg) for one study on brain glioma in male rats (National Toxicology Program 2018b) (Table 5).

2.*2.2.2. Factors potentially decreasing certainty*

- RoB: No downgrade was applied because all five chronic cancer bioassays were rated as definitively or probably low RoB across all domains and considered Tier 1.
- Unexplained inconsistency: Downgrading was not proposed because glial cell-derived neoplasms were also found in the NTP study in male rats exposed to GSM, starting on GD5,with experimental conditions very similar to the CDMA exposed rats. Even though these findings were not statistically significant, no glioma were found in the sham-exposed group (Figure 11B). It might be expected that the frequency modulation for GSM or CDMA does not result in different cancer outcomes, but the difference in modulation of GSM and the CDMA may explain the different outcomes.

Heterogeneity in study designs (such as species, strain, sex and exposure characteristics might explain some of the apparent inconsistency across the chronic bioassays. While not significant, the study by Falcioni *et al*. showed an increase in malignant tumors in glial cells of female Sprague Dawley rats at much lower SAR levels. Of note, the sensitivity domain 'exposure contrast' was rated as having 'major concerns' (Falcioni et al. 2018).

- Indirectness: Downgrading for indirectness was not applied. Strong evidence on cancer in experimental animals is relevant to the identification of a carcinogenic hazard to humans (Baan et al. 2019). The assessment of brain gliomas in chronic cancer bioassays was considered direct with respect to the PECOS for animal model relevance. There was no downgrade for indirectness regarding the PECOS for human relevance, numbers for brain neoplasms and how they are related to the human brain cancers (Krewski et al. 2019). In the absence of information to the contrary, results from chronic cancer bioassays are typically presumed relevant to humans (Mevissen et al. 2022; National Toxicology Program 2019a).
- Imprecision: as in preclinical animal studies (Hooijmans et al. 2018), the direction of effect in animal toxicology studies is generally considered more important than the exact magnitude of the effect. The concept of statistical power (as a pre-study concept related to the post-study concept of precision) generally takes into account size of study population, prevalence of exposure, significance level, and desired power for an anticipated effect size. These determinants of study power are all defined in GLP cancer bioassay guidelines and have been adhered to in the two studies with an increased incidence in cancer, therefore there is no reason for downgrading the certainty for imprecision.
- Publication bias: No downgrade was applied for publication bias.

2.2.2.2. Factors potentially increasing certainty

- Dose-response: Several team members applied this upgrade factor. The calculation of the BMD for the NTP study resulted in a statistically significant increase of glial cell derived neoplasms per increment of exposure.

2.2. Heart

2.2.2. Characteristics of included studies for heart-related carcinogenicity

Four publications reported experiments that assessed heart lesions including predominantly malignant schwannomas and related lesions (Falcioni et al. 2018; Heikkinen et al. 2006; National Toxicology Program 2018a; b). Three studies were chronic cancer bioassays (National Toxicology Program 2018a; b) or natural life span (Falcioni et al. 2018), and one initiation-(co-)promotion study (Heikkinen et al. 2006). Among these, 3 studies were conducted in rats (2,365 animals), mostly in Sprague Dawley rats (Falcioni et al. 2018; National Toxicology Program 2018b) with one study using Wistar rats (Heikkinen et al. 2006), and one study was conducted in mice (720 animals) with two experiments, one in males and one in females (National Toxicology Program 2018a).

****Figure 13 is a thumbnail image of an interactive graphic in HAWC. The graphic presents the included studies and endpoints assessed across studies (Panel A) and summarizes study characteristics for the included studies (Panel B). Additional details on the studies can be accessed through the interactive dashboard as well as download of the underlying data in tabular format using the 'Actions' option.

This is a thumbnail image of an interactive dashboard available on HAWC that is filterable by citation, physiological or organ system, species, strain, and frequency. **Panel A** Numbers in the heatmap represent the number of studies that investigated a health system within a study design. If a study evaluated multiple health outcomes/endpoints or presented several experiments, it is shown here multiple times. **Panel B** provides additional experimental detail about the studies, including exposure characteristics. The table in the Panel B is only a partial representation of the full table available in HAWC.

2.2.2. Risk of bias and sensitivity considerations for heart-related carcinogenicity

There were no significant concerns for RoB in the chronic bioassay studies as shown in Figure 14 (Falcioni et al. 2018; National Toxicology Program 2018a; b). The only domain judgement of concern was an apparent lack of blinding of researchers in one study (Heikkinen et al. 2006) (Supplementary Figure S3A). Thus, all studies were considered Tier 1, and no sensitivity analyses were conducted based on RoB. Two studies had major sensitivity concerns regarding exposure contrast (Supplementary Figure S3B) (Falcioni et al. 2018; Heikkinen et al. 2006). More specifically, the concerns were due to the large uncertainty in the SAR in both the high and the low RF EMF exposure groups, which overlap.

2.2.2. Summary of the evidence for heart-related carcinogenicity

****Figure 15A presents an exposure response array of malignant neoplasms of the heart across the five experiments (presented in four studies) to summarize exposure-specific effects. This figure provides a graphical overview of findings across the studies, with respect to whether there was no effect (null), a statistically significant increase or decrease in incidence, or a statistically significant trend. The Bayesian Average 1% BMD was calculated when dose-response or a trend for an increase or decrease in the outcome was found. Evidence of dose response is presented in Figure 15B, which represents a thumbnail image of an interactive graphic in HAWC where study details are presented with the results. Supplementary Figure S3C displays the respective data for benign neoplasms and Supplementary Figure S3D shows preneoplastic lesions.

2.2.2.2. Effects of RF EMF on heart schwannoma in chronic bioassays

2.2.2.2.2. Rats

The studies by NTP (National Toxicology Program 2018b) and the Ramazzini Institute (Falcioni et al. 2018) provided evidence for an exposure-related effect, namely an increase in the tumor incidence of malignant heart schwannomas in male rats when compared to the sham-controlled animals.

In the CDMA experiment (National Toxicology Program 2018b), malignant heart schwannomas were found in all exposed male rat groups with the effect being SAR/dose-dependent, and the incidence in the 6 W/kg group was significantly increased, and a significant trend for dose-response was calculated by the authors. Malignant schwannomas were not found in the respective sham controls. A statistically significant increase in the incidence of endocardial Schwann cell hyperplasia was also found in the 6 W/kg group in male rats. Interpretation of these findings was not complicated by effects on body weight or clinical indications of toxicity as mean body weights of exposed groups of both sexes were similar to those of the sham controls, and no exposure-related clinical observations were found. With respect to historical controls, malignant schwannoma incidences exceed the number of these tumor types in historical controls in the NTP study (National

Toxicology Program 2018b). The survival at the end of the 2-year study was significantly greater in the male RF EMF exposed groups due to a chronic progressive nephropathy in the kidney of the sham control male rats. However, the statistical analysis for heart lesions was adjusted for the difference in survival. No effect was found in female in both the CDMA and the GSM experiment. In male rats in the GSM experiment, malignant heart schwannomas were found in all three dose/SAR groups with the highest incidence at the highest SAR values (6 W/kg) of 5.6%, but the increase was not statistically significant (Figure 15B) (National Toxicology Program 2018b).

The chronic bioassay performed by Falcioni et al. showed a statistically significant increase in malignant heart schwannomas in male rats at the highest SAR levels of 0.1 W/kg with pairwise testing when compared to sham-exposed rats (Falcioni et al. 2018). The SAR levels used were much lower (0.001, 0.03, and 0.1 W/kg) compared to those used in the NTP study (1.5, 3, and 6 W/kg). In addition, possible differences, especially the 10-fold difference in the BMD or heart schwannomas, between the Falcioni study and the NTP study in male rats might also be related to the different start of the exposure (GD5 in the NTP study), daily exposure times, namely 19 hours/day and 8 hours/day, respectively. Non-significant increases in Schwann cell hyperplasia were found at the high exposure for both males and females. Of note, the study by Falcioni et al (2018) was rated as 'major concerns' for the sensitivity domain of 'exposure contrast' where the relatively low exposure levels may preclude detection of an increased risk.

2.2.2.2.2. Mice

One chronic bioassay performed in mice, showed no effect of RF EMF on heart neoplasms or a dose-related effect regardless of the sex used (National Toxicology Program 2018a).

2.2.2.2. Effects of RF EMF on neoplastic lesions in initiation-(co-)promotion studies

2.2.2.2.2. Rats

The study by Heikkinen et al., an initiation (co)-promotion study, did not show effects on neoplasms or neoplastic lesions of the heart (Heikkinen et al. 2006).

2.2.2.2.2. Mice

No initiation-(co-)promotion studies were performed in mice.

2.2.2.2. Effects of RF EMF on benign tumors or preneoplastic of the heart

2.2.2.2.2. Rats and mice

No effects of RF EMF were found for benign tumors (Supplementary Figure S3C) or preneoplastic lesions (Supplementary Figure S3D) of the heart when compared to sham-exposed animals in any of the studies evaluated.

2.2.2. Certainty assessment for heart-related carcinogenicity

All experimental studies in nature resulting in an initial rating for the certainty in the animal studies as high (Mevissen et al. 2022). Certainty was not downgraded for any factor (Table 7). A meta-analysis was not conducted due to variation in study design across the experiments, which included 3 chronic bioassays and one initiation-(co-)promotion study.

Even among the 3 chronic bioassays, key features of study design varied considerably (i.e., species, SAR modulation/frequency/exposure levels, exposure duration per day, and presence or absence of *in utero* exposure) (Table 6). Further, the small number of studies precluded meaningful subgroup analyses. A narrative description of the evidence was considered based on the findings from the 3 chronic cancer bioassays because this is the most direct study design for assessing carcinogenicity (Falcioni et al. 2018; National Toxicology Program 2018b). The BMD provides data on the increase in the risk per dose unit (SAR in W/kg) for both studies (Table 7).

2.2.2.2. Consideration of Factors potentially decreasing certainty

- RoB: No downgrade was applied because studies were all Tier 1 with definitively or probably low RoB judgements across the vast majority of domains.
- Unexplained inconsistency: No downgrade was applied for unexplained inconsistency because study designs are quite heterogeneous, which is why the data were not considered suitable for meta-analysis. That said, there was consistency in findings from the chronic bioassays for increased heart schwannoma incidence in male Sprague Dawley rats, whereas no significant increases were found in female rats or mice of both sexes. The apparent inconsistency in findings was considered to be plausibly explainable and may represent a potential susceptibility based on species and sex.
- Indirectness: There was no downgrade for indirectness because findings were based on chronic cancer bioassays. Regarding human relevance, systematic analyses of animal heart neoplasms and how they relate to the human heart cancers are not available (Krewski et al. 2019). However, in the absence of strong evidence that the tumors in animals are induced by a mechanism that cannot operate in humans, results from chronic cancer bioassays are typically presumed relevant to humans (Mevissen et al. 2022; National Toxicology Program 2019a).
- Imprecision: As in preclinical animal studies (Hooijmans et al. 2018), the direction of effect in animal toxicology studies is generally considered more important than the exact magnitude of the effect. The concept of statistical power (as a pre-study concept related to the post-study concept of precision) generally takes into account size of study population, prevalence of exposure, significance level, and desired power for an anticipated effect size. These determinants of study power are all defined in GLP cancer bioassay guidelines and have been adhered to in the two studies with an increased incidence in cancer as 50 animals per treatment arm are considered standard for chronic cancer bioassay.
- Publication bias: No downgrade was applied for publication bias.

2.2.2.2. Consideration of factors potentially increasing certainty

No factors for increasing certainty were applied because the initial certainty was high and there were no downgrades. However, below we summarize the upgrade factors described in our protocol, noting instances when they seemed applicable.

- Dose response: The NTP study reported a statistically significant monotonic positive trend across all tested dose/SAR levels in male rats. Further, Falcioni et al reported a pair-wise statistically significant increase at the highest dose/SAR level.

- Rare outcomes: Malignant schwannoma is a rare tumor in rodents. We note that even if we had applied a downgrade factor, we would have upgraded because of the rarity of the tumor types.

2.2. Adrenal gland

2.2.2. Characteristics of included studies for adrenal gland-related carcinogenicity

Fifteen experiments presented in 12 publications assessed tumors of the adrenal gland (Chou et al. 1992; Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Jauchem et al. 2001; La Regina et al. 2003; National Toxicology Program 2018a; b; Oberto et al. 2007; Tillmann et al. 2007; Toler et al. 1997) (19). Five publications report chronic cancer bioassays (Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018a; b; Tillmann et al. 2007), two publications report initiation-(co-)promotion studies (Heikkinen et al. 2006; Heikkinen et al. 2001), and five studies on tumor-prone mice (Frei et al. 1998a; Frei et al. 1998b; Jauchem et al. 2001; Oberto et al. 2007; Toler et al. 1997). Among these studies a total of 1,456 rats and 2,207 mice were used.

****Figure 16 is a thumbnail image of an interactive graphic in HAWC. The graphic presents the included studies and endpoints assessed across studies (Panel A) and summarizes study characteristics for the included studies (Panel B). Additional details on the studies can be accessed through the interactive dashboard as well as download of the underlying data in tabular format using the 'Actions' option.

This is a thumbnail image of an interactive dashboard available on HAWC that is filterable by citation, health system, species, strain, and frequency. **Panel A** Numbers in the heatmap represent the number of studies that investigated a health system within a study design. If a study evaluated multiple health outcomes or presented several experiments, it is shown here multiple times. **Panel B** provides additional experimental detail about the studies, including exposure characteristics. The table in the Panel B is only a partial representation of the full table available in HAWC.

2.2.2. Risk of bias and sensitivity considerations for adrenal gland-related carcinogenicity

Overall, there were no significant concerns for RoB in chronic bioassays (Figure 17), Initiation- (co)-promotion (Supplementary Figure S4A), tumor-prone (Supplementary Figure S4B) and all were considered Tier 1. Four studies had major sensitivity concerns (Supplementary Figure S4C), especially regarding exposure contrast (Chou et al. 1992; Heikkinen et al. 2006; Jauchem et al. 2001) that are either due to a wide and overlapping range of SAR levels (Heikkinen et al. 2006; Jauchem et al. 2001), or very low SAR levels that were close to background levels (Jauchem et al. 2001). One study had major sensitivity concerns related to endpoint sensitivity and specificity (Oberto et al. 2007).

2.2.2. Summary of the evidence for adrenal gland-related carcinogenicity

****Figure 18A presents an exposure response array of malignant or combined (benign and malignant) pheochromocytoma of the adrenal gland presented in 4 publications to summarize exposure-specific effects. This figure provides a graphical overview of findings across the studies,

with respect to whether there was no effect (null), a statistically significant increase or decrease in incidence, or a statistically significant trend. Evidence of dose response for malignant or combined (benign and malignant) pheochromocytomas is presented in Figure 18B, which represents a thumbnail image of an interactive graphic in HAWC where study details can be easily read. The 1% Bayesian Average BMD was not calculated because there was no dose-response or trend for an increase or decrease in pheochromocytoma.

2.2.2.2. Effects of RF EMF on tumors of the adrenal gland, pheochromocytomas

2.2.2.2.2. Rats and mice

Some evidence for RF EMF-related effects were provided for pheochromocytomas in the NTP rat study (National Toxicology Program 2018b). More specifically, pheochromocytomas (benign, or combined) were increased in the F1 generation male Sprague Dawley rats exposed to 900 MHz (GSM) at SAR levels of 1.5 and 3 W/kg, but no dose-dependent effect was observed (Figures 18 A and 18B). In female Sprague Dawley rats exposed to 900 MHz (CDMA), an increase in benign pheochromocytomas was observed at a SAR level of 1.5 W/kg (Supplementary Figure S4F), and an increase in benign, malignant, or complex pheochromocytomas was observed at 1.5 W/kg when compared to the respective sham controls (Figure 18B). None of these findings were dose-dependent (Figure 18B). Statistically significant decreases in pheochromocytomas in female mice were found in the initiation-(co-)promotion study by Heikkinen et al. at the highest dose of either 0.35 or 1.5 W/kg in the two experiments where ionizing radiation has been used for tumor initiation (Heikkinen et al. 2001) (Supplementary Figures S4D, S4E).

Single pheochromocytomas, benign and malignant, were provided in the NTP mouse report (National Toxicology Program 2018a) in male and female mice. However, the effects were not statisticall significant in any of the dose/SAR groups investigated (Figure 18B, Supplementary Figures S4F, S4G).

2.2.2.2. Effects of RF EMF on other malignant tumors than pheochromocytomas of the adrenal gland

2.2.2.2.2. Rats and mice

Results for malignant adrenal tumors besides pheochromocytomas were null regardless of the species, sex or study type. Effects of RF EMF on other benign tumors than pheochromocytomas and preneoplastic lesions of the adrenal gland were not found.

2.2.2.2.2. Rats and mice

No effects of RF EMF were found for benign neoplasms besides pheochromocytomas of the adrenal gland in any species, sex or study type as illustrated in Supplementary Figure S4H.

With respect to increases in preneoplastic lesions, in the F1 generation, female Sprague Dawley rats exposed to 900 MHz (GSM), adrenal cortex hyperplasia was reported at 3 and 6 W/kg, adrenal cortex vacuolization at 6 W/kg, and a trend for an increase in hyperplasia of the adrenal medulla (National Toxicology Program 2018a) (Supplementary Figure S4I, S4J). However, the effects were not dose-dependent. An increase in adrenal hypertrophy was also seen in the F1 generation of male Sprague Dawley in the 900 MHz (GSM) study at 6 W/kg.

Certainty assessment for adrenal gland-related carcinogenicity

The initial rating for the certainty in the animal studies was high. Certainty was downgraded by one level for unexplained inconsistency (Table 9). A meta-analysis was not conducted due to variation in study design across the experiments, which included 5 chronic bioassays, two one initiation-(co-)promotion studies and 5 studies in tumor-prone mice. Even among the 5 chronic bioassays, key features of study design varied considerably (i.e., species, SAR modulation/frequency/exposure levels, and presence or absence of *in utero* exposure) (Table 8). Further, the small number of studies precluded meaningful subgroup analyses. A narrative description of the evidence was considered based on the findings from the 5 chronic bioassays because this is the most direct study design for assessing carcinogenicity. No BMD was calculated for any of the studies for pheochromocytoma because due to lack of clear monotonic dose response.

2.2.2.2. Consideration of Factors potentially decreasing certainty

- *Risk of bias:* No downgrade was applied because studies were all Tier 1 with definitely or probably low risk of bias judgements across the vast majority of domains.

Certainty was downgraded one level because the positive findings were reported only at low or low and medium-level of exposure in two experiments: F1 generation male rats (pheochromocytomas regardless of being benign or malignant or a combination at 900 MHz GSM SAR levels of 1.5 and 3 W/kg), and female Sprague Dawley (benign pheochromocytomas at a 900 MHz CDMA SAR level of 3 W/kg) and (National Toxicology Program 2018b). This may raise concern about the biological plausibility of these findings.

- Unexplained inconsistency: (National Toxicology Program 2018b)Certainty was not further downgraded for unexplained inconsistency even though there was some inconsistency in the results across the 15 experiments which may be related to heterogeneity of the study design (the animal model, species, strain, exposure conditions). Most notably, the exposure with effects in pheochromocytomas was found in the NTP rat study at developmental exposure (National Toxicology Program 2018b). Furthermore, the animals were allowed to roam freely in their cages in the NTP study as well as in the Chou study (Chou et al. 1992; National Toxicology Program 2018b), whereas all other studies used a Ferris wheel setup where animals were contained within narrow tubes during exposure. The NTP study used SAR levels up to 6 W/kg (National Toxicology Program 2018b). With the exception of the NTP rat studies, most studies reported no statistically significant increases in pheochromocytoma and related lesions in the adrenal gland. In the tumor-prone animals as well as in all other mouse studies, there were no statistically significant increases, and no dose-dependent trends were found in any of the studies. It is important to note that there were considerable differences in in the exposure conditions and exposure onset used in the NTP study compared to all other studies. Besides the NTP study, three other studies were performed in rats (Chou et al. 1992; Heikkinen et al. 2006; La Regina et al. 2003). One also used male Sprague Dawley rats, but the exposure characteristics were different regarding frequency and SAR level (2.45 GHz, SAR: 0.15 to 0.4 W/kg, depending on the weight of the animals (Chou et al. 1992), the other study was done in female Wistar rats, but the SAR levels used were substantially lower (900 MHz, SAR: 0.3, 0.9 W/kg) (Heikkinen et al. 2006) compared to the NTP study (Heikkinen et al. 2006). Furthermore, the quality of the studies had some limitations (see RoB and sensitivity). Altogether, the fact that the NTP study (National Toxicology Program 2018b) describes positive findings in adrenal pheochromocytomas and related histopathologic lesions in male Sprague

Dawley rats at developmental, even though not dose-dependent, supports the certainty evaluation of this finding.

- Indirectness: Regarding directness to the animal PECOS, no downgrade was applied for indirectness because findings were based on chronic cancer bioassays. Strong evidence on cancer in experimental animals is relevant to the identification of a carcinogenic hazard to humans (IARC SP165). Regarding the PECOS for human relevance, there is no *a priori* mechanistic or other biological evidence to suggest animal models are not reasonable for predicting potential cancer effects in humans for the tumor types considered in this analysis. Further, in additional analyses of the tumor site concordance between experimental animals and humans, Krewski et al. compared agent-specific concordance between (groups of) cancer sites with sufficient evidence in humans and in animals (the latter required replication of positive results at the same specific site in at least two animal experiments). The overlap for endocrine tumors was 100%, albeit based on only two agents (Krewski et al. 2019). That stated, in most hazard characterization frameworks, tumor site concordance between animal models and humans is not required to identify an exposure as a potential carcinogen. When present, tumor site concordance can increase confidence in the findings, particularly for human risk assessment.

- Imprecision: As in preclinical animal studies (Hooijmans et al. 2018), the direction of effect in animal toxicology studies is generally considered more important than the exact magnitude of the effect. The concept of statistical power (as a pre-study concept related to the post-study concept of precision) generally takes into account size of study population, prevalence of exposure, significance level, and desired power for an anticipated effect size. These determinants of study power are all defined in GLP cancer bioassay guidelines and have been adhered to in the two studies with an increased incidence in cancer. Therefore, there is no reason for downgrading the certainty for imprecision.

- Publication bias: No downgrade was applied for publication bias.

2.2.2.2. Consideration of Factors potentially increasing certainty

No upgrade factors were applied. The non-neoplastic data provided additional data on occurrence of abnormal cell growth and cell populations in the NTP GSM rat experiment where a pheochromocytoma response was observed, supporting plausibility of a RF EMF effect on adrenal tissue as tested. However, the certainty of evidence was not upgraded due to substantial inconsistency between the NTP study results compared to all other studies and the lack of a dose response (extending to high doses) in the NTP study.

2.2. Liver

2.2.2. Characteristics of included studies for liver-related carcinogenicity

Twenty-six experiments presented in 16 publications assessed neoplastic lesions, and related histopathologic abnormalities in liver, and in 14 of these studies, specific tumor types were reported (Chou et al. 1992; Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Imaida et al. 1998a; Imaida et al. 1998b; Jauchem et al. 2001; La Regina et al. 2003; Lerchl et al. 2015; National Toxicology Program 2018a; b; Oberto et al. 2007; Tillmann et al. 2007; Tillmann et al. 2010; Toler et al. 1997). Among these 26 experiments, 10 were conducted in mice (3,201 animals), and 6 in rats (1,483 animals). The 16 publications include 5 chronic cancer bioassays (Chou et al. 1992; La Regina et al. 2003;

National Toxicology Program 2018a; b; Tillmann et al. 2007), 6 initiation(co-)promotion studies (Heikkinen et al. 2006; Heikkinen et al. 2001; Imaida et al. 1998a; Imaida et al. 1998b; Lerchl et al. 2015; Tillmann et al. 2010), and 5 studies with tumor-prone mice (Frei et al. 1998a; Frei et al. 1998b; Jauchem et al. 2001; Oberto et al. 2007; Toler et al. 1997).

****Figure 19 is a thumbnail image of an interactive graphic in HAWC. The graphic presents the included studies and endpoints assessed across studies (Panel A) and summarizes study characteristics for the included studies (Panel B). Additional details on the studies can be accessed through the interactive dashboard as well as download of the underlying data in tabular format using the 'Actions' option.

2.2.2. Risk of bias and sensitivity considerations

There were some concerns for RoB in some studies (Figure 20 and Supplementary Figures S5A, S5B). Six studies had limitations regarding the blinding during the study (Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Imaida et al. 1998a; Jauchem et al. 2001), and one study had major limitations regarding randomization (Lerchl et al. 2015). Two studies (Imaida et al. 1998a; Imaida et al. 1998b) have high RoB in the temperature domain. Thus, these two studies were considered Tier 2, whereas the other studies were considered Tier 1. Four studies had major sensitivity issues regarding exposure contrast (Chou et al. 1992; Heikkinen et al. 2006; Imaida et al. 1998b; Jauchem et al. 2001) (Supplementary Figure S5C).

2.2.2. Summary of the evidence for liver-related carcinogenicity

****Figure 21 presents an exposure response array of malignant neoplasms of the liver across studies to summarize exposure-specific level effects. This figure presents a graphical overview of the findings across studies with respect to whether there was no effect (null), a statistically significant increase of decrease in incidence, or a statistically significant trend. Data on combined carcinomas and adenomas was typically not presented (and could not be calculated) in the available studies. Only one study presented the incidence of both tumor types separately (National Toxicology Program 2018a). The NTP study reported a statistically non-significant increase of liver adenomas and carcinomas (combined) at mid-level dose in male rats for the CDMA exposure. The 1% Bayesian BMD was not calculated because there was no dose-response or trend for an increase or decrease in liver neoplasms.

2.2.2.2. Effects of RF EMF on other malignant tumors of the liver in chronic bioassays

2.2.2.2.2. Mice

In the chronic bioassays, the only increased malignant neoplastic finding was for liver hepatoblastoma in male B6C3F1/N mice exposed to 1900 MHz CDMA at a SAR level of 5 W/kg (National Toxicology Program 2018a) (Figures 21 and 22). However, this finding did not display evidence for dose-response. In the same study, a statistically significant decrease in hepatocellular carcinoma was found in male mice at a SAR level of 2.5 W/kg compared to sham exposure.

No effects on malignant liver tumors were found in the other 3 chronic cancer bioassays in mice when compared to the sham controls with the tumor types in the study by La Regina and colleagues not being mentioned (Chou et al. 1992; La Regina et al. 2003; Tillmann et al. 2007). No increases in combined carcinomas and adenomas were observed in any study, although the number of studies reporting this information was small (National Toxicology Program 2018a).

2.2.2.2.2. Rats

In the one chronic cancer bioassay (CDMA) performed in male Sprague Dawley rats. In male rats, there were increased incidences of hepatocellular adenoma in the 1.5 and 3 W/kg groups compared to sham controls. There were also hepatocellular carcinomas, one each in the 3 and 6 W/kg groups (Figure 22). The incidences of hepatocellular adenoma or carcinoma (combined) were not statistically significant (CDMA). In 6 W/kg females, there was a decreased incidence of hepatocellular adenoma with a significant negative trend (National Toxicology Program 2018a).

2.2.2.2. Effects of RF EMF on malignant tumors of the liver in initiation-(co-)promotion studies

2.2.2.2.2. Mice

In the tumor initiation-(co-)promotion studies, Lerchl et al. 2015 reported a significant increase in malignant liver carcinomas in mice at 0.04, 0.4 and 2 W/kg in a mixed group of male and female mice with the effect not being dose-dependent (Lerchl et al. 2015) (Supplementary Figure S5D). In the Tillman 2010 initiation-(co-)promotion study with male and female mice, a difference in liver carcinomas was only found when ENU-treated mice were compared to cage controls, and a combined exposure with RF EMF and ENU resulted in a significant increase in tumor incidence of liver carcinomas at 0.58 W/kg but not at 2 W/kg (Supplementary Figure S5D) (Tillmann et al. 2010).

2.2.2.2.2. Rats

Heikkinen and colleagues reported no effects of RF EMF in female rats when compared to sham-exposed animals (Supplementary Figure S5D) (Heikkinen et al. 2006).

2.2.2.2. Effects of RF EMF on malignant tumors of the liver in tumor-prone mice

No effects of RF EMF on malignant neoplasms were observed in tumor-prone mice of either sex (Supplementary Figure S5E) (Frei et al. 1998a; Frei et al. 1998b; Oberto et al. 2007; Toler et al. 1997). In the studies performed in tumor-prone mice, a significant trend was found in liver vascular tumors (considered a malignant neoplasm) in one study in male tumor-prone (for lymphomas) with the effect being statistically significant at the highest SAR/dose level of 4 W/kg (Oberto et al. 2007) (Supplementary Figure S5E).

2.2.2.2. Effects of RF EMF on benign tumors

2.2.2.2.2. Mice

RF EMF did not have compelling effects on benign liver tumors in mice regardless of the sex and the study type (Supplementary Figure S5F). The histopathological findings for the liver showed some effects (both increases and decreases) in hepatocellular benign neoplasms in three studies in mice, namely two chronic bioassays (National Toxicology Program 2018a; Tillmann et al. 2007) (Supplementary Figure S5F). With respect to benign neoplastic findings, a trend for a significant and dose-dependent decrease in benign liver

adenomas was found in the study of Tillmann et al. for male mice regardless of the frequency (900 MHz or 1747 MHz, DCS) (Supplementary Figure S5F) (Tillmann et al. 2007). In both NTP experiments with mice, some hepatocellular adenomas were increased, but no dose-response was found (Supplementary Figure S5F) (National Toxicology Program 2018a). No effect of RF EMF was found in initiation-(co-)promotion or tumor-prone mice (Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Lerchl et al. 2015; Oberto et al. 2007; Tillmann et al. 2010; Toler et al. 1997).

2.2.2.2.2. Rats

No effects of RF EMF exposure on benign tumors were found in rats when compared to sham-controls. In the chronic cancer bioassays, the only preneoplastic lesion observed was for mixed foci in male Sprague-Dawley rats exposed to 900 MHz CDMA at a SAR level of 3 W/kg (Supplementary Figure S5G). However, this finding did not display evidence of dose-response gradient. Foci of cellular alteration, with a statistically significant positive trend were found in one study in female Wistar rats using an initiation-(co-)promotion model (Heikkinen et al. 2006).

Journal Pre-proofs

2.2.2. Certainty assessment for liver-related carcinogenicity

The initial rating for the certainty in the animal studies was set to high. The evidence was downgraded by one level because a statistically significant increase in hepatoblastomas was only reported in one chronic cancer bioassay. No downgrade for directness was utilized and the rationale for this is provided below. No factors that potentially increase certainty were utilized. Thus, the overall certainty was considered moderate (Table 11).

Sixteen studies investigated tumors of the liver with 5 of these being chronic cancer bioassays, 6 initiation-(co-)promotion studies and 5 studies in tumor-prone mice. Overall, a total of 1,483 rats and 3,201 mice was used. None of the findings was dose-dependent. A meta-analysis was not conducted due to variation in study design across the experiments. Even among the chronic cancer bioassays, key features of study design varied considerably (i.e., species, SAR modulation/frequency/exposure levels, and presence or absence of *in utero* exposure) (Table 10). Further, the small number of studies precluded meaningful subgroup analyses. A narrative description of the evidence was considered based on the findings from the 5 chronic bioassays because this is the most direct study design for assessing carcinogenicity (Falcioni et al. 2018; National Toxicology Program 2018b). The CoE was evaluated as moderate for hepatoblastomas due to a downgrade by one level because of unexplained inconsistency, and absence of factors increasing confidence. The BMD was not calculated for hepatoblastomas and hepatocellular carcinomas due to the lack of a monotonous dose-response.

2.2.2.2. Consideration of Factors potentially decreasing certainty

- *Risk of bias:* No downgrade was applied because studies were all Tier 1 or Tier 2 with definitely or probably low risk of bias judgments across the vast majority of domains.

- Unexplained inconsistency: Certainty was downgraded for unexplained inconsistency in hepatoblastoma findings. A statistically significant increase in hepatoblastomas was observed in male B6C3F1/N mice exposed to 1900 MHz CDMA at a SAR level of 5 W/kg(National Toxicology Program 2018a). However, the effect was only observed in a middle dose. No effects were found male mice in the GSM experiment. Besides the NTP study in male mice, there is no support for an increased risk of hepatoblastomas from a second study. Similarly for hepatocellular adenomas and/or carcinomas the inconsistency cannot be easily explained.

- Indirectness: Regarding directness to the animal PECOS, no downgrade was applied. The studies in tumor-prone animals or using an initiation-(co-)promotion design are considered appropriate models for assessment of liver carcinogenicity in rodents.

- Publication bias: No downgrade was applied for publication bias.

2.2.2.2. Consideration of Factors potentially increasing certainty

- None applied

2.2. Lung

2.2.2. Characteristics of included studies for lung-related carcinogenicity

Eleven publications reported neoplastic lesions and related histopathology findings of the lung (Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Jauchem et al. 2001; Lerchl et al. 2015; National Toxicology Program 2018a; Oberto et al. 2007; Sommer et al. 2004; Tillmann et al. 2007; Tillmann et al. 2010).

Among these publications, 23 studies were conducted in mice (2,800), and 8 in rats (1,296). The 11publications include 2 chronic bioassays (National Toxicology Program 2018a; Tillmann et al. 2007), 4 initiation-promotion studies (Heikkinen et al. 2006; Heikkinen et al. 2001; Lerchl et al. 2015; Tillmann et al. 2010), and 5 studies in tumor-prone mice (Frei et al. 1998a; Frei et al. 1998b; Jauchem et al. 2001; Oberto et al. 2007; Sommer et al. 2004). In addition, there were 3 more bioassays, where only premalignant lesions, or metastases from other neoplasias occurred in the lung (Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018b) or neoplasms were only given for one dose/SAR (National Toxicology Program 2018b). These studies are included in the overview figure (Figure 23).

2.2.2. Risk of bias and sensitivity considerations

There were no concerns for RoB in the 2 chronic bioassays (Figure 24), some concern in one domain in the initiation-promotion studies (Supplementary Figure S6A) and some concern in one domain in 4 studies with tumor-prone mice, S6B). Six studies had major limitations regarding the blinding during the study (Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Jauchem et al. 2001; Oberto et al. 2007; Sommer et al. 2004), and one study had major limitations regarding randomization (Lerchl et al. 2015). All studies were considered Tier 1.

No significant sensitivity concerns were identified in the chronic bioassays (Supplementary Figure S6C). Two other studies had major sensitivity issues regarding exposure contrast (Heikkinen et al. 2006; Jauchem et al. 2001) (Supplementary Figures S6D, S6E).

2.2.2. Summary of the evidence for lung-related carcinogenicity

****Figure 25 presents the exposure response array of malignant neoplasms of the lung across studies to summarize exposure-specific effects. This figure presents a graphical overview of the findings across studies with respect to whether there was no effect (null), a statistically significant increase or decrease in incidence, or a statistically significant trend. One study presented the incidence of bronchioalveolar adenoma or carcinoma (combined) and both tumor types separately (National Toxicology Program 2018a). The NTP study reported a statistically significant positive trend of both bronchioalveolar adenoma or carcinoma (combined) for the GSM exposure in male mice. A dose-dependent not statistically significant increase in carcinoma was found (National Toxicology Program 2018a). In two initiation-promotion studies (Lerchl et al. 2015; Tillmann et al. 2010) with mice, performed in a similar way but different mouse strains and slightly different SAR levels, significant increases in bronchioalveolar adenoma and also for carcinoma at the middle dose/SAR level of 0.58 W/kg were found in male and female B63F1 mice (combined) at 1966 MHz UMTS at chronic exposure for 24 months (Tillmann et al. 2010). Statistically significant increases in bronchioalveolar adenoma in male CH3/HeNCrl mice and female C57Bi/6N mice (combined) for 72 weeks at 1966 MHz, were found in all doses/SAR levels (Lerchl et al. 2015).

2.2.2.2. Effects of RF EMF on neoplastic lesions of the lung in chronic bioassays

2.2.2.2.2. Mice and rats

In mice, bronchioalveolar adenomas and carcinomas as well as the combination of both neoplasms were demonstrated in the other experiments in the two bioassays for male and female mice (CDMA, GSM), but, besides the statistically significant positive trend in adenomas or carcinomas (combined) in male mice at GSM exposure (National Toxicology Program 2018a), the findings were not statistically significant, nor was a trend for dose-response reported (Figures 25 and 26). The tumor incidence for all tumors found was lower in female mice when compared to male mice in the NTP study (National Toxicology Program 2018a), and the same holds true for the Tillmann study, except for bronchoalveolar adenomas (Tillmann et al. 2010).

In rats, occasional findings of lung neoplasms were reported in the NTP study, but only at certain SAR levels.

2.2.2.2. Effects of RF EMF on neoplastic lesions in the initiation-(co)promotion studies

2.2.2.2.2. Mice and rats

As stated above, in B63F1mice (combined males and females) exposed at 1966 MHz for 24 months (20 hours/day), statistically significant increases were found in bronchoalveolar adenomas and carcinomas (combined) at the middle dose, a SAR level of 0.5 W/kg (Tillmann et al. 2010). In a study with a similar design, statistically significant increases in bronchioalveolar adenoma in male CH3/HeNCrl mice and female C57Bi/6N mice (combined) for 72 weeks at 1966 MHz, were found at all doses/SAR levels (Lerchl et al. 2015) (Supplementary Figures S6F and S6G).

In another study in female CBA/S mice, exposed for 1.5 hours/day, 5 days/week for 78 weeks at 902 MHz GSM or continuous exposure, no significant increases in lung neoplastic lesions were reported (Heikkinen et al. 2001). Similar to the NTP study in rats, only very few lung tumors were found in female Wistar rats, and GSM exposure did not affect the tumor incidence of any of lung neoplastic lesions (Heikkinen et al. 2006).

2.2.2.2. Effects of RF EMF on neoplastic lesions of the lung in studies with tumor-prone mice

No effects of RF EMF on malignant or benign neoplasms were observed in tumor-prone mice of either sex (Supplementary Figures S6H and S6I) (Frei et al. 1998a; Frei et al. 1998b; Oberto et al. 2007). In the study by Oberto et al., an increase in bronchoalveolar adenomas was found at the highest dose of 4 W/kg in male pim1 mice exposed for 1 hour/day, 5 days/week for 18 months at a 217 Hz pulsed 900 MHz TF EMF (Oberto et al. 2007) (Supplementary Figures S6H and S6I). The studies by Frei et al. used female C3H/HeJ mice that were exposed at 2.54 GHz for 20 hours/day for 18 months and a SAR of 0.3 W/kg, and no differences in lung tumors were obtained when sham and RF EMD-exposed mice were compared (Frei et al. 1998a; Frei et al. 1998b).

2.2.2. Certainty assessment for lung-related carcinogenicity

The initial rating for the certainty in the animal studies was set to high. The evidence was downgraded by one level for unexplained inconsistency in bronchioalveolar adenoma and carcinoma findings in one cancer bioassay (National Toxicology Program 2018a). No downgrade for directness was utilized and the rationale for this is provided below. No factors

that potentially increase certainty were utilized. Thus, the overall certainty was considered moderate (Table 11).

Sixteen studies investigated tumors of the lung with 2 of these studies who reported lung neoplasias are chronic cancer bioassays, 4 initiation-(co-)promotion studies and 5 studies in tumor-prone mice. In these studies, a total of 1,296 rats and 2,800 mice was used. A statistically significant positive trend was found in one chronic bioassay in bronchoalveolar adenomas or carcinomas (combined). None of the direct pairwise comparisons was statistically significant (National Toxicology Program 2018a). A meta-analysis was not conducted due to variation in study design across the experiments. Even among the chronic cancer bioassays, key features of study design varied considerably (i.e., species, SAR modulation/frequency/exposure levels) (Table 12). Further, the small number of studies precluded meaningful subgroup analyses. A narrative description of the evidence was considered based on the findings from the 2 chronic bioassays because this is the most direct study design for assessing carcinogenicity (National Toxicology Program 2018a; Tillmann et al. 2007). The CoE was evaluated as moderate for lung neoplasms due to a downgrade by one level because of unexplained inconsistency, and absence of factors increasing confidence. The BMD was calculated for bronchioalveolar adenomas or carcinomas (combined) (0.729, CI low 0.197), and bronchioalveolar carcinomas (2.324, 95% CI 0.497, 24.318).

2.2.2.2. Consideration of Factors potentially decreasing certainty

- *Risk of bias:* No downgrade was applied because studies were all Tier 1 with definitively or probably low risk of bias judgments across the vast majority of domains. The chronic bioassays (National Toxicology Program 2018a; Tillmann et al. 2007) had no limitations in RoB or study sensitivity.

- Unexplained inconsistency: Certainty was downgraded for unexplained inconsistency in bronchioalveolar adenoma and carcinoma findings. A statistically significant positive trend was found for bronchioalveolar adenomas or carcinomas (combined) observed in male B6C3F1/N mice exposed to 1900 MHz GSM (National Toxicology Program 2018a). None of the direct comparisons per SAR level (2.5, 5, and 10 w/kg) were statistically significant. No effects were found male mice in the CDMA experiment, and no effects were reported in female mice (National Toxicology Program 2018a). Besides the NTP study at GSM exposure in male mice, there is some support for an increased risk of adenomas (Lerchl et al. 2015; Tillmann et al. 2010) and carcinomas (Tillmann et al. 2010) in initiation-promotion(-co)promotion studies. The inconsistency for these neoplasms cannot be easily explained because both studies followed a similar design. However, there are differences in the mouse strains and the SAR levels.

- Indirectness: Regarding directness to the animal PECOS, no downgrade was applied. The initiation-(co-)promotion design is considered appropriate model for assessment of lung carcinogenicity in rodents.

- Publication bias: No downgrade was applied for publication bias.

2.2.2.2. Consideration of Factors potentially increasing certainty

- None applied

3. Discussion

The work in this publication represents a systematic review of the carcinogenic effects of radiofrequency exposure in laboratory animals in manuscripts published until 2022. A summary of the confidence judgments for each of the main types of cancer evaluated (lymphoma, brain, heart, adrenal gland,liver, and lung) are summarized in Table 13. High CoE can be interpreted as the true relationship is highly likely to be reflected in the apparent relationship. Moderate CoE indicates the true relationship may be reflected in the apparent relationship.

2.2. Summary of Findings

A total of 52 animal cancer bioassays, namely, chronic cancer bioassays, initiation-(co-) promotion studies, and studies in tumor-prone mice were eligible and evaluated according to the protocol (Mevissen et al. 2022). Of these, chronic rodent bioassays are generally considered the most informative study design for detecting chemical carcinogens, recognizing that factors such as how animals are exposed, sex, species, strain, stubstrain, background rates and variability in spontaneous tumors can all affect specific study outcomes (Bannasch et al. 1986; Della Porta and Dragani 1990). Besides different study designs, a variety of different animal models, exposure metrics, duration periods of exposure, and other parameters were reviewed. Chronic rodent bioassays have reached a high standard of performance for detection of chemical carcinogens. Nevertheless, different öis, length of the testing, selection and use of strains, especially variability of spontaneous tumors, affect the outcome (Bannasch et al. 1986; Della Porta and Dragani 1990). Overall, the evidence base is very heterogeneous and does not support reaching strong conclusions on the impact of specific study design variables, including radiation type, species, strain as potential sources of variation in results across studies. To help address the focused question on radiation type specific effects, supplementary Figures (S7, S8, S9, S10, S11, S12) were developed for tumor sites where multiple forms of radiation were administered across studies (or study arms in the same publication), and some evidence of carcinogenicity was observed in at least one study (or study arm). For high CoE, two positive long-term bioassays (for 2 sites, heart (heart schwannoma) and brain (glioma)) with the positive studies differing in several regards (animal model, radiation type, dosage) were identified. For lung adenomas or carcinomas (combined), one chronic bioassay showed a significant positive trend in male mice exposed to GSM (National Toxicology Program 2018a). No pairwise comparisons were statistically significant at any SAR level investigated in that study. Two initiation-(co-)promotion studies with a similar design gave some evidence for an increase in bronchoalveolar adenomas and/or carcinomas in mice of both sexes (combined) (Lerchl et al. 2015; Tillmann et al. 2010).

The team concluded there was high CoE for an increased incidence of malignant schwannomas of the heart based mainly on results from two chronic bioassays in male rats. With respect to the magnitude of the effect, a significant trend for an increase in endocardial malignant schwannomas in male Sprague Dawley rats was observed in one study with a statistically significant trend in malignant schwannomas across the dose range of a whole

body average SAR from 1.5 W/kg to 6 W/kg given over 106 weeks and the increase also being statistically significant at 6 W/kg (National Toxicology Program 2018b). In a second study, an increased incidence in total schwannoma of the heart in male Sprague Dawley rats was observed at 0.1 W/kg (whole body average SAR) given over life span (Falcioni et al. 2018). Sex differences in these rare tumors, namely schwannoma in Sprague Dawley rats have been reported in historic controls with an incidence of 0.7 in males rats and 0.1 in female rats (Kumar A. 2024). Regarding the directness of the PECOS of the animal model to the evidence in humans, heart schwannomas are rare tumors in humans. However, neoplasms derived from Schwann cells can basically occur at any place in the body where nerve sheaths exists. Schwannomas are the most common neoplasms of the peripheral nerve sheath, which also include neurofibromas, perineurinomas, granular cell tumors and malignant peripheral nerve sheath tumors. A vestibular schwannoma is also known as an acoustic neuroma or acoustic neurinoma, a neoplasm for which some epidemiological studies have reported associations for increased risks in humans (Cardis et al. 2010; Hardell et al. 2002; Hardell et al. 2003; IARC 2013; Pettersson et al. 2014).

The evaluation of neoplastic tumors in the brain was rated moderate to high CoE for an increased incidence of glial cell-derived tumors at whole body average SAR at 6 W/kg over 106 weeks, based mainly on the NTP study (National Toxicology Program 2018b). In the study by Anderson et al. (2004), in a post publication analysis, an increase in oligodendrogliomas with a statistically significant trend across the dose range of 0.16 W/kg to 1.6 W/kg (whole body average SAR) given for 2 hours/day over 5 days/week over 100 weeks was observed when compared to historical controls (Anderson et al. 2004). Further, results reported by Falcioni (Falcioni et al. 2018) in another rat strain (male Fischer 344 rats), even though of borderline statistical significance support this evaluation. A moderate certainty for an increased incidence of neoplastic tumors was also found for the adrenal gland, namely pheochromocytoma at a whole-body average SAR level of 1.5 W/kg. Moderate CoE was concluded for an increased incidence of hepatoblastoma at a whole-body average SAR of 5 W/kg given over 106 weeks. We also concluded that the CoE for an increased incidence of lymphoma was moderate at a whole-body average SAR of 2.5 W/kg given for 108 weeks in mice. The evaluation of neoplastic tumors in the lung was rated moderate CoE for an increased positive trend for the incidence of bronchoalveolar adenoma and carcinoma (SAR: 2.5-10 W/kg) (National Toxicology Program 2018a) with some support of significant increases in one or both tumor types from two initiation-(co-)promotion studies (Lerchl et al. 2015; Tillmann et al. 2010).

For other cancer sites, this systematic review provides little evidence that RF EMF exposure causes cancer of the following systems, gastrointestinal/digestive, kidney, mammary gland, urinary, endocrine, musculoskeletal, reproductive, and auditory as shown in Supplementary Table 1.

2.2. Comparison with other studies

In 2011, the IARC Monograph Program performed an evaluation on RF EMF and a Working Group classified RF EMF as 'possibly carcinogenic to humans' (Class 2B) based upon limited evidence of carcinogenicity in humans and experimental animals (IARC 2013). Our analysis found additional support for evidence of carcinogenicity in animals based on studies published since the IARC review, in particular (Falcioni et al. 2018; National Toxicology Program 2018b).

In contrast to the recent systematic review by Pinto et al. (2023), who performed a meta-analysis for cancers with ≥3 studies we used a narrative approach (Pinto et al. 2023)because we considered the data too heterogeneous to support a meaningful quantitative integration (or subgroup analyses). Pinto et al. (2023) considered the evidence for carcinogenicity to be weaker than our judgements. This is perhaps especially true for

brain Pinto et al. (2023) also concluded that the limited number of studies for most cancer types limits the ability to consider results definitively conclusive.

2.2. Strengths and limitations of the review process

Most of the studies that support our evaluations with high or moderate CoE were published after the IARC Monographs meeting in 2011, and the evidence came primarily from two-year rodent bioassays. Overall, the eligible studies were performed within a timeframe of about 40 years. Overarching limitations to the evidence are mostly based on the heterogeneity in the study designs including different species and strains, study length, and different exposure conditions and levels. Therefore, consistency was difficult to evaluate. Due to the heterogeneity, per our protocol, a quantitative meta-analysis was not considered appropriate due to the small number of studies with sufficiently similar experimental design.

One challenge we faced was characterizing the magnitude of the effect size in a manner that could be considered analogous to presentation of relative risk estimates used to summarize findings from epidemiological studies (see Amendments). Ultimately, the direction of the effect was described, and the lowest SAR level at which an increase in cancer incidence occurred was noted. The 1% Bayesian Average BMDs were calculated when dose-response effects were identified. These BMDs may inform future health assessments that may be based on evidence from animal cancer bioassays, without making any assumptions regarding interspecies extrapolations with respect to incidence of carcinogenic findings.

While not limitations of the review *per se*, we did have some observations on our experience with the OHAT framework, which uses GRADE as it's foundation with some adjustments specifically to animal evidence (National Toxicology Program 2019a). With respect to use of the OHAT RoB tool, some of the RoB elements used for human clinical trial studies seemed contrived when applied to animal cancer bioassay studies. In particular, the assessment of whether research personnel allocating animals could not foresee which administered dose or exposure level is going to be assigned at the start of a study ('blinding before the study'). Raters generally considered not blinding at this step not likely to be a significant source of bias for animal studies. In addition, ideally RoB analyses include information on the predicted direction and magnitude of the potential bias.

Certain systematic review frameworks in environmental health introduce the additional assessment of study sensitivity or informativeness, which can identify study limitations that fall outside of RoB (IARC, RoC, IRIS). To assess study sensitivity for this systematic review, we adapted the tool from the RoC covering four domains, such as endpoint sensitivity of certain study design or exposure contrasts. We found the assessment of study sensitivity worthwhile as a complement to the RoB analysis, although we did not identify many chronic cancer bioassays studies with significant sensitivity concerns.

Like Hooijmans et al. (2018), our expedience indicates that additional method refinements and clarifications are needed when using GRADE-based approaches for reaching CoE conclusions from animal evidence (Hooijmans et al. 2018). Similar to Hooijmans et al. (2018), we separated considerations of directness into two levels. One level was consideration of directness with respect to the animal evidence (e.g., we considered chronic cancer bioassays to be more direct than inhiation/promotion studies). The other level was directness to the animal evidence to humans. Here we considered animal studies as relevant to the identification of a carcinogenic hazard to humans.

In our assessment, we experienced challenges given the diversity of the evidence base. This impacted judgments on whether to conduct a meta-analysis as well as judgements on

unexplained inconsistency and precision. Compared to experimental human evidence, a greater degree of heterogeneity in response may be expected animal evidence given the wide range of model systems and exposure paradigms used. For example, a high degree of consistency in findings from various inbred strains, outbred strains, and tumor-prone (and other modified) models would not necessarily be expected (Vesterinen et al. 2014). Thus, plausible reasons for inconsistency were present and we seldomly downgraded for unexplained inconsistency. We note that the OHAT refinement to GRADE allows for consistency across model systems to be used as an upgrade factor. Another issue we identified is the transparent consideration of supportive evidence into the structured framework, e.g., preneoplastic histopathology findings such as hyperplasia. The OHAT/GRADE upgrade factor of impact of residual confounding is typically contextualized as being applicable to epidemiological evidence. We did not utilize this factor in our analyses and are unaware of its use in other assessments of animal evidence using OHAT/GRADE. Thus, the applicability and how to operationalize this consideration to animal evidence should be considered as part of future method refinement. Worth noting is the existence a project group focused on analysis of animal evidence operating under the auspices of the GRADE Working Group (https://www.gradeworkinggroup.org, see “Groups and projects”). The goal of this project group is to advance and clarify methodology for applying GRADE to animal evidence.

The grading of the evidence as performed here in the context of the adapted GRADE and OHAT frameworks assessed the level of evidence by cancer site and did not develop an overall carcinogenicity conclusion. In contrast cancer hazard identifications as conducted by well-established national and international programs (e.g., U.S. RoC, IARC) synthesize the evidence and the resulting evaluation across all tumor sites. Typically, only a very few high-quality chronic cancer bioassays are available for any given agent, tested according to GLP guidelines at high levels of exposure to maximize the informativeness of the studies while avoiding toxicity and related threats to validity. Together, the evaluation framework applied here and the relatively low levels of exposure in most RF EMF cancer bioassays might result in a potential underestimation of a potential cancer hazard when compared with well-established programs.

2.2. Implications of the results for practice, policy, and future research

The results of this systematic review provide high or moderate CoE for several cancer sites relevant to cancer hazard identification for humans. In contrast to some uncertainties of translation of SAR/doses and cancer effects to human risk assessment, some of the SAR levels where effects were observed are within the magnitude of the whole body average SAR that might be experienced in humans. However, the type of exposure (whole body versus localized), intensity of exposure and duration of exposure must also be considered when translating the effect sizes to cancer risk in humans.

Strong evidence on cancer in experimental animals is relevant to the identification of a carcinogenic hazard to humans (Baan et al. 2019). Regarding the PECOS for human relevance, there is no a priori compelling mechanistic or other biological evidence to suggest animal models are not reasonable for predicting potential effects in humans. Further, in additional analyses of the tumor site concordance between experimental animals and humans, Krewski et al. compared agent-specific concordance between (groups of) cancer sites with sufficient evidence in humans and in animals (the latter required replication of positive results at the same specific site in at least two animal experiments). It has to be noted that these estimates are expected to underestimate concordance (Krewski et al. 2019). When present, tumor site concordance can increase confidence in the findings, particularly for human risk assessment. We note that the two tumor types with high CoE in animals in this systematic review are the same as those identified with limited evidence in humans by the IARC Working Group. Of note, “limited evidence of carcinogenicity in

humans" is defined in the Preamble to the IARC Monographs as "a causal interpretation of the positive association observed in the body of evidence on exposure to the agent and cancer is credible, but chance, bias, or confounding could not be ruled out with reasonable confidence."

For heart schwannomas, a significant increase was observed at a whole body average SAR of 6 W/kg (CDMA) and a significant SAR-related trend was observed in the NTP study. A statistically significant trend for an increase, and a significant increase at the highest SAR level was also observed in the Falcioni study where male Sprague Dawley rats were exposed to whole body average SAR of 0.001, 0.03 and 0.1 W/kg which is in the range of human exposures. Neoplasms derived from Schwann cells can basically occur at any location in the body where nerve sheath exist, and these neoplasms are the most common tumors of the peripheral nerve sheath.

Our systematic review is about carcinogenesis in experimental animals. As part of this work we identified the whole body average SAR levels where increases in cancer were observed, but no thresholds of carcinogenicity could be identified based on the data. Mechanistic understanding may help in the animal to human extrapolation of the exposure response relationship of RF EMF carcinogenicity and for setting limit values. One of the characteristics at higher SAR levels is an increase in body temperature, which has been suggested as a potential mechanism of carcinogenicity in animals (Hinchliffe et al. 2021; Zhang et al. 1995). While there is no scientific consensus about the relevance of this mechanism, and evidence for other mechanisms of the carcinogenicity of RF EMF emerges, this systematic review assessed whether RF EMF exposure-induced heating was measured in included studies as a domain of our specifically adapted RoB tool. All the informative studies on carcinogenicity have been evaluated of being at low RoB because they assessed tissue/body heating, but increases in core body temperature cannot be completely excluded. Conceptually, some hypothetical temperature effects at SAR levels above 1.5 W/kg and related body temperature increases of less than 1°C cannot be ruled out with certainty. In addition, oxidative stress has been demonstrated in various studies, even at low SAR levels, and permanent increase in oxidative stress may lead to cancer (Schuermann and Mevissen 2021). Specifically, hereto and within these limits, we propose some additional considerations on RF EMF-induced heating and how this compares between experimental animals and humans, primarily from the physical exposure perspective (Annex 2). This must, however, not be misunderstood as an endorsement of the temperature hypothesis of RF EMF carcinogenicity or concerns of potential temperature-related bias of our certainty assessments.

Given the high CoE for the carcinogenicity of RF-EMF-based on studies in experimental animals, and also taking into account reduction, replacement, refinement, and responsibility (4R) principle, with certain exceptions on frequencies of future networks, simple attempts at replication of current studies are not recommended, because they will never be full replicates (e.g., due to shifts in inbred rodent strains). Rather, we suggest to employ available biospecimens from informative chronic cancer bioassays to investigate mechanisms of RF-EMF carcinogenesis to inform inter-species extrapolation, also building on the systematic reviews on oxidative stress (Meyer et al. 2024; Schuermann and Mevissen 2021).

Frequencies that will be used in current and future 5G networks (3.5 – 60 GHz) are anticipated occur long before animal cancer data will be available to affect exposure guidelines. Such studies require many years to design, conduct and report which makes it a challenge to keep up with technology changes. If such studies are performed, they should include intermediate endpoints that may help to identify and disentangle mechanisms of the carcinogenicity of RF EMF. For instance, chronic bioassay studies could monitor serum stress hormones and ROS levels, heart rate and other metabolic endpoints that may help to

identify how organs or systems may be dysregulated from RF EMF exposure and which ones may be particularly susceptible to RF EMF.

The use of 3D models like organoids originating from humans induced pluripotent stem cells (iPSC) might also strengthen the mechanistic evidence. These emerging tools are focused on a system, tissue or even a tumor and also allow for mechanistic research. Regarding the CoE for this up-to-date systematic review, further research on brain tumors and related premalignant lesions would be of high of priority as well as research on schwannomas. In order to investigate SAR/dose response, different levels of SAR values, below and above realistic SAR levels of human exposure would be important for human risk assessment. Taking into account the long latency of solid tumors in humans, epidemiological studies on forthcoming wireless frequencies and evaluations will only be informative after the new technology has been in place for 10, 20 or more years. Despite all necessary efforts to refine and reduce studies with animals, chronic cancer bioassays are still the best understood tool for cancer hazard identification and risk assessment in scenarios where epidemiological studies in humans are likely not yet fully informative and mechanisms of carcinogenicity are not yet understood.

2.2.2. Final conclusions

The findings of this systematic review indicate that there is evidence that RF EMF exposure increases the incidence of cancer in experimental animals with the CoE being strongest for malignant heart schwannomas and gliomas.

Despite the high level of certainty that evidence of carcinogenicity in experimental animals may predict a carcinogenic hazard to humans, extrapolation of risk from cancer bioassays to humans is particularly complex for RF EMF. Without an understanding of the mechanism of the carcinogenicity of RF-EMF the choice of exposure metric for risk extrapolation (whole body versus localized), intensity or cumulative exposure whether or not a monotonic dose-response holds for carcinogenic effects, and whether SAR is the appropriate dose metric for adverse effects induced by RF-EMF may be critical.

Besides the integration of the sensitivity domain adopted from the RoC, more work is needed to tailor the GRADE approach to assessing the CoE from animal cancer bioassays designed to identifying risks of environmental agents.

2.2. Other information

The protocol has been registered at PROSPERO (No. 42021265563).

2.2. Support

2.2.2. Provide sources of support

This review was partly funded by the World Health Organization (WHO). WHO provided a first draft of the protocol and some methodological support on the implementation of the GRADE framework.

2.2.2. Disclaimer

The views expressed are those of the authors and do not necessarily represent the views or policies of the U.S. EPA. Any mention of trade names, products, or services does

not imply an endorsement by the U.S. government or the U.S. EPA. The U.S. EPA does not endorse any commercial products, services, or enterprises.

2.2.2. Availability of other material

All the materials are publicly available in HAWC EMF (Animal Toxicology) (2022) | HAWC (hawcproject.org).

References

Adey, W.R.; Byus, C.V.; Cain, C.D.; Higgins, R.J.; Jones, R.A.; Kean, C.J.; Kuster, N.; MacMurray, A.; Stagg, R.B.; Zimmerman, G. Spontaneous and nitrosourea-induced primary tumors of the central nervous system in Fischer 344 rats exposed to frequency-modulated microwave fields. Cancer Res 2000;60:1857-1863

Adey, W.R.; Byus, C.V.; Cain, C.D.; Higgins, R.J.; Jones, R.A.; Kean, C.J.; Kuster, N.; MacMurray, A.; Stagg, R.B.; Zimmerman, G.; Phillips, J.L.; Haggren, W. Spontaneous and nitrosourea-induced primary tumors of the central nervous system in Fischer 344 rats chronically exposed to 836 MHz modulated microwaves. Radiat Res 1999;152:293-302

Anane, R.; Dulou, P.E.; Taxile, M.; Geffard, M.; Crespeau, F.L.; Veyret, B. Effects of GSM-900 microwaves on DMBA-induced mammary gland tumors in female Sprague-Dawley rats. Radiat Res 2003;160:492-497

Anderson, L.E.; Sheen, D.M.; Wilson, B.W.; Grumbein, S.L.; Creim, J.A.; Sasser, L.B. Two-year chronic bioassay study of rats exposed to a 1.6 GHz radiofrequency signal. Radiat Res 2004;162:201-210

Baan, R.A.; Stewart, B.W.; Straif, K. Tumour site concordance and mechanisms of carcinogenesis. in: Baan R.A., Stewart B.W., Straif K., eds. IARC Scientific publication Lyon, France: IARC, France; 2019

Bannasch, P.; Griesemer, R.A.; Anders, F.; Becker, R.; Cabral, J.R.; Della Porta, G.; Feron, V.J.; Henschler, D.; Ito, N.; Kroes, R.; et al. Long-term assays for carcinogenicity in animals. IARC Sci Publ 1986:13-83

Bartsch, H.; Bartsch, C.; Seebald, E.; Deerberg, F.; Dietz, K.; Vollrath, L.; Mecke, D. Chronic exposure to a GSM-like signal (mobile phone) does not stimulate the development of DMBA-induced mammary tumors in rats: results of three consecutive studies. Radiat Res 2002;157:183-190

Bellossi, A.; Dubost, G.; Moulinoux, J.P.; Himdi, M.; Ruelloux, M.; Rocher, C. Biological effects of millimeter-wave irradiation on mice - Preliminary results. Ieee T Microw Theory 2000;48:2104-2110

Cardis, E.; Deltour, I.; Vrijheid, M.; Combalot, E.; Moissonnier, M.; Tardy, H.; Armstrong, B.; Giles, G.; Brown, J.; Siemiatycki, J.; Parent, M.E.; Nadon, L.; Krewski, D.; McBride, M.L.; Johansen, C.; Collatz, C.H.; Auvinen, A.; Kurttio, P.; Lahkola, A.; Salminen, T.; Hours, M.; Bernard, M.; Montestruq, L.; Schuz, J.; Berg-Beckhoff, G.; Schlehofer, B.; Blettner, M.; Sadetzki, S.; Chetrit, A.; Jarus-Hakak, A.; Lagorio, S.; Iavarone, I.; Takebayashi, T.; Yamaguchi, N.; Woodward, A.; Cook, A.; Pearce, N.; Tynes, T.; Blaasaas, K.G.; Klaeboe, L.; Feychting, M.; Lonn, S.; Ahlbom,

A.; McKinney, P.A.; Hepworth, S.J.; Muir, K.R.; Swerdlow, A.J.; Schoemaker, M.J.; Grp, I.S. Brain tumour risk in relation to mobile telephone use: results of the INTERPHONE international case-control study. Int J Epidemiol 2010;39:675-694

Chagnaud, J.L.; Moreau, J.M.; Veyret, B. No effect of short-term exposure to GSM-modulated low-power microwaves on benzo(a)pyrene-induced tumours in rat. Int J Radiat Biol 1999;75:1251-1256

Chou, C.K.; Guy, A.W.; Kunz, L.L.; Johnson, R.B.; Crowley, J.J.; Krupp, J.H. Long-term, low-level microwave irradiation of rats. Bioelectromagnetics 1992;13:469-496

de Seze, R.; Poutriquet, C.; Gamez, C.; Maillot-Marechal, E.; Robidel, F.; Lecomte, A.; Fonta, C. Repeated exposure to nanosecond high power pulsed microwaves increases cancer incidence in rat. PLoS One 2020;15:e0226858

Della Porta, G.; Dragani, T.A. Long-term assays for carcinogenicity. Teratog Carcinog Mutagen 1990;10:137-145

EPA. Guidelines for carcinogen risk assessment. Washington, DC: EPA; 2005

EPA, U.S. ORD Staff Handbook for Developing IRIS Assessments. 2022;

Falcioni, L.; Bua, L.; Tibaldi, E.; Lauriola, M.; De Angelis, L.; Gnudi, F.; Mandrioli, D.; Manservigi, M.; Manservisi, F.; Manzoli, I.; Menghetti, I.; Montella, R.; Panzacchi, S.; Sgargi, D.; Strollo, V.; Vornoli, A.; Belpoggi, F. Report of final results regarding brain and heart tumors in Sprague-Dawley rats exposed from prenatal life until natural death to mobile phone radiofrequency field representative of a 1.8GHz GSM base station environmental emission. Environ Res 2018;165:496-503

Frei, M.R.; Berger, R.E.; Dusch, S.J.; Guel, V.; Jauchem, J.R.; Merritt, J.H.; Stedham, M.A. Chronic exposure of cancer-prone mice to low-level 2450 MHz radiofrequency radiation. Bioelectromagnetics 1998a;19:20-31

Frei, M.R.; Jauchem, J.R.; Dusch, S.J.; Merritt, J.H.; Berger, R.E.; Stedham, M.A. Chronic, low-level (1.0 W/kg) exposure of mice prone to mammary cancer to 2450 MHz microwaves. Radiat Res 1998b;150:568-576

Goldstein, L.S.; Kheifets, L.; Van Deventer, E.; Repacholi, M. Further comments on "Long-term exposure of Emu-Pim1 transgenic mice to 898.4 MHz microwaves does not increase lymphoma incidence" by Utteridge et al. (Radiat. Res. 158, 357-364 2002). Radiat Res 2003;159:835; author reply 835-836

Hardell, L.; Hallquist, A.; Mild, K.H.; Carlberg, M.; Pahlson, A.; Lilja, A. Cellular and cordless telephones and the risk for brain tumours. Eur J Cancer Prev 2002;11:377-386

Hardell, L.; Mild, K.H.; Carlberg, M. Further aspects on cellular and cordless telephones and brain tumours. Int J Oncol 2003;22:399-407

Heikkinen, P.; Ernst, H.; Huuskonen, H.; Komulainen, H.; Kumlin, T.; Maki-Paakkanen, J.; Puranen, L.; Juutilainen, J. No effects of radiofrequency radiation on 3-chloro-4-(dichloromethyl)-5-hydroxy-2(5H)-furanone-induced tumorigenesis in female Wistar rats. Radiat Res 2006;166:397-408

Journal Pre-proofs

Heikkinen, P.; Kosma, V.M.; Alhonen, L.; Huuskonen, H.; Komulainen, H.; Kumlin, T.; Laitinen, J.T.; Lang, S.; Puranen, L.; Juutilainen, J. Effects of mobile phone radiation on UV-induced skin tumourigenesis in ornithine decarboxylase transgenic and non-transgenic mice. Int J Radiat Biol 2003;79:221-233

Heikkinen, P.; Kosma, V.M.; Hongisto, T.; Huuskonen, H.; Hyysalo, P.; Komulainen, H.; Kumlin, T.; Lahtinen, T.; Lang, S.; Puranen, L.; Juutilainen, J. Effects of mobile phone radiation on X-ray-induced tumorigenesis in mice. Radiat Res 2001;156:775-785

Higashikubo, R.; Culbreth, V.O.; Spitz, D.R.; LaRegina, M.C.; Pickard, W.F.; Straube, W.L.; Moros, E.G.; Roti, J.L.R. Radiofrequency electromagnetic fields have no effect on the in vivo proliferation of the 9L brain tumor. Radiation Research 1999;152:665-671

Hinchliffe, A.; Kogevinas, M.; Perez-Gomez, B.; Ardanaz, E.; Amiano, P.; Marcos-Delgado, A.; Castano-Vinyals, G.; Llorca, J.; Moreno, V.; Alguacil, J.; Fernandez-Tardon, G.; Salas, D.; Marcos-Gragera, R.; Aragones, N.; Guevara, M.; Gil, L.; Martin, V.; Benavente, Y.; Gomez-Acebo, I.; Santibanez, M.; Angel Alba, M.; Garcia, A.M.; Pollan, M.; Turner, M.C. Occupational Heat Exposure and Breast Cancer Risk in the MCC-Spain Study. Cancer Epidemiol Biomarkers Prev 2021;30:364-372

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.; Schunemann, H.J.; Langendam, M.W.; Group, G.W. Facilitating healthcare decisions by assessing the certainty in the evidence from preclinical animal studies. PLoS One 2018;13:e0187271

Hruby, R.; Neubauer, G.; Kuster, N.; Frauscher, M. Study on potential effects of "902-MHz GSM-type Wireless Communication Signals" on DMBA-induced mammary tumours in Sprague-Dawley rats. Mutation research 2008;649:34-44

Huang, T.Q.; Lee, J.S.; Kim, T.H.; Pack, J.K.; Jang, J.J.; Seo, J.S. Effect of radiofrequency radiation exposure on mouse skin tumorigenesis initiated by 7,12-dimethybenz[alpha]anthracene. Int J Radiat Biol 2005;81:861-867

IARC. Non-ionizing radiation, Part 2: Radiofrequency electromagnetic fields. IARC monographs on the evaluation of carcinogenic risks to humans / World Health Organization, International Agency for Research on Cancer 2013;102:1-460

IARC. Preamble to the IARC Monographs. IARC Monographs on the evaluation of carcinogenic risks to humans 2019;

Imaida, K.; Kuzutani, K.; Wang, J.; Fujiwara, O.; Ogiso, T.; Kato, K.; Shirai, T. Lack of promotion of 7,12-dimethylbenz[a]anthracene-initiated mouse skin carcinogenesis by 1.5 GHz electromagnetic near fields. Carcinogenesis 2001;22:1837-1841

Imaida, K.; Taki, M.; Watanabe, S.; Kamimura, Y.; Ito, T.; Yamaguchi, T.; Ito, N.; Shirai, T. The 1.5 GHz electromagnetic near-field used for cellular phones does not promote rat liver carcinogenesis in a medium-term liver bioassay. Jpn J Cancer Res 1998a;89:995-1002

Imaida, K.; Taki, M.; Yamaguchi, T.; Ito, T.; Watanabe, S.; Wake, K.; Aimoto, A.; Kamimura, Y.; Ito, N.; Shirai, T. Lack of promoting effects of the electromagnetic near-field used for cellular phones (929.2 MHz) on rat liver carcinogenesis in a medium-term liver bioassay. Carcinogenesis 1998b;19:311-314

Jauchem, J.R.; Ryan, K.L.; Frei, M.R.; Dusch, S.J.; Lehnert, H.M.; Kovatch, R.M. Repeated exposure of C3H/HeJ mice to ultra-wideband electromagnetic pulses: lack of effects on mammary tumors. Radiat Res 2001;155:369-377

Jin, Y.B.; Lee, H.J.; Seon Lee, J.; Pack, J.K.; Kim, N.; Lee, Y.S. One-year, simultaneous combined exposure of CDMA and WCDMA radiofrequency electromagnetic fields to rats. Int J Radiat Biol 2011;87:416-423

Kohnken, R.; Porcu, P.; Mishra, A. Overview of the Use of Murine Models in Leukemia and Lymphoma Research. Front Oncol 2017;7:22

Krewski, D.; Rice, J.M.; Bird, M.; Milton, B.; Collins, B.; Lajoie, P.; Billard, M.; Grosse, Y.; Cogliano, V.J.; Caldwell, J.C.; Rusyn, II; Portier, C.J.; Melnick, R.L.; Little, J.; Zielinski, J.M. Analysis of tumour site concordance. in: Baan R.A., Stewart B.W., Straif K., eds. Tumour Site Concordance and Mechanisms of Carcinogenesis. Lyon (FR); 2019

Kumar A., B.M., Laast V., Sharma A. Historical control background incidence of spontaneous neoplastic lesions of Sprague-Dawley rats in 104-week toxicity studies. Toxicol Pathol 2024;51:329-356

La Regina, M.; Moros, E.G.; Pickard, W.F.; Straube, W.L.; Baty, J.; Roti Roti, J.L. The effect of chronic exposure to 835.62 MHz FDMA or 847.74 MHz CDMA radiofrequency radiation on the incidence of spontaneous tumors in rats. Radiat Res 2003;160:143-151

Lee, H.J.; Jin, Y.B.; Lee, J.S.; Choi, S.Y.; Kim, T.H.; Pack, J.K.; Choi, H.D.; Kim, N.; Lee, Y.S. Lymphoma development of simultaneously combined exposure to two radiofrequency signals in AKR/J mice. Bioelectromagnetics 2011;32:485-492

Lerchl, A.; Klose, M.; Grote, K.; Wilhelm, A.F.; Spathmann, O.; Fiedler, T.; Streckert, J.; Hansen, V.; Clemens, M. Tumor promotion by exposure to radiofrequency electromagnetic fields below exposure limits for humans. Biochem Biophys Res Commun 2015;459:585-590

Louis, D.N.; Perry, A.; Wesseling, P.; Brat, D.J.; Cree, I.A.; Figarella-Branger, D.; Hawkins, C.; Ng, H.K.; Pfister, S.M.; Reifenberger, G.; Soffietti, R.; von Deimling, A.; Ellison, D.W. The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol 2021;23:1231-1251

Mason, P.A.; Walters, T.J.; DiGiovanni, J.; Beason, C.W.; Jauchem, J.R.; Dick, E.J., Jr.; Mahajan, K.; Dusch, S.J.; Shields, B.A.; Merritt, J.H.; Murphy, M.R.; Ryan, K.L. Lack of effect of 94 GHz radio frequency radiation exposure in an animal model of skin carcinogenesis. Carcinogenesis 2001;22:1701-1708

Mevissen, M.; Ward, J.M.; Kopp-Schneider, A.; McNamee, J.P.; Wood, A.W.; Rivero, T.M.; Thayer, K.; Straif, K. Effects of radiofrequency electromagnetic fields (RF EMF) on cancer in laboratory animal studies. Environ Int 2022;161:107106

Meyer, F.; Bitsch, A.; Forman, H.J.; Fragoulis, A.; Ghezzi, P.; Henschenmacher, B.; Kellner, R.; Kuhne, J.; Ludwig, T.; Sachno, D.; Schmid, G.; Tsaioun, K.; Verbeek, J.; Wright, R. The effects of radiofrequency electromagnetic field exposure on biomarkers of oxidative stress in vivo and in vitro: A systematic review of experimental studies. Environ Int 2024;194:108940

Morgan, R.L.; Thayer, K.A.; Bero, L.; Bruce, N.; Falck-Ytter, Y.; Ghersi, D.; Guyatt, G.; Hooijmans, C.; Langendam, M.; Mandrioli, D.; Mustafa, R.A.; Rehfuess, E.A.; Rooney, A.A.; Shea, B.;

Silbergeld, E.K.; Sutton, P.; Wolfe, M.S.; Woodruff, T.J.; Verbeek, J.H.; Holloway, A.C.; Santesso, N.; Schunemann, H.J. GRADE: Assessing the quality of evidence in environmental and occupational health. Environ Int 2016;92-93:611-616

National Toxicology Program, N. Handbook for Preparing Report on Carcinogens Monographs. Part E: Cancer Studies in Experimental Animals 2015a;

National Toxicology Program, N. OHAT Risk of Bias Rating Tool for Human and Animal Studies. Office of Health Assessment and Translation. 2015b;

National Toxicology Program, N. Toxicology and carcinogenesis studies in B6C3F1/N mice exposed to whole-body radio frequency radiation at a frequency (1,900 MHz) and modulations (GSM and CDMA) used by cell phones. Natl Toxicol Program Tech Rept Series 2018a:1-260

National Toxicology Program, N. Toxicology and carcinogenesis studies in Hsd: Sprague Dawley sd rats exposed to whole-body radio frequency radiation at a frequency (900 MHz) and modulations (GSM and CDMA) used by cell phones Natl Toxicol Program Tech Rept Series 2018b:1-384

National Toxicology Program, N. Handbook for Conducting a Literature-based Health Assessment Using OHAT Approach for Systematic Review and Evidence Integration. 2019a;

National Toxicology Program, N. OHAT Risk of Bias Rating Tool for Human and Animal Studies. in: NIEHS/NTP, ed; 2019b

Oberto, G.; Rolfo, K.; Yu, P.; Carbonatto, M.; Peano, S.; Kuster, N.; Ebert, S.; Tofani, S. Carcinogenicity study of 217 Hz pulsed 900 MHz electromagnetic fields in Pim1 transgenic mice. Radiat Res 2007;168:316-326

OECD. Detailed review paper on transgenic rodent mutation assays. 2009;103

OECD. Guidance document 116 on the conduct and design of chronic toxicity and carcinogenicity studies, supporting test guidelines 451 and 453, 2nd edition 2012;116

OECD. Guidance document 116 on the conduct and design of chronic toxicity and carcinogenicity studies, supporting test guidelines 451, 452 and 453. 2014;

OECD. OECD guidelines for the testing of chemicals: carcinogenicity studies. 2018;

Ouadah, N.S.; Lecomte, A.; Robidel, F.; Olsson, A.; Deltour, I.; Schuz, J.; Blazy, K.; Villegier, A.S. Possible effects of radiofrequency electromagnetic fields on in vivo C6 brain tumors in Wistar rats. J Neurooncol 2018;140:539-546

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.; Hrobjartsson, 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. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71

Paulraj, R.; Behari, J. Effects of low level microwave radiation on carcinogenesis in Swiss Albino mice. Molecular and cellular biochemistry 2011;348:191-197

Journal Pre-proofs

Pettersson, D.; Mathiesen, T.; Prochazka, M.; Bergenheim, T.; Florentzson, R.; Harder, H.; Nyberg, G.; Siesjo, P.; Feychting, M. Long-term mobile phone use and acoustic neuroma risk. Epidemiology 2014;25:233-241

Pinto, R.; Ardoino, L.; Villani, P.; Marino, C. In Vivo Studies on Radiofrequency (100 kHz–300 GHz)|Electromagnetic Field Exposure and Cancer: A Systematic Review. International Journal of Environmental Research and Public Health 2023;20

Preskorn, S.H.; Edwards, W.D.; Justesen, D.R. Retarded tumor growth and greater longevity in mice after fetal irradiation by 2450-MHz microwaves. J Surg Oncol 1978;10:483-492

Repacholi, M.H.; Basten, A.; Gebski, V.; Noonan, D.; Finnie, J.; Harris, A.W. Lymphomas in E mu-Pim1 transgenic mice exposed to pulsed 900 MHZ electromagnetic fields. Radiat Res 1997;147:631-640

Safety, I.C.o.E. C951-2019_ IEEE Standards for safety levels with respect t human exposure to electric_ magnetic and electromagnetic fiels_ 0 Hz to 300GHz.pdf>. 2019;

Salford, L.G.; Brun, A.; Persson, B.R.R. Brain tumour development in rats exposed to electromagnetic fields used in wireless cellular communication. Wirel Netw 1997;3:463-469

Salford, L.G.; Brun, A.; Persson, B.R.R.; Eberhardt, J. Experimental Studies of Brain-Tumor Development during Exposure to Continuous and Pulsed 915 Mhz Radiofrequency Radiation. Bioelectroch Bioener 1993;30:313-318

Saran, A.; Pazzaglia, S.; Mancuso, M.; Rebessi, S.; Di Majo, V.; Tanori, M.; Lovisolo, G.A.; Pinto, R.; Marino, C. Effects of exposure of newborn patched1 heterozygous mice to GSM, 900 MHz. Radiat Res 2007;168:733-740

Schuermann, D.; Mevissen, M. Manmade Electromagnetic Fields and Oxidative Stress-Biological Effects and Consequences for Health. Int J Mol Sci 2021;22

Shirai, T.; Ichihara, T.; Wake, K.; Watanabe, S.; Yamanaka, Y.; Kawabe, M.; Taki, M.; Fujiwara, O.; Wang, J.; Takahashi, S.; Tamano, S. Lack of promoting effects of chronic exposure to 1.95-GHz W-CDMA signals for IMT-2000 cellular system on development of N-ethylnitrosourea-induced central nervous system tumors in F344 rats. Bioelectromagnetics 2007;28:562-572

Shirai, T.; Kawabe, M.; Ichihara, T.; Fujiwara, O.; Taki, M.; Watanabe, S.; Wake, K.; Yamanaka, Y.; Imaida, K.; Asamoto, M.; Tamano, S. Chronic exposure to a 1.439 GHz electromagnetic field used for cellular phones does not promote N-ethylnitrosourea induced central nervous system tumors in F344 rats. Bioelectromagnetics 2005;26:59-68

Smith, P.; Kuster, N.; Ebert, S.; Chevalier, H.J. GSM and DCS wireless communication signals: combined chronic toxicity/carcinogenicity study in the Wistar rat. Radiat Res 2007;168:480-492

Sommer, A.M.; Bitz, A.K.; Streckert, J.; Hansen, V.W.; Lerchl, A. Lymphoma development in mice chronically exposed to UMTS-modulated radiofrequency electromagnetic fields. Radiat Res 2007;168:72-80

Sommer, A.M.; Streckert, J.; Bitz, A.K.; Hansen, V.W.; Lerchl, A. No effects of GSM-modulated 900 MHz electromagnetic fields on survival rate and spontaneous development of lymphoma in female AKR/J mice. BMC cancer 2004;4:77

Journal Pre-proofs

Szmigielski, S.; Szudzinski, A.; Pietraszek, A.; Bielec, M.; Janiak, M.; Wrembel, J.K. Accelerated development of spontaneous and benzopyrene-induced skin cancer in mice exposed to 2450-MHz microwave radiation. Bioelectromagnetics 1982;3:179-191

Szudzinski, A.; Pietraszek, A.; Janiak, M.; Wrembel, J.; Kalczak, M.; Szmigielski, S. Acceleration of the development of benzopyrene-induced skin cancer in mice by microwave radiation. Arch Dermatol Res 1982;274:303-312

Tillmann, T.; Ernst, H.; Ebert, S.; Kuster, N.; Behnke, W.; Rittinghausen, S.; Dasenbrock, C. Carcinogenicity study of GSM and DCS wireless communication signals in B6C3F1 mice. Bioelectromagnetics 2007;28:173-187

Tillmann, T.; Ernst, H.; Streckert, J.; Zhou, Y.; Taugner, F.; Hansen, V.; Dasenbrock, C. Indication of cocarcinogenic potential of chronic UMTS-modulated radiofrequency exposure in an ethylnitrosourea mouse model. Int J Radiat Biol 2010;86:529-541

Toler, J.C.; Shelton, W.W.; Frei, M.R.; Merritt, J.H.; Stedham, M.A. Long-term, low-level exposure of mice prone to mammary tumors to 435 MHz radiofrequency radiation. Radiat Res 1997;148:227-234

Utteridge, T.D.; Gebski, V.; Finnie, J.W.; Vernon-Roberts, B.; Kuchel, T.R. Long-term exposure of E-mu-Pim1 transgenic mice to 898.4 MHz microwaves does not increase lymphoma incidence. Radiat Res 2002;158:357-364

Verbeek, J.; Oftedal, G.; Feychting, M.; van Rongen, E.; Rosaria Scarfi, M.; Mann, S.; Wong, R.; van Deventer, E. Prioritizing health outcomes when assessing the effects of exposure to radiofrequency electromagnetic fields: A survey among experts. Environ Int 2021;146:106300

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. Meta-analysis of data from animal studies: a practical guide. J Neurosci Methods 2014;221:92-102

Wu, R.Y.; Chiang, H.; Shao, B.J.; Li, N.G.; Fu, Y.D. Effects of 2.45-GHz microwave radiation and phorbol ester 12-O-tetradecanoylphorbol-13-acetate on dimethylhydrazine-induced colon cancer in mice. Bioelectromagnetics 1994;15:531-538

Yu, D.; Shen, Y.; Kuster, N.; Fu, Y.; Chiang, H. Effects of 900 MHz GSM wireless communication signals on DMBA-induced mammary tumors in rats. Radiat Res 2006;165:174-180

Zhang, Z.F.; Vena, J.E.; Zielezny, M.; Graham, S.; Haughey, B.P.; Brasure, J.; Marshall, J.R. Occupational exposure to extreme temperature and risk of testicular cancer. Arch Environ Health 1995;50:13-18

Ziegelberger, G.C., Rodney J.; Feychting, Maria; Green, Adèle C.; Hirata, Akimasa; d'Inzeo, Guglielmo; Marino, Carmela; Miller, Sharon A.; Oftedal, Gunnhild; Okuno, Tsutomu; van Rongen, Eric; Röösli, Martin; Sienkiewicz, Zenon; Watanabe, Soichi. . ICNIRP Note: Critical Evaluation of Two Radiofrequency Electromagnetic Field Animal Carcinogenicity Studies Published in 2018. Health Phys 2020;118:525-532

Zook, B.C.; Simmens, S.J. The effects of 860 MHz radiofrequency radiation on the induction or promotion of brain tumors and other neoplasms in rats. Radiat Res 2001;155:572-583

Journal Pre-proofs

Zook, B.C.; Simmens, S.J. The effects of pulsed 860 MHz radiofrequency radiation on the promotion of neurogenic tumors in rats. Radiat Res 2006;165:608-615

Figure 1. From Initial Confidence by Key Features of Study Design to Confidence in the Body of Evidence considering Factor Decreasing and Increasing Confidence (from OHAT, NTP 2019)(National Toxicology Program 2019b)

Figure 2. PRISMA Study Flow Diagram

Figure 3A. Risk of Bias Analysis. Risk of bias results for included studies **with** chronic bioassays **(see interactive graphic for rating rationales)**

Figure 3B. Risk of Bias Analysis. Risk of bias results for included studies **with** initiation-(co) promotion studies (**see interactive graphic for rating rationales)**

Figure 3C. Risk of Bias Analysis. Risk of bias results for included studies **with** tumor-prone mice **(see interactive graphic for rating rationales)**

Figure 4A. Sensitivity Analysis. Sensitivity results for included studies **with** chronic bioassays **(see interactive graphic for rating rationales)**

Figure 4B. Sensitivity Analysis. Sensitivity results for included studies **with** initiation-(co-)promotion studies **(see interactive graphic for rating rationales)**

Figure 4C. Sensitivity Analysis. Sensitivity results for included studies with tumor prone mice (see interactive graphic for rating rationales)

Figure 5. Overview of lymphomas reported in included studies

Figure 6. Risk of bias results for studies assessing lymphoma-related carcinogenicity in chronic bioassays (see interactive graphic for rating rationales)

Figure 7. Lymphoma-related malignant neoplasms from chronic bioassay studies – Exposure Array. The dashed lines separate species and sex. For additional details see interactive graphic in HAWC. (Anderson et al. 2004; Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018a; Tillmann et al. 2007)

Figure 8. Summary of malignant neoplasms from chronic bioassays - Effect Size. The dashed lines separate species and sex. This is a thumbnail image of an interactive graphic in HAWC. (Anderson et al. 2004; Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018a; b; Tillmann et al. 2007)

Figure 9. Overview of brain carcinogenicity endpoints reported in included studies This is a thumbnail image of an interactive dashboard available on HAWC that is filterable by citation, health system, species, strain, and frequency. **Panel A** Numbers in the heatmap represent the number of studies that investigated a health system within a study design. If a study evaluated multiple health outcomes or presented several experiments, it is shown here multiple times. **Panel B** provides additional experimental detail about the studies, including exposure characteristics. The table in the Panel B is only a partial representation of the full table available in HAWC.

Figure 10. Risk of bias results for studies assessing brain-related carcinogenicity of chronic bioassay studies (see interactive graphic for rating rationales)

Figure 11A. Brain-related carcinogenicity of malignant neoplasms derived from glial cells - Exposure Array – of chronic bioassays. The dashed lines separate species and sex. For additional details see interactive graphic in HAWC (Anderson et al. 2004; Falcioni et al. 2018; La Regina et al. 2003; National Toxicology Program 2018a; b). Data not shown for studies where data were only presented in figures or described qualitatively.

Figure 11B. Brain-related carcinogenicity of malignant neoplasms derived from glial cells - Effect Size – of chronic bioassays. The dashed lines separate species and sex. This is a thumbnail image of an interactive graphic in HAWC . Data not shown for studies where data were only presented in figures or described qualitatively.

Figure 12A. Brain-related carcinogenicity of malignant neoplasms other than glia cell neoplasms – Exposure array – of chronic bioassays. The dashed lines separate species and sex. For additional details see interactive graphic in HAWC (Anderson et al. 2004; Falcioni et al. 2018; La Regina et al. 2003; National Toxicology Program 2018b). Data not shown for studies where data were only presented in figures or described qualitatively.

Figure 12B. Brain-related carcinogenicity of malignant neoplasms other than those derived from glial cells – Effect Size – of chronic bioassays. The dashed lines separate species and sex. This is a thumbnail image of an interactive graphic in HAWC . We added the NTP study with mice even though no brain tumors were listed in their tables of mouse tumors in the technical report (National Toxicology Program 2018a). Data not shown for studies where data were only presented in figures or described qualitatively.

Journal Pre-proofs

Figure 13. Overview of heart carcinogenicity endpoints reported in included studies

Figure 14. Risk of bias results for chronic bioassay studies assessing heart-related carcinogenicity (see Interactive graphic)

Figure 15A. Heart-related carcinogenicity of malignant neoplasms – Exposure Array. For additional details see interactive graphic in HAWC. The dashed lines separate species and sex (Falcioni et al. 2018; National Toxicology Program 2018a; b); the blue dashed line in the NTP 2018 rat study at 900 MHz CDMA displays a statistically significant trend.

Figure 15B. Heart-related carcinogenicity of malignant neoplasms – Effect Size. The dashed lines separate species and sex. This is a thumbnail image of an interactive graphic in HAWC. (Falcioni et al. 2018; National Toxicology Program 2018a; b)

Figure 16. ***Overview of adrenal gland carcinogenicity endpoints reported in included studies***

Figure 17. **Risk of bias assessment for chronic bioassays on adrenal gland-related carcinogenicity** (see interactive graphic for rating rationales)

Figure 18A. presents an exposure response array of malignant or combined (benign, malignant) pheochromocytomas of the adrenal gland. The dashed lines separate species and sex. For additional details see interactive graphic in HAWC. (Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018a; b; Tillmann et al. 2007)

Figure 18B. presents the effect size (tumor incidence) of malignant or combined (benign, malignant) pheochromocytomas of the adrenal gland for chronic bioassay studies. The dashed lines separate species and sex. This is a thumbnail image of an interactive graphic in HAWC.

Figure 19. Overview of liver carcinogenicity endpoints reported in included studies.This is a thumbnail image of an interactive dashboard available on HAWC that is filterable by citation, health system, species, strain, and frequency. **Panel A** Numbers in the heatmap represent the number of studies that investigated a health system within a study design. If a study evaluated multiple health outcomes or presented several experiments, it is shown here multiple times. **Panel B** provides additional experimental detail about the studies, including exposure characteristics. The table in the Panel B is only a partial representation of the full table available in HAWC. Studies where data are presented on figures only, and data could not be extracted are not shown in the figure.

Figure 20. Risk of bias results for chronic bioassay studies assessing liver-related carcinogenicity (see interactive graphic for rating rationales).

Figure 21. presents an exposure response array of malignant neoplasms of the liver across studies (chronic bioassays) to summarize exposure-specific effects interactive graphic. The dashed lines

separate species, strains and sex. (Chou et al. 1992; National Toxicology Program 2018a; b; Tillmann et al. 2007)

Figure 22. presents a thumbnail image of the effect size (tumor incidence) of malignant neoplasms of the liver across studies (chronic bioassays). The dashed lines separate species and sex. This is a thumbnail image of an interactive graphic in HAWC. The dashed lines separate species, strains and sex.

Figure 23. is a thumbnail image of an interactive graphic in HAWC. The graphic presents the included studies and endpoints assessed across studies (Panel A) and summarizes study characteristics for the included studies (Panel B). Additional details on the studies can be accessed through the interactive dashboard as well as download of the underlying data in tabular format using the 'Actions' option.

Figure 24. Risk of bias results for chronic bioassay studies assessing lung-related carcinogenicity (see interactive graphic for rating rationales).

Figure 25. presents the exposure response array of malignant neoplasms of the lung across studies (chronic bioassays) to summarize exposure-specific effects interactive graphic. The dashed lines separate species, strains and sex (National Toxicology Program 2018a; Tillmann et al. 2007).

Figure 26. presents a thumbnail image of the effect size (tumor incidence) of malignant neoplasms of the lung across studies (chronic bioassays). The dashed lines separate species and sex.This is a thumbnail image of an interactive graphic in HAWC. The dashed lines separate species, strains and sex (National Toxicology Program 2018a; Tillmann et al. 2007).

Table 1. PECOS Eligibility Criteria

PECOS element	Description
Populations	*Inclusion criteria:* Studies conducted with non-human mammalian animal species (whole organism), of any life-stage (including preconception, *in utero*, lactation, peripubertal, and adult stages), of any strain, sub-strain and sex including transgenic animals (tumor-prone animals).
Exposures	*Inclusion criteria:* Studies that have applied electric, magnetic or electromagnetic fields in the frequency range of 100 kHz to 300 GHz and that 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. • **body/tissue/sample internal exposure metrics** measured or calculated for the particular conditions of the experiment (i.e., specific absorption rate or SAR induced electric field strength, internal magnetic field strength). • **body/tissue/sample internal exposure metrics** describing superficial absorption at frequencies above 6 GHz measured or calculated for the particular conditions of the experiment (incident power flux density,

PECOS element	Description

	incident energy density, transmitted (absorbed) power flux density, transmitted (absorbed) energy.
	• **body/tissue/sample external exposure metrics** [external electric field strength (E>1 V/m or E>√10*background level in unshielded environment, otherwise no restriction), external magnetic field strength (H>2.7 mA/m or H>√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)].
	
	These exposure metrics had to be measured or calculated at the location of the exposed body in the approximate far-field of the field source. 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.
	
	• mobile phones or other RF-generating devices as source of exposure (without reporting of metrics given above and
	  • With output controlled by appropriate software or hardware operated close to the tissue/sample, provided that the output power and the distance to the sample are reported to enable inference of the exposure).
	  • In Global System for Mobile communications (GSM) mode with an active call operated close to the body/tissue/sample

Journal Pre-proofs

PECOS element	Description
	Studies involving other exposures, e.g., in animal models where a tumor initiator and/or tumor promotor was used, will be included only if they include an experimental arm with exposure to RF EMF only. Studies involving co-exposure to a tumor initiator, tumor promotor or co-promotor need to provide the dose and timing of the administration, as well as the source of the agent. *Exclusion criteria:* Studies with exposure of laboratory animals outside the considered frequency range, 100 kHz – 300 GHz, as well as studies without data on exposure metrics. In case a mobile phone was used, and it was not operated in GSM or code division multiple access (CDMA) mode and the outcome power was not controlled by hard- or software, or no active call was established and maintained through the experiment. Studies with an exposure duration <10 days for long-term animal cancer bioassays including initiation-(co-)promotion studies or less than 1 day in transgenic rodent assay with at least one dose-level tested.
Comparators	*Inclusion criteria:* Studies that have compared exposure to a concurrent control, namely a sham-exposed group or a group that has been exposed to a substantially lower level of RF EMF or a non-exposed control group (cage control with the laboratory environment). Studies with historic controls can also be included, especially for rare tumors, under certain conditions (i.e., same sex, same species, same strain, same gender, same diet, same laboratory environment, species that originates from the same laboratory within a time frame of not more than 5 years). *Exclusion criteria: Studies* without sham-exposed animals or other controls (e.g., cage) with the RF EMF exposure in the controls being substantially lower than the RF EMF exposure.
Outcomes	*Inclusion criteria:* Studies that have evaluated the incidence of one or more of the following cancer-related endpoints in any organ, tissue or body fluids of the laboratory animals: malignant tumors, benign tumors, or combinations of benign and malignant tumors, and preneoplastic lesions. Studies must report effect estimate(s) or sufficient data to calculate an effect estimate.
Study Type	*Inclusion criteria:* Long-term carcinogenicity studies, initiation-(co-)promotion studies, co-carcinogenesis studies, studies in tumor-prone animals, and implantation (cancer cells) studies by all routes of exposure. Medium-term duration tests for carcinogenicity on the development of proliferative lesions in a single tissue, e.g., foci of alteration in the liver, are also considered.

Journal Pre-proofs

PECOS element	Description
	Exclusion criteria: Therapeutic studies, e.g., electroporation, microthermal applications, magnetic resonance tomography were considered out of scope and tracked as ineligible.

For further details of exposure inclusion/exclusion criteria see Mevissen et al. 2022 and for exposure guidelines for human populations, including metrics, see ICNIRP 2020 and IEEE 2019 (Safety 2019; Ziegelberger 2020).

Table 2. Comparison of chronic cancer bioassays reporting lymphoma-related carcinogenicity outcomes

Author/ Year	Species (Group No.)	Strain	Sex	No animals per group	In utero exposure	Frequency (MHz)	Modulation	SAR	Exposure whole body (wb)/local	Duration (weeks)	Duration (h/d, d/w)	Type of Tumor (b/m)
Anderson, 2004	Rat	F344 F1	M + F	90	N/A	1620		0.16, 1.6	Local (head)	104	2h/d, 5d/w	b+m
Chou, 1992	Rat	SD	M	100	N/A	2450	MW	0.4	wb	100	21.5h/d, 7d/w	b+m
La Regina, 2003	Rat	F344	M + F	40	N/A	835.62	FDMA	1.3	wb	104	4h/d, 5d/w,	b+m
	Rat	F344	M + F	40	N/A	847.74	CDMA	1.3	wb	104	4h/d, 5d/w,	b+m
NTP, 2018	Rat	SD	M + F	90	GD6	900	GSM	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10'OFF	b+m
	Rat	SD	M + F	90	GD6	900	CDMA	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10'OFF	b+m
	Mouse	B6C3F1/N	M + F	90	N/A	1900	GSM	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10'OFF	b+m
	Mouse	B6C3F1/N	M + F	90	N/A	1900	CDMA	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10'OFF	b+m

Tillman, 2007	Mouse	B6C3F1/N	M + F	50	N/A	1747	DCS	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m
	Mouse	B6C3F1/N	M + F	50	N/A	902	GSM	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m

Table 3. Certainty Assessment for Lymphoma-Related Carcinogenicity

		Factors Decreasing Certainty "—" if no concerns, "↓"if serious concerns to downgrade certainty					**Factors Increasing Certainty** "—" not detected or applicable, "↑ "if sufficient to upgrade certainty					
	INITIAL CERTAINTY RATING (# of studies)	**Risk of Bias**	**Unexplained Inconsistency**	**Indirectness**	**Imprecision**	**Publication Bias**	**Large Magnitude**	**Dose Response**	**Residual Confounding**	**Consistency Across Species/Models**	**Rare Outcomes**	**FINAL CERTAINTY RATING**
Malignant Lymphoma (all systems)	High (4 rat, 12 mouse)	--	↓	--	--	--	--	--	--	--	--	Moderate for an increased incidence of lymphoma. The calculations of the

												1% BMD were not statistically significant for any of the studies.

Citations: *Mouse* (Frei et al. 1998a; Heikkinen et al. 2006; Heikkinen et al. 2001; Lee et al. 2011; National Toxicology Program 2018b; Oberto et al. 2007; Repacholi et al. 1997; Sommer et al. 2004; Tillmann et al. 2007; Tillmann et al. 2010; Utteridge et al. 2002 Lerchl et al., 2015). *Rat* (Anderson et al. 2004; Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018b).

Table 4. Comparison of chronic cancer bioassays reporting brain-related carcinogenicity outcomes

Author/ Year	Species (Group No.)	Strain	Sex	No animals per group	In utero exposure	Frequency (MHz)	Modulation	SAR	Exposure whole body (wb)/local	Duration (weeks)	Duration (h/d, d/w)	Type of Tumor (b/m)
Anderson, 2004	Rat	F344 F1	M+F	90	N/A	1620		0.16, 1.6	Local (head)	104	2h/d, 5d/w	b+m
Falcioni, 2018	Rat	SD	M+F	409-811	GD12	1800	GSM	0.001, 0.03, 0.1	wb	GD12-152	19h/d; 7d/w	b+m
La Regina, 2003	Rat	F344	M+F	40	N/A	835.62	FDMA	1.3	wb	104	4h/d, 5d/w,	b+m
	Rat	F344	M+F	40	N/A	847.74	CDMA	1.3	wb	104	4h/d, 5d/w	b+m

NTP, 2018	Rat	SD	M +F	90	GD6	900	GSM	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
	Rat	SD	M +F	90	GD6	900	CDMA	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
	Mouse	B6C3F 1/N	M +F	90	N/A	1900	GSM	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10 'OFF	b+m
	Mouse	B6C3F 1/N	M +F	90	N/A	1900	CDMA	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10 'OFF	b+m

Table 5. Certainty Assessment for Brain-Related Carcinogenicity

		Factors Decreasing Certainty "—" if no concerns, "↓"if serious concerns to downgrade certainty					**Factors Increasing Certainty** "—" not detected or applicable, "↑ "if sufficient to upgrade certainty					
	INITIAL CERTAINTY RATING (# of studies)	**Risk of Bias**	**Unexplained Inconsistency**	**Indirectness**	**Imprecision**	**Publication Bias**	**Large Magnitude**	**Dose Response**	**Residual Confounding**	**Consistency Across Spes/Models**	**Rare Outcomes**	**FINAL CERTAINTY RATING**

Journal Pre-proofs

Brain	High (10 rat, 6 mouse)	--	--	--	--	--	--	(+)	--	--	--	High for an increased incidence of gliomas at 6 W/kg (whole body average SAR) for 9 hours and 10 minutes/day over 106 weeks (National Toxicology Program 2018b). The 1% BMD is 4.244 (95% CI 2.70, 10.24). Increased incidence of oligodendroglioma s with a statistically significant trend

Journal Pre-proofs

												observed across the dose range of 0.16 W/kg to 1.6 W/kg (whole body average SAR) given for 2 hours/ day over 5 days/ week over 100 weeks (Anderson et al. 2004).

Citations **:** Mouse (Heikkinen et al. 2001; Lerchl et al. 2015; National Toxicology Program 2018a; Saran et al. 2007; Sommer et al. 2004; Tillmann et al. 2010) and Rat (Adey et al. 2000; Adey et al. 1999; Anderson et al. 2004; Falcioni et al. 2018; Heikkinen et al. 2006; National Toxicology Program 2018b; Shirai et al. 2007; Shirai et al. 2005; Zook and Simmens 2001; 2006).

Table 6. Comparison of chronic cancer bioassays reporting heart-related carcinogenicity outcomes

Author/ Year	Species (Group No.)	Strain	Sex	No animals per group	In utero exposure	Frequency (MHz)	Modulation	SAR	Exposure whole body (wb)/local	Duration (weeks)	Duration (h/d, d/w)	Type of Tumor (b/m)
Falcioni, 2018	Rat	SD	M +F	409-811	GD12	1800	DCS	0.001, 0.03	wb	GD12-152	19h/d; 7d/w	b+m
NTP, 2018	Rat	SD	M +F	90	GD6	900	GSM	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w, 10'ON/10' OFF	b+m

Rat	SD	M +F	90	GD6	900	CDMA	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w, 10'ON/10' OFF	b+m
Mouse	B6C3F 1/N	M +F	90	N/A	1900	DCS	2.5, 5, 10	wb	107	9h10''/d; 7d/w, 10'ON/10' OFF	b+m
Mouse	B6C3F 1/N	M +F	90	N/A	1900	CDMA	2.5, 5, 10	wb	107	9h10''/d; 7d/w, 10'ON/10' OFF	b+m

Table 7. Certainty Assessment for Heart-Related Carcinogenicity

		Factors Decreasing Certainty "—" if no concerns, "↓"if serious concerns to downgrade certainty					**Factors Increasing Certainty** "—" not detected or applicable, "↑ "if sufficient to upgrade certainty					
	INITIAL CERTAINTY RATING (# of studies)	**Risk of Bias**	**Unexplained Inconsistency**	**Indirectness**	**Imprecision**	**Publication Bias**	**Large Magnitude**	**Dose Response**	**Residual Confounding**	**Consistency Across Species/Models**	**Rare Outcomes**	**FINAL CERTAINTY RATING**
Heart	High (3 rat, 1 mouse)*	--	--	--	--	--	--	+	--	--	+	High for an increased incidence of malignant schwannomas was observed with a statistically

												significant trend across the dose range of 1.5 W/kg to 6 W/kg whole body average SAR for 9 hours and 10 minutes/day for 7 days/week over 106 weeks (NTP 2018a) with a 1% Bayesian Average BMD of 1.918 (95% CI 0.71, 4.145). An increased incidence of total

Journal Pre-proofs

												schwannomas was also observed with statistically significance at the highest dose 0.1 W/kg (Falcioni et al. 2018) The 1% BMD is 0.177 (95% CI 0.125, 0.241).

Citations**:** mouse (National Toxicology Program 2018a) and rat (Falcioni et al. 2018; Heikkinen et al. 2006; National Toxicology Program 2018b).
*No factors for increasing certainty were applied because the initial certainty was high and there were no downgrades, but we note that the 'dose-response' and the 'rare outcomes' factors would apply as a factors increasing the certainty.

Table 8. Comparison of chronic cancer bioassays reporting adrenal gland-related carcinogenicity outcomes

Author/ Year	Species (Group No.)	Strain	Sex	No animals per group	In utero exposure	Frequency (MHz)	Modulation	SAR	Exposure whole body (wb)/local	Duration (weeks)	Duration (h/d, d/w)	Type of Tumor (b/m)
Chou, 1992	Rat	SD	M	100	N/A	2450	MW	0.4	wb	100	21.5h/d, 7d/w	b+m

La Regina, 2003	Rat	F344	M + F	40	N/A	835.62	FDMA	1.3	wb	104	4h/d, 5d/w,	b+m
	Rat	F344	M + F	40	N/A	847.74	CDMA	1.3	wb	104	4h/d, 5d/w,	b+m
NTP, 2018	Rat	SD	M + F	90	GD6	900	GSM	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w, 10'ON/10' OFF	b+m
	Rat	SD	M + F	90	GD6	900	CDMA	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w, 10'ON/10' OFF	b+m
	Mouse	B6C3F1/N	M + F	90	N/A	1900	GSM	2.5, 5, 10	wb	107	9h10''/d; 7d/w, 10'ON/10 'OFF	b+m
	Mouse	B6C3F1/N	M + F	90	N/A	1900	CDMA	2.5, 5, 10	wb	107	9h10''/d; 7d/w, 10'ON/10 'OFF	b+m
Tillman, 2007	Mouse	B6C3F1/N	M + F	50		1747	DCS	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m
	Mouse	B6C3F1/N	M + F	50		902	GSM	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m

Table 9. Certainty Assessment for Adrenal Gland-Related Carcinogenicity

		Factors Decreasing Certainty "—" if no concerns, "↓"if serious concerns to downgrade certainty	**Factors Increasing Certainty** "—" not detected or applicable, "↑ "if sufficient to upgrade certainty	

Journal Pre-proofs

	INITIAL CERTAINTY RATING (# of studies)	Risk of Bias	Unexplained Inconsistency	Indirectness	Imprecision	Publication Bias	Large Magnitude	Dose Response	Residual Confounding	Consistency Across Species/ Models	Rare Outcomes	FINAL CERTAINTY RATING
Adrenal Gland	High (5 rat, 7 mouse)	--	↓	--	--	--	--	--	--	--	--	Moderate for an increased incidence of malignant neoplasms (pheochromocytomas).

Citations: *Mouse* (Frei et al. 1998b; Heikkinen et al. 2001; Jauchem et al. 2001; National Toxicology Program 2018a; Oberto et al. 2007; Tillmann et al. 2007; Toler et al. 1997) and *Rat* (Chou et al. 1992; Falcioni et al. 2018; Heikkinen et al. 2006; La Regina et al. 2003; National Toxicology Program 2018b).

Table 10. Comparison of chronic cancer bioassays reporting liver-related carcinogenicity outcomes

Author/ Year	Species (Group No.)	Strain	Sex	No animals per group	In utero exposure	Frequency (MHz)	Modulation	SAR	Exposure whole body (wb)/local	Duration (weeks)	Duration (h/d, d/w)	Type of Tumor (b/m)
Chou, 1992	Rat	SD	M	100	N/A	2450	MW	0.4	wb	100	21.5h/d, 7d/w	b+m

La Regina, 2003	Rat	F344	M + F	40	N/A	835.62	FDMA	1.3	wb	104	4h/d, 5d/w,	b+m
	Rat	F344	M + F	40	N/A	847.74	CDMA	1.3	wb	104	4h/d, 5d/w,	b+m
NTP, 2018	Rat	SD	M + F	90	GD6	900	GSM	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
	Rat	SD	M + F	90	GD6	900	CDMA	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
	Mouse	B6C3F1/N	M + F	90	N/A	1900	GSM	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
	Mouse	B6C3F1/N	M + F	90	N/A	1900	CDMA	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
Tillmann, 2007	Mouse	B6C3F1/N	M + F	50		1747	DCS	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m
	Mouse	B6C3F1/N	M + F	50		902	GSM	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m

Table 11. Certainty Assessment for Liver-Related Carcinogenicity

		Factors Decreasing Certainty "—" if no concerns, "↓"if serious concerns to downgrade certainty	**Factors Increasing Certainty** "—" not detected or applicable, "↑ "if sufficient to upgrade certainty	

Journal Pre-proofs

	INITIAL CERTAINTY RATING (# of studies)	Risk of bias	Unexplained Inconsistency	Indirectness	Imprecision	Publication Bias	Large Magnitude	Dose Response	Residual Confounding	Consistency Across Speies/Models	Rare Outcomes	FINAL CERTAINTY RATING
Liver	High (2 rat, 14 mouse)	--	↓	--	--	--	--	--	--	--	--	Moderate for an increased incidence of hepatoblastomas.

Citations: Mouse (Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Imaida et al. 1998a; Imaida et al. 1998b; Jauchem et al. 2001; Lerchl et al. 2015; National Toxicology Program 2018a; Oberto et al. 2007; Tillmann et al. 2007; Tillmann et al. 2010; Toler et al. 1997) and Rat (Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018b).

Table 12. Comparison of chronic cancer bioassays reporting lung-related carcinogenicity outcomes

Author/ Year	Species (Group No.)	Strain	Sex	No animals per group	In utero exposure	Frequency (MHz)	Modulation	SAR	Exposure whole body (wb)/local	Duration (weeks)	Duration (h/d, d/w)	Type of Tumor (b/m)
NTP, 2018	Rat	SD	M + F	90	GD6	900	GSM	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10' OFF	b+m

	Rat	SD	M + F	90	GD6	900	CDMA	1.5, 3, 6	wb	GD5-107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
NTP, 2018	Mouse	B6C3F1/N	M + F	90	N/A	1900	GSM	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
	Mouse	B6C3F1/N	M + F	90	N/A	1900	CDMA	2.5, 5, 10	wb	107	9h10''/d; 7d/w 10'ON/10' OFF	b+m
Tillmann, 2007	Mouse	B6C3F1/N	M + F	50		1747	DCS	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m
	Mouse	B6C3F1/N	M + F	50		902	GSM	0.4, 1.3, 4	wb	104	2h/d, 5d/w	b+m

Table 13. Certainty Assessment for Lung-Related Carcinogenicity

		Factors Decreasing Certainty "—" if no concerns, "↓"if serious concerns to downgrade certainty					**Factors Increasing Certainty** "—" not detected or applicable, "↑ "if sufficient to upgrade certainty					
	INITIAL CERTAINTY RATING (# of studies)	**Risk of bias**	**Unexplained Inconsistency**	**Indirectness**	**Imprecision**	**Publication Bias**	**Large Magnitude**	**Dose Response**	**Residual Confounding**	**Consistency Across Speies/Models**	**Rare Outcomes**	**FINAL CERTAINTY RATING**
Lung	High (8 rat, 23	--	↓	--	--	--	--	--	--	--	--	Moderate for an increased

	mouse)											incidence of lung tumors (bronchioalveolar adenoma or carcinoma) BMD 01: 0.729 (95% CI, 0.197)

Citations: Mouse (Frei et al. 1998a; Frei et al. 1998b; Heikkinen et al. 2006; Heikkinen et al. 2001; Jauchem et al. 2001; Lerchl et al. 2015; National Toxicology Program 2018a; Oberto et al. 2007; Sommer et al. 2007; Tillmann et al. 2007; Tillmann et al. 2010) and Rat (Chou et al. 1992; La Regina et al. 2003; National Toxicology Program 2018b)

Table 14. Summary of Review Findings

Certainty assessment *						Summary of findings	Certainty	Importance of Outcome
No of experiments	Risk of bias	Unexplained Inconsistency	Indirectness	Imprecision	Other considerations			
Lymphoma								
7 rat, 16 mouse	--	↓	--	--	--	Increased incidence of lymphoma was observed at 2.5 W/kg, 1900 MHz GSM and CDMA and at 5 W/kg, 1900 MHz GSM (whole body average SAR) for 9 hours and 10 minutes/day over 108 weeks in B6C3F1/N female mice (National Toxicology Program 2018a). One positive	Moderate	Serious

						study in tumor-prone female Pim-1 mice reported an increase in lymphoma at the highest SAR value ((1.5 W/kg) at 217 Hz pulsed 900 MHz GSM (Repacholi et al. 1997).		
Brain Neoplasia								
16 rat, 7 mouse	--	--	--	--	--	Increased incidence of gliomas at 6.0 W/kg (whole body average SAR). 900 MHz CDMA for 9 hours and 10 minutes/day over 106 weeks in male SD rats (National Toxicology Program 2018b). Increased incidence of oligodendrogliomas in male F344 rats with a statistically significant trend observed across the dose range of 0.16 W/kg to 1.6 W/kg at 1.6 GHz (whole body average SAR) given for 2 hours/day over 5 days/week over 100 weeks.	High	Serious
Heart Neoplasia								
4 rat, 2 mouse	--	--	--	--	--	Increased incidence of malignant schwannomas in male SD rats was observed with a statistically significant trend across the dose range of 1.5 W/kg to 6 W/kg whole body average SAR for 9 hours and 10 minutes/day, 900 MHz CDMA for 7 days/week over 106 weeks (NTP 2018a). An increased incidence of total schwannomas was also observed in male SD rats with statistically significance at the highest dose 0.1W/kg at 900 MHz GSM for 2 hours/day	High	Serious

Journal Pre-proofs

						for 5 weeks (Falcioni et al. 2018).		
Adrenal Gland Neoplasia								
6 rat, 11 mouse	--	↓	--	--	--	Increased incidence of malignant neoplasms (pheochromocytomas) in female SD rats at a 1.5 W/kg whole body average SAR, 900 MHz CDMA and male SD rats (F1) at 1.5 and 3.0 W/kg, 900 MHz GSM given for 9 hours and 10 minutes/day, 7 days/week over 106/7 weeks (National Toxicology Program 2018b). An increase in benign pheochromocytoma was found in the same study female SD rats at 900 MHz CDMA at 3.0 W/kg, and at 1.5 and 3.0 W/kg in male F1 generation SD rats at 900 MHz GSM (National Toxicology Program 2018b). A significant increase in hyperplasia was found in female rats at 3.0 and 6.0 W/kg with the trend being statistically significant (National Toxicology Program 2018b).	Moderate	Serious
Liver Neoplasia								
8 rat, 13 mouse	--	↓	--	--	--	Increased incidence of hepatoblastomas and hepatocellular carcinoma in male B6C3F1/N mice at 5.0 W/kg and 2.5 W/kg whole body average SAR, respectively at 1900 MHz CDMA given for 9 hours and 10 minutes/day for 7 days/week over 104/6 weeks (National Toxicology Program 2018a).	Moderate	Serious

Journal Pre-proofs

						In an initiation-promotion study, and increase in hepatocellular carcinoma was found in male C3H/HeNCrl mice, C57Bl/6N mice, females (combined), at all SAR levels investigated (0.04, 0.4 and 2 W/kg), exposed to 1966 MHz UMTS from gestation day 12 to 54 weeks (Lerchl et al. 2015). A increase in hepatocellular carcinoma was also found in B6C3F1 male and female (combined) mice at 0.58 W/kg, 1966 MHz UMTS exposed for 20 hours/day, 7 days/week and up to 24 months in an initiation-promotion study (Tillmann et al. 2010). In the same study, preneoplastic lesions, namely hepatocellular foci were increased at the highest SAR level (5.8 W/kg).		
Lung Neoplasia								
8 rat, 13 mouse	--	↓	--	--	--	Significant positive trend for an increased incidence of alveolar/bronchiolar adenomas or carcinomas (combined) in male B6C3F1/N mice at 1900 MHz GSM given for 9 hours and 10 minutes/day for 7 days/week over 108 weeks (National Toxicology Program 2018a). In the initiation-promotion studies by Tillmann (2010) and Lerchl (2015), increases in lung neoplasms were found: In male C3H/HeNCrl mice, and female C57Bl/6N mice (combined) an increase in bronchiolar adenomas was found at at all SAR levels	Moderate	Serious

						investigated (0.04, 0.4 and 2 W/kg), exposed to 1966 MHz UMTS from gestation day 12 to 54 weeks (Lerchl et al. 2015). A increase in bronchoalveolar adenoma and carcinoma was also found in B6C3F1 male and female (combined) mice at 0.58 W/kg, 1966 MHz UMTS exposed for 20 hours/day, 7 days/week and up to 24 months in an initiation-promotion study (Tillmann et al. 2010).		

*Upgrade factors are not presented because none were utilized.

Journal Pre-proofs