CCARS
Comité científico asesor en radiofrecuencias y salud

**Campos electromagnéticos y salud pública**

# Funciones Cognitivas y radiofrecuencias: la evidencia científica más robusta revisada por la OMS

Un análisis sistemático concluye que la evidencia epidemiológica disponible sobre los posibles efectos a largo plazo de la exposición a **CEM-RF** sobre la función cognitiva es escasa y presenta importantes limitaciones.

**Madrid, 4 de Noviembre de 2026**

# 4. Funciones cognitivas

Este cuarto bloque temático integra la evidencia sobre función cognitiva en relación con la exposición a CEM-RF. Como en los apartados previos, la selección del tema responde al proceso promovido por la OMS para priorizar las áreas de incertidumbre y relevancia en salud pública relacionadas con la exposición a CEM-RF (Verbeek et al., 2021). Dentro de ese programa, la función cognitiva fue uno de los resultados seleccionados para una evaluación sistemática. Dada la propia naturaleza de la pregunta, que comprende tanto posibles efectos inmediatos durante o tras la exposición como cambios asociados a exposiciones más prolongadas, la evidencia se abordó mediante dos revisiones complementarias: una experimental, basada en estudios aleatorizados en humanos con exposición controlada, y otra observacional, basada en estudios longitudinales en condiciones de vida real. Ambas disponen de protocolos prospectivos registrados en PROSPERO (Benke et al., 2022; Pophof et al., 2021), incorporan evaluación del riesgo de sesgo mediante OHAT y valoración de la certeza mediante GRADE, y fueron desarrolladas en el marco del programa de revisiones sistemáticas promovido por la OMS, con financiación total o parcial de esta organización.

La revisión experimental (Pophof et al., 2024) responde a la pregunta sobre los posibles efectos inmediatos o a corto plazo de la exposición a CEM-RF sobre el rendimiento cognitivo humano, comparando condiciones de exposición con condiciones de control o exposición simulada (*sham*) en estudios experimentales aleatorizados, con cegamiento de los participantes respecto a la condición de exposición.

Su protocolo, registrado en PROSPERO (CRD42021236168), estableció una búsqueda amplia en PubMed, Embase, Scopus, Web of Science y EMF-Portal y una estrategia de síntesis mediante metaanálisis de efectos aleatorios utilizando *Hedges' g*, diferenciando las medidas relacionadas con la precisión (*accuracy*) de aquellas relacionadas con la velocidad (*speed*). La taxonomía adoptada comprende siete grandes dominios cognitivos: orientación y atención, percepción, memoria, funciones verbales y habilidades lingüísticas, construcción y rendimiento motor, formación de conceptos y razonamiento y funciones ejecutivas, subdivididos a su vez en diferentes clases de funciones y pruebas.

El corpus experimental final comprendió 76 estudios descritos en 80 publicaciones, con un total de 3.846 participantes; de ellos, 50 estudios, correspondientes a 52 publicaciones y 2.433 participantes, aportaron datos utilizables para los metaanálisis. Predominaron las exposiciones localizadas en la cabeza y las señales GSM de 900 MHz en *uplink*. El riesgo de sesgo se evaluó mediante la herramienta OHAT, adaptada cuando fue necesario a los diseños cruzados y clasificando los estudios en tres niveles o *tiers*. La certeza se valoró mediante GRADE para los distintos desenlaces cognitivos. Este diseño experimental permite evaluar de forma controlada posibles cambios agudos y sutiles en variables como el tiempo de reacción, la velocidad de procesamiento o la exactitud de las respuestas, reduciendo la influencia de la confusión y estableciendo con claridad la temporalidad entre exposición y rendimiento. No obstante, la caracterización espacial de la exposición cerebral fue insuficiente en numerosos estudios, circunstancia que, entre otras razones, llevó a los autores a no realizar finalmente un análisis dosis-respuesta.

La revisión observacional (Benke et al., 2024) evalúa, por el contrario, los posibles efectos a más largo plazo, durante meses o años, de la exposición a CEM-RF sobre la función cognitiva humana. Su pregunta PECO comprende tanto la función cognitiva global como dominios específicos, entre ellos atención compleja, función ejecutiva, aprendizaje y memoria, lenguaje, capacidad perceptivo-motora y cognición social. La revisión fue registrada prospectivamente en PROSPERO (CRD42021257548) y reportada conforme a PRISMA.

Se realizaron búsquedas en PubMed, Embase, PsycInfo y EMF-Portal sin restricciones por fecha o idioma, y fueron elegibles estudios de cohortes y casos y controles en población general o trabajadora, de cualquier edad o sexo, siempre que comparasen al menos dos niveles de exposición y dispusieran de un seguimiento mínimo de seis meses. Los estudios transversales fueron excluidos debido a su incapacidad para establecer adecuadamente la temporalidad entre exposición y desenlace.

La exposición observacional podía proceder de fuentes personales de campo cercano, ambientales de campo lejano u ocupacionales. Dado que la dosis de energía absorbida en el tejido relevante, dependiente de la SAR, la duración y las características de la exposición, resulta difícil de determinar en estudios epidemiológicos, los autores admitieron diferentes aproximaciones indirectas, entre ellas intensidad de campo eléctrico, densidad de potencia, uso autoinformado de teléfonos móviles o inalámbricos y distintas métricas ocupacionales, incluidas matrices empleo-exposición.

Se excluyeron, en cambio, aproximaciones consideradas poco válidas, como la distancia autodeclarada a antenas o la distancia calculada a estaciones base de telefonía móvil. Esta limitación de la exposición resulta especialmente importante porque el uso del teléfono constituye simultáneamente un *proxy* de exposición a CEM-RF y una conducta compleja que incorpora llamadas, mensajería, juegos, redes sociales y otras actividades potencialmente relacionadas por sí mismas con el rendimiento cognitivo.

La evidencia observacional finalmente incluida fue considerablemente más reducida que la experimental: cinco estudios, correspondientes a cuatro cohortes y 4.639 participantes de Australia, Singapur y Suiza. La mayoría de los participantes eran niños o adolescentes, mientras que sólo una cohorte incluyó adultos de 55 años o más. La principal fuente de exposición fue el uso personal de teléfonos móviles o inalámbricos; sólo un estudio evaluó además exposición ambiental y ninguno investigó exposición ocupacional.

Los metaanálisis se realizaron mediante modelos de efectos fijos cuando fue posible combinar cuantitativamente los resultados. Aunque el protocolo había previsto modelos de efectos aleatorios, los autores modificaron este planteamiento porque sólo se disponía de dos estudios procedentes del mismo grupo de investigación para los principales metaanálisis, siguiendo las recomendaciones metodológicas para síntesis con muy pocos estudios. El riesgo de sesgo se evaluó mediante OHAT y la certeza mediante GRADE, considerando las modificaciones del enfoque OHAT para estudios de exposición ambiental.

Esta revisión observacional aporta, por tanto, una perspectiva que no pueden ofrecer los estudios experimentales de corta duración: la evaluación de posibles cambios cognitivos asociados a exposiciones mantenidas durante meses o años y estudiados longitudinalmente en contextos cotidianos. Sin embargo, esta mayor proximidad a las condiciones reales de exposición conlleva una menor capacidad de control experimental.

Las principales limitaciones se relacionan con la utilización de *proxies* de exposición, la posibilidad de confusión residual, la escasa variedad de poblaciones y fuentes estudiadas y la heterogeneidad entre las pruebas utilizadas para evaluar los diferentes dominios cognitivos. La propia revisión señala además que no existen estudios elegibles sobre exposición ambiental y cognición en adultos o personas mayores ni sobre exposición ocupacional en adultos, y que varios dominios cognitivos apenas han sido investigados.

La presentación de este bloque en dos subapartados responde, por tanto, a preguntas distintas y complementarias. La revisión experimental aborda si pueden detectarse efectos agudos sobre el rendimiento cognitivo bajo condiciones de exposición controladas, mientras que la observacional examina si exposiciones más prolongadas se asocian con cambios en la función cognitiva en condiciones de vida real. En el primer caso, la aleatorización, el cegamiento y el control experimental de la exposición permiten evaluar de forma más directa posibles efectos inmediatos y reducir la influencia de la confusión; en el segundo, los estudios longitudinales permiten analizar cambios cognitivos asociados a periodos de exposición más prolongados, aunque con mayor dependencia de medidas indirectas de exposición y del control estadístico de posibles factores de confusión.

En el análisis CCARS que sigue mantenemos, por tanto, la trazabilidad PECO y la separación explícita entre ambas líneas de evidencia. Este enfoque permite interpretar de forma coordinada la evidencia experimental sobre efectos agudos y la evidencia observacional sobre exposiciones más prolongadas, sin asumir que ambas responden exactamente a la misma pregunta ni extrapolar más allá de lo que permite cada diseño. La ausencia de un efecto cognitivo inmediato en un experimento controlado no permite por sí sola excluir posibles efectos asociados a exposiciones mantenidas durante meses o años, del mismo modo que una asociación observacional a largo plazo no demostraría por sí misma un efecto causal de los CEM-RF.

En suma, el bloque de funciones cognitivas se organiza en dos subapartados: estudios experimentales en humanos, centrados en los posibles efectos inmediatos o a corto plazo bajo exposición controlada, y estudios observacionales humanos, enfocados en posibles cambios a más largo plazo, durante meses o años, en población general y trabajadora. La combinación de ambas perspectivas proporciona una síntesis complementaria y metodológicamente estructurada para valorar, con certeza graduada, la relación entre exposición a CEM-RF y función cognitiva, atendiendo tanto a la mayor validez interna de los estudios experimentales como a la aplicabilidad de la evidencia longitudinal obtenida en condiciones cotidianas.

# 4.1. Revisión estudios experimentales

## 4.1.1. Referencia completa del artículo

**Título:** *The effect of exposure to radiofrequency electromagnetic fields on cognitive performance in human experimental studies: Systematic review and meta-analyses*. Pophof B., Kuhne J., Schmid G., Weiser E., Dorn H., Henschenmacher B., Burns J., Danker-Hopfe H., Sauter C. **Revista y datos bibliográficos:** Environment International, 191, 108899 (2024). **DOI:** 10.1016/j.envint.2024.108899. **Registro de protocolo:** PROSPERO CRD42021236168.

## 4.1.2. Objetivos y preguntas de investigación

El objetivo principal de Pophof et al. (2024) fue evaluar los posibles efectos de la exposición a corto plazo a campos electromagnéticos de radiofrecuencia (CEM-RF) sobre el rendimiento cognitivo humano en estudios experimentales. La revisión se planteó específicamente determinar si, bajo condiciones controladas de exposición, el rendimiento en diferentes funciones cognitivas difiere entre una condición de exposición a CEM-RF y una condición de control, constituida por exposición simulada (*sham*), ausencia de exposición por encima del nivel de fondo o un nivel inferior de exposición. El interés se centra, por tanto, en efectos agudos o a corto plazo, diferenciándose de la revisión observacional complementaria de Benke et al. (2024), orientada a posibles efectos asociados a exposiciones mantenidas durante meses o años.

La pregunta de investigación se estructuró mediante el marco PECO. La población (P) estuvo constituida por seres humanos, sin restricciones generales por edad, sexo o estado de salud. Se contemplaron, por tanto, participantes sanos y personas con diferentes condiciones de salud, con el propósito de abarcar también posibles grupos susceptibles. La exposición (E) comprendió CEM-RF dentro del intervalo de frecuencias de 100 kHz a 300 GHz, aplicados experimentalmente bajo condiciones controladas. Los estudios debían disponer de al menos dos condiciones de exposición y cumplir criterios previamente establecidos destinados a garantizar un contraste de exposición suficiente y una caracterización adecuada de los CEM-RF. El comparador (C) podía consistir en una exposición simulada (*sham*), ausencia de exposición adicional respecto del nivel de fondo o una exposición a CEM-RF de menor intensidad.

El desenlace (O) fue el rendimiento cognitivo, clasificado conforme a la taxonomía neuropsicológica de Lezak et al. (2012) en siete grandes dominios: orientación y atención; percepción; memoria; funciones verbales y habilidades lingüísticas; construcción y rendimiento motor; formación de conceptos y razonamiento; y funciones ejecutivas. Dentro de estos dominios se consideraron diferentes subclases y pruebas cognitivas. El rendimiento se evaluó separando dos dimensiones fundamentales: medidas relacionadas con la exactitud, como respuestas correctas, incorrectas u omitidas, y medidas relacionadas con la velocidad, como tiempo de reacción, velocidad de respuesta o tiempo de procesamiento. Esta distinción se mantuvo posteriormente en la síntesis cuantitativa, realizándose metaanálisis separados para las medidas de exactitud y velocidad.

La revisión se restringió a estudios experimentales humanos aleatorizados, de grupos paralelos o diseño cruzado (*cross-over*), realizados en laboratorio o en otros entornos bajo condiciones controladas y en los que, como mínimo, los participantes desconocieran si se encontraban en la condición de exposición o de control. De este modo, la pregunta no se limitó a comprobar si existían diferencias estadísticamente significativas entre exposición y *sham*, sino que pretendió sintetizar cuantitativamente la magnitud y dirección de los posibles efectos de la exposición a CEM-RF sobre las distintas dimensiones del rendimiento cognitivo y establecer la certeza de la evidencia disponible para cada desenlace mediante GRADE.

Un aspecto metodológicamente relevante es que el protocolo contemplaba también la posible evaluación de la influencia de las características de la exposición sobre los resultados, incluida la relación entre nivel de exposición y efecto. Esta cuestión resulta particularmente importante en estudios de exposición localizada de la cabeza, dado que una eventual respuesta cognitiva podría depender no sólo del nivel global de SAR, sino también de su distribución espacial en las regiones cerebrales implicadas en la tarea evaluada. Como veremos posteriormente, las limitaciones en la caracterización espacial de la exposición de muchos estudios condicionaron finalmente la posibilidad de investigar adecuadamente esta cuestión.

## 4.1.3. Metodología

La revisión se desarrolló a partir de un protocolo metodológico previamente establecido, publicado por Pophof et al. (2021) y registrado prospectivamente en PROSPERO (CRD42021236168). La metodología se basó en el *WHO Handbook for Guideline Development* y la revisión se informó de acuerdo con las directrices PRISMA. El protocolo definió previamente la pregunta PECO, los criterios de elegibilidad, la estrategia de búsqueda, los procedimientos de selección y extracción, la evaluación del riesgo de sesgo, el plan de síntesis cuantitativa y la valoración de la certeza de la evidencia. Durante la realización de la revisión se introdujeron algunas modificaciones respecto al protocolo, fundamentalmente para adaptar la clasificación de los desenlaces a la gran cantidad y diversidad de pruebas cognitivas encontradas y para adecuar determinados análisis a las características de la evidencia finalmente disponible.

### *Búsqueda y fuentes*

La búsqueda bibliográfica se realizó en cinco bases de datos: PubMed, Embase, Scopus, Web of Science y EMF-Portal, esta última especializada en literatura científica sobre los efectos de los campos electromagnéticos. Las búsquedas comenzaron en mayo de 2021 y se actualizaron en todas las bases hasta el 2 de marzo de 2023. La estrategia se desarrolló inicialmente para PubMed y posteriormente se adaptó a las restantes bases, combinando términos relativos a la exposición y los desenlaces definidos en la pregunta PECO. En PubMed y Embase se utilizaron además términos de vocabulario controlado. Los autores ajustaron la estrategia con el objetivo de mantener un equilibrio adecuado entre sensibilidad y precisión.

Como procedimiento complementario, se revisaron las listas de referencias de las revisiones relevantes identificadas mediante la búsqueda electrónica con el fin de localizar estudios potencialmente elegibles que no hubieran sido recuperados por las bases bibliográficas. No se establecieron restricciones por fecha de publicación ni por idioma.

Sin embargo, únicamente se admitieron artículos revisados por pares que proporcionasen datos primarios; quedaron excluidas revisiones, declaraciones, informes, artículos de opinión, comentarios, editoriales, resúmenes de congresos y actas de reuniones científicas. Cuando fue necesario traducir resúmenes o artículos para determinar su elegibilidad, los autores utilizaron Google Translate.

### *Criterios de elegibilidad*

Se incluyeron estudios experimentales aleatorizados en humanos, tanto de grupos paralelos como de diseño cruzado (*cross-over*), realizados en laboratorio bajo condiciones controladas. Los estudios debían comparar al menos dos niveles de exposición a CEM-RF, incluyendo como referencia una condición *sham* o control, y al menos los participantes debían desconocer la condición de exposición a la que estaban siendo sometidos. Se excluyeron explícitamente los estudios no aleatorizados y aquellos en los que los participantes conocían la condición de exposición, debido al mayor riesgo de confusión y sesgo de selección y a la disponibilidad de estudios experimentales de mayor calidad metodológica.

La exposición debía corresponder a campos eléctricos, magnéticos o electromagnéticos en el intervalo de 100 kHz a 300 GHz y estar suficientemente caracterizada. Se excluyeron los estudios con un contraste insuficiente entre condiciones de exposición, con caracterización inadecuada de la exposición o con coexposición a campos electromagnéticos situados fuera del intervalo de frecuencias especificado. No se establecieron restricciones generales respecto a edad, sexo o estado de salud, por lo que podían incluirse participantes sanos o con determinadas condiciones clínicas, así como personas con intolerancia ambiental idiopática atribuida a campos electromagnéticos (IEI-EMF).

El desenlace de interés fue el rendimiento cognitivo. Las pruebas se clasificaron conforme a la taxonomía de Lezak et al. (2012) en siete grandes dominios: orientación y atención; percepción; memoria; funciones verbales y habilidades lingüísticas; construcción y rendimiento motor; formación de conceptos y razonamiento; y funciones ejecutivas. Se utilizaron además las clasificaciones propuestas por Fulda y Schulz para pruebas adicionales. Cuando una prueba concreta no aparecía clasificada en estas fuentes, se asignó de acuerdo con la capacidad cognitiva principal requerida para su realización. Dentro de cada subclase se diferenciaron las medidas relacionadas con la exactitud, por ejemplo, respuestas correctas, incorrectas u omitidas, de las relacionadas con la velocidad, tiempo de reacción, velocidad de respuesta o tiempo de procesamiento.

Esta clasificación de los desenlaces representa una de las modificaciones relevantes respecto al protocolo. Debido a la gran cantidad de datos disponibles, especialmente para orientación y atención y, en menor medida, formación de conceptos y razonamiento, y a que algunos estudios proporcionaban resultados para numerosas pruebas diferentes, mantener las categorías inicialmente previstas habría generado importantes problemas de unidad de análisis y habría impedido incorporar una proporción considerable de los datos originales a los metaanálisis.

Los autores optaron, por ello, por utilizar categorías más detalladas correspondientes a las subclases definidas por Lezak, permitiendo diferenciar mejor las distintas capacidades cognitivas y aumentar el número de resultados susceptibles de síntesis cuantitativa.

### *Selección y extracción de datos*

Los registros bibliográficos se gestionaron inicialmente mediante Covidence. La selección se realizó en dos fases. En primer lugar, dos revisores evaluaron de forma independiente los títulos y resúmenes y excluyeron los registros claramente irrelevantes. Posteriormente, dos revisores examinaron de manera independiente los textos completos y aplicaron los criterios de inclusión y exclusión. Los desacuerdos se resolvieron mediante discusión y, cuando no se alcanzaba consenso, se recurría a un tercer revisor. Las publicaciones excluidas tras la evaluación del texto completo y sus motivos de exclusión quedaron documentados. Cuando varios artículos informaban resultados procedentes de un mismo estudio, fueron considerados conjuntamente como un único estudio; a la inversa, cuando una publicación describía varios experimentos independientes, estos se trataron como estudios separados.

Para cada estudio se extrajo información sobre la publicación y las características del diseño, país, financiación y conflictos de interés; características de la población, incluido número de participantes, edad y sexo; parámetros de exposición, como intensidad o nivel, frecuencia y modulación; pruebas y desenlaces cognitivos y sus resultados numéricos; y aspectos metodológicos relevantes para la posterior evaluación del riesgo de sesgo, como aleatorización, contrabalanceo y cegamiento. Aunque inicialmente se había previsto utilizar Covidence también para esta fase, los autores emplearon finalmente una plantilla estandarizada en Microsoft Excel debido al elevado volumen de información. Un revisor realizó la extracción y un segundo comprobó los datos frente a la publicación original para garantizar su integridad y exactitud. Los autores de la revisión que también lo eran de algún estudio primario incluido no participaron en la extracción de los datos de sus propios trabajos.

Cuando faltaba información necesaria para determinar la elegibilidad de un estudio, evaluar su riesgo de sesgo o incorporarlo a la síntesis cuantitativa, se contactó por correo electrónico con el autor correspondiente. En ausencia de respuesta se realizaron dos nuevos intentos; si tampoco se obtenía la información, los datos se consideraban ausentes.

### *Medidas de efecto y síntesis cuantitativa*

La síntesis cuantitativa se realizó mediante metaanálisis de efectos aleatorios para aquellos desenlaces en los que existían al menos dos estudios suficientemente comparables. Dada la diversidad de pruebas y escalas utilizadas para evaluar una misma función cognitiva, los efectos se expresaron mediante la diferencia de medias estandarizada corregida para muestras pequeñas, g de Hedges, con su intervalo de confianza del 95 %.

Los metaanálisis se realizaron por subclase cognitiva y separadamente para las medidas de exactitud y velocidad, evitando combinar resultados que, aunque procedieran de una misma prueba o dominio general, representasen dimensiones distintas del rendimiento cognitivo.

La heterogeneidad estadística se cuantificó mediante $I^2$. Cuando $I^2$ era superior al 30 %, considerado potencialmente indicativo de al menos heterogeneidad moderada, se exploraron posibles fuentes de variabilidad mediante análisis de subgrupos.

Entre los factores examinados se incluyeron el riesgo de sesgo, *tiers* 1 y 2 frente a *tier* 3, diseño *cross-over* frente a grupos paralelos, tamaño muestral inferior a 30 frente a 30 participantes o más, características de la población, características de la población relacionadas con IEI-EMF, mediante comparaciones limitadas entre adultos sin IEI-EMF y grupos que incluían adultos con y sin IEI-EMF, adultos frente a niños o adolescentes, sexo (en la práctica, estudios realizados sólo en hombres frente a estudios que incluían hombres y mujeres, dado que sólo existía un estudio exclusivamente femenino), región corporal expuesta, cabeza frente a cuerpo entero, fuente de financiación y tipo de cegamiento. No se realizaron los análisis separados de personas con y sin IEI-EMF previstos inicialmente, debido al reducido número de estudios disponibles. Los análisis estadísticos se realizaron mediante *Comprehensive Meta-Analysis V3*.

El protocolo había previsto explorar también características específicas de la exposición, incluida una posible relación exposición-respuesta. Sin embargo, este análisis no se realizó finalmente. Los autores justificaron esta decisión porque los resultados mostraban predominantemente ausencia de efecto con independencia del nivel de exposición y, sobre todo, porque en muchos estudios la caracterización dosimétrica no proporcionaba información espacial suficientemente detallada sobre la exposición en las diferentes regiones cerebrales. Por razones semejantes, tampoco se realizaron los análisis de subgrupos inicialmente previstos según modulación, duración de la exposición, campo cercano frente a campo lejano y frecuencia. Los autores consideraron que la concentración de numerosos estudios y de varios de los valores atípicos en una misma categoría de exposición habría limitado la capacidad de estos análisis para explicar la heterogeneidad.

El posible sesgo de publicación se examinó mediante inspección visual de *funnel plots* para todos los metaanálisis. Cuando un análisis incluía diez o más estudios se aplicó además la prueba de Egger, utilizando $p < 0,1$ como umbral indicativo de posible asimetría.

### *Evaluación del riesgo de sesgo*

El riesgo de sesgo se evaluó mediante la herramienta *Risk of Bias Rating Tool for Human and Animal Studies* desarrollada por la Office of Health Assessment and Translation del National Toxicology Program (NTP/OHAT).

Se utilizaron principalmente las preguntas diseñadas para ensayos clínicos humanos, pero, dado que esta herramienta está orientada fundamentalmente a estudios de grupos paralelos, se añadieron tres preguntas para valorar adecuadamente los diseños *cross-over*: una adaptada de la herramienta OHAT para estudios experimentales en animales y dos procedentes de Cochrane RoB 2.0. De esta forma se incorporaron específicamente cuestiones relevantes para los diseños cruzados, como posibles efectos de periodo y *carry-over*.

Debe señalarse que el procedimiento concreto para transformar las valoraciones de los distintos dominios de riesgo de sesgo en una clasificación global por *tiers* no había sido especificado en el protocolo. Los autores reconocen esta desviación y justifican detalladamente en la revisión los criterios finalmente utilizados, señalando que podría constituir una fuente potencial de sesgo, especialmente por su influencia sobre la posterior valoración GRADE.

Cada ítem se clasificó en cuatro categorías: riesgo de sesgo definitivamente bajo, probablemente bajo, probablemente alto, categoría en la que se incluyó también la ausencia de información o riesgo incierto, y definitivamente alto.

La evaluación fue realizada independientemente por dos revisores y los desacuerdos se resolvieron mediante discusión. Como en la extracción de datos, los autores de la revisión no evaluaron el riesgo de sesgo de estudios propios.

A partir de las evaluaciones por dominios, los estudios se clasificaron globalmente en tres niveles o *tiers* de acuerdo con el enfoque OHAT. Los autores consideraron especialmente importantes dos dominios: el cegamiento y la evaluación del desenlace. Los estudios con todos los dominios clasificados como definitivamente o probablemente bajos se asignaron al *tier* 1, permitiéndose como máximo un dominio probablemente alto siempre que no correspondiera a una de esas dos cuestiones clave.

Se clasificaron en *tier* 3 los estudios con riesgo definitivamente alto en dos o más dominios o en cualquiera de los dos dominios clave, así como determinadas combinaciones de riesgo alto en otros dominios con riesgo probablemente alto en uno de los dominios clave. Los restantes estudios se clasificaron en *tier* 2. Esta decisión implica, por ejemplo, que los estudios únicamente cegados para los participantes, pero no para el personal investigador, fueran clasificados en el nivel global de mayor riesgo de sesgo.

Aunque la caracterización de la exposición no fue seleccionada como uno de esos dos dominios clave para establecer los *tiers*, los autores subrayan su importancia para la interpretación mecanística y, especialmente, para cualquier análisis exposición-respuesta. Todos los estudios incluidos garantizaban un contraste suficiente para considerar que los participantes habían sido realmente expuestos, pero en algunos casos el nivel absoluto y, sobre todo, la distribución espacial de la exposición en la cabeza no podía determinarse con precisión.

### *Certeza de la evidencia*

La certeza de la evidencia se evaluó para cada cuerpo de evidencia mediante el enfoque GRADE. Al tratarse de estudios experimentales aleatorizados, la evidencia partió de un nivel inicial de certeza alta. Posteriormente se consideraron cinco posibles motivos de degradación: limitaciones metodológicas o riesgo de sesgo, indirectitud, inconsistencia, imprecisión y posible sesgo de publicación.

Los autores establecieron criterios explícitos para operacionalizar estas degradaciones. Por riesgo de sesgo, se reducía un nivel cuando entre el 50 % y menos del 75 % de los estudios de un determinado resultado pertenecían al *tier* 3 y dos niveles cuando esa proporción alcanzaba o superaba el 75 %.

No se aplicaron degradaciones por indirectitud, puesto que los criterios de inclusión habían restringido la evidencia a estudios directamente relacionados con la pregunta PECO. Para inconsistencia, se reducía un nivel cuando $I^2$ se situaba entre el 30 % y el 90 %, y dos cuando era igual o superior al 90 %; no obstante, si los análisis de subgrupos mostraban que la heterogeneidad desaparecía en los estudios de *tiers* 1–2 o en aquellos con al menos 30 participantes, se consideraba explicada y no se degradaba por este criterio.

La imprecisión se valoró atendiendo tanto a la magnitud de la g de Hedges como a la amplitud de su IC95 %. Cuando la estimación se situaba entre −0,2 y 0,2 pero el intervalo incluía −0,2 o +0,2, se degradaba un nivel la certeza respecto a la ausencia de efecto; cuando la estimación se situaba fuera de ese intervalo, pero el IC95 % incluía cero, se degradaba un nivel respecto a la presencia de un efecto.

Se aplicaba una degradación adicional cuando el IC95 % era más amplio que el intervalo comprendido entre −1,0 y 1,0. Finalmente, la certeza se reducía por posible sesgo de publicación únicamente cuando coincidían un resultado de la prueba de Egger con $p < 0,1$ y una asimetría compatible en la inspección visual del *funnel plot*. Los autores señalan expresamente que esta operacionalización de la inconsistencia y la imprecisión fue más estricta que la prevista inicialmente en el protocolo.

La certeza final se clasificó como alta, moderada, baja o muy baja. Para comunicar las conclusiones asociadas a cada nivel se utilizó la terminología estandarizada propuesta por Santesso et al. (2020), de modo que la formulación de los resultados, por ejemplo, que la exposición «no reduce», «probablemente no reduce» o «puede no reducir» el rendimiento cognitivo, dependió explícitamente de la certeza alcanzada para cada desenlace y no únicamente de la significación estadística de la estimación agrupada.

## 4.1.4. Resultados

### *Flujo PRISMA*

La búsqueda bibliográfica identificó inicialmente 57.543 registros. Tras eliminar los duplicados, quedaron 23.450 registros únicos que fueron sometidos al cribado por título y resumen. De ellos, 245 publicaciones pasaron a evaluación en texto completo. Finalmente, 80 publicaciones, correspondientes a 76 estudios, cumplieron los criterios de inclusión de la revisión. Los estudios incluidos comprendieron en conjunto 3.846 participantes.

No todos los estudios proporcionaron información numérica suficiente para su incorporación a la síntesis cuantitativa. Cincuenta estudios, descritos en 52 publicaciones y con un total de 2.433 participantes, aportaron datos cuantitativos utilizables; 36 proporcionaron datos para metaanálisis de medidas relacionadas con la exactitud y 40 para medidas relacionadas con la velocidad, con solapamiento entre ambos grupos. Los 26 estudios restantes no proporcionaron información numérica adecuada para los metaanálisis y fueron considerados en la síntesis narrativa. Todos los estudios incluidos compararon el rendimiento bajo exposición a CEM-RF con una condición *sham*.

### *Características de los estudios incluidos*

Los 76 estudios se realizaron entre 1989 y 2021 en 19 países e incluyeron participantes de diferentes edades, sexos y estados de salud. Los 50 estudios incorporados a los metaanálisis procedían de 15 países, fueron publicados entre 2001 y 2021 y comprendieron 2.433 participantes. Predominaron claramente los estudios en adultos; la evidencia disponible para niños y adolescentes fue mucho más reducida y la correspondiente a personas de edad avanzada especialmente escasa.

Los estudios utilizaron fundamentalmente diseños experimentales cruzados (*cross-over*), aunque también se incluyeron diseños de grupos paralelos. En general, las muestras fueron pequeñas, característica habitual de este tipo de estudios experimentales. Entre los 50 estudios con información cuantitativa, el tamaño muestral mediano fue de 24 participantes.

Existió, además, una notable heterogeneidad metodológica en aspectos potencialmente relevantes para el rendimiento cognitivo, como el entrenamiento previo en las tareas, el orden de presentación de estas, el contrabalanceo de las condiciones experimentales, la separación temporal entre sesiones, la duración de la exposición y el control de posibles efectos residuales entre sesiones en los diseños *cross-over*.

La situación experimental predominante fue la exposición localizada de la cabeza, especialmente mediante señales GSM de 900 MHz en *uplink*, aunque las características concretas de frecuencia, modulación, duración e intensidad variaron entre estudios.

La exposición localizada en la cabeza resulta particularmente relevante para esta pregunta de investigación por la exposición comparativamente elevada del cerebro que puede producirse durante una llamada telefónica. Sin embargo, aunque todos los estudios incluidos presentaban contraste suficiente entre las condiciones experimental y *sham*, la caracterización dosimétrica no siempre permitía conocer adecuadamente el nivel absoluto y, especialmente, la distribución espacial de la exposición dentro del cerebro.

Las pruebas cognitivas fueron clasificadas en los siete dominios de Lezak et al.: orientación y atención; percepción; memoria; funciones verbales y habilidades lingüísticas; construcción y rendimiento motor; formación de conceptos y razonamiento; y funciones ejecutivas.

Dentro de ellos se establecieron subclases más específicas y se analizaron separadamente las medidas de velocidad y exactitud. La evidencia se concentró especialmente en orientación y atención, mientras que fue mucho más limitada para otros dominios. Sólo dos estudios aportaron información sobre funciones verbales y habilidades lingüísticas y cuatro sobre construcción y rendimiento motor. No se identificó ningún estudio elegible para el dominio de funciones ejecutivas.

### *Resultados globales de los metaanálisis*

El resultado más consistente de la revisión fue que ninguno de los metaanálisis mostró un efecto estadísticamente significativo de la exposición a CEM-RF frente a la condición *sham* sobre el rendimiento cognitivo, atendiendo tanto a los intervalos de confianza de la g de Hedges como a las correspondientes pruebas *z*. Este patrón se observó en las medidas relacionadas con la velocidad y con la exactitud y en los distintos dominios y subclases para los que fue posible realizar una síntesis cuantitativa.

Esto no implica que todas las estimaciones fueran exactamente nulas ni que la certeza fuese idéntica para todos los desenlaces. Las estimaciones puntuales oscilaron en ambas direcciones y algunos análisis mostraron heterogeneidad o intervalos de confianza amplios. Por ello, la interpretación de los autores se realizó combinando magnitud del efecto, precisión, heterogeneidad, riesgo de sesgo y certeza GRADE, en lugar de basarse únicamente en la significación estadística.

### *Orientación y atención*

Orientación y atención fue el dominio con mayor cantidad de evidencia y comprendió numerosas subclases relacionadas con capacidad atencional, atención focalizada, vigilancia, atención selectiva y dividida, velocidad de procesamiento y memoria de trabajo.

Para capacidad atencional (*attentional capacity*), cinco estudios pudieron combinarse para las medidas de exactitud. La g de Hedges fue 0,024 (IC95 % −0,10 a 0,15), con $I^2$ = 28 % y 473 participantes. La estimación se situó prácticamente en el valor nulo y los autores concluyeron que la exposición a CEM-RF produce poca o ninguna diferencia en la exactitud. No se dispuso de medidas de velocidad susceptibles de metaanálisis para esta subclase.

Para concentración o atención focalizada, la estimación para velocidad fue g = 0,005 (IC95 % −0,17 a 0,18; $I^2$ = 7 %; 132 participantes), mientras que para exactitud fue g = 0,097 (IC95 % −0,05 a 0,24; $I^2$ = 0 %; 217 participantes). Ambos resultados fueron compatibles con poca o ninguna diferencia entre exposición y *sham*. Para vigilancia, la velocidad presentó g = 0,118 (IC95 % −0,04 a 0,28; $I^2$ = 41 %; 247 participantes) y la exactitud g = 0,042 (IC95 % −0,09 a 0,18; $I^2$ = 0 %; 199 participantes). Tampoco se observó evidencia de deterioro del rendimiento.

En atención selectiva, las estimaciones fueron g = 0,080 (IC95 % −0,09 a 0,25; $I^2$ = 63 %; 452 participantes) para velocidad y g = 0,178 (IC95 % −0,02 a 0,38; $I^2$ = 68 %; 480 participantes) para exactitud. Aunque las estimaciones puntuales se situaron por encima de cero, los intervalos de confianza incluyeron el valor nulo y existió heterogeneidad moderada o sustancial. Para atención dividida, los resultados fueron igualmente compatibles con poca o ninguna diferencia: g = −0,010 (IC95 % −0,14 a 0,12; $I^2$ = 5 %; 307 participantes) para velocidad y g = −0,089 (IC95 % −0,35 a 0,18; $I^2$ = 53 %; 167 participantes) para exactitud.

La velocidad de procesamiento fue evaluada mediante distintos tipos de tareas. En las tareas de tiempo de reacción simple, el efecto agrupado fue g = 0,069 (IC95 % −0,02 a 0,16; $I^2$ = 29 %; 820 participantes), que constituye uno de los conjuntos de datos más amplios de la revisión. En tareas de elección entre dos alternativas, la velocidad presentó g = −0,023 (IC95 % −0,13 a 0,08; $I^2$ = 0 %; 401 participantes) y la exactitud g = −0,063 (IC95 % −0,38 a 0,25; $I^2$ = 63 %; 117 participantes). En tareas con más de dos alternativas, los valores correspondientes fueron g = −0,054 (IC95 % −0,14 a 0,03; $I^2$ = 0 %; 544 participantes) para velocidad y g = −0,129 (IC95 % −0,30 a 0,04; $I^2$ = 0 %; 131 participantes) para exactitud. Ninguna de estas estimaciones indicó un deterioro estadísticamente significativo asociado a la exposición.

Para otras tareas de velocidad de procesamiento, la g fue 0,067 (IC95 % −0,12 a 0,26; $I^2$ = 38 %; 249 participantes) en velocidad y 0,036 (IC95 % −0,08 a 0,15; $I^2$ = 0 %; 354 participantes) en exactitud. Nuevamente, los resultados fueron compatibles con poca o ninguna diferencia.

La memoria de trabajo dentro de este dominio fue analizada particularmente mediante tareas *n-back*. Para las diferentes variantes entre 0-*back* y 3-*back*, las estimaciones de g de Hedges oscilaron entre −0,090 (IC95 % −0,18 a 0,01) y 0,060 (IC95 % −0,06 a 0,18), con $I^2$ = 0 % en todos estos análisis y entre 237 y 474 participantes. Los autores concluyeron que la exposición produce poca o ninguna diferencia tanto en velocidad como en exactitud. Para *mental tracking*, la exactitud presentó igualmente una estimación próxima a cero, g = −0,047 (IC95 % −0,15 a 0,05; $I^2$ = 0 %; 438 participantes).

En conjunto, por tanto, el dominio con mayor cantidad de evidencia no mostró un patrón compatible con un deterioro agudo del rendimiento atencional, de la velocidad de procesamiento o de la memoria de trabajo como consecuencia de la exposición a CEM-RF.

### *Percepción y memoria*

La evidencia sobre percepción fue considerablemente menor. Para percepción visual y auditiva, la estimación agrupada fue g = −0,015 (IC95 % −0,23 a 0,195; $I^2$ = 0 %; 84 participantes) en las medidas de velocidad y g = 0,035 (IC95 % −0,13 a 0,199; $I^2$ = 0 %; 137 participantes) para exactitud. Ambos resultados fueron compatibles con poca o ninguna diferencia entre exposición y *sham*.

En el dominio de memoria, para memoria verbal y visual la velocidad presentó g = 0,042 (IC95 % −0,15 a 0,23; $I^2$ = 0 %; 102 participantes). Para exactitud, la estimación fue g = −0,087 (IC95 % −0,38 a 0,20; 625 participantes), pero en este caso existió una heterogeneidad considerable, $I^2$ = 85 %. La estimación agrupada siguió siendo compatible con ausencia de efecto, aunque la elevada heterogeneidad redujo la confianza con la que podía interpretarse este resultado.

### *Funciones verbales, rendimiento motor, razonamiento y funciones ejecutivas*

La evidencia para funciones verbales y habilidades lingüísticas fue muy escasa. Sólo dos estudios evaluaron expresión verbal y no pudo realizarse metaanálisis porque uno de ellos no proporcionaba valores numéricos adecuados. Ninguno de los dos comunicó resultados indicativos de efectos estadísticamente significativos de la exposición sobre velocidad o exactitud.

Para construcción y rendimiento motor, concretamente habilidades motoras, la síntesis de las medidas de velocidad produjo la estimación más extrema de toda la revisión: g = −0,919 (IC95 % −3,09 a 1,26; $I^2$ = 96 %; 42 participantes). Sin embargo, el intervalo de confianza fue extraordinariamente amplio, incluyó ampliamente el valor nulo, la heterogeneidad fue extrema y el número de participantes muy reducido.

Por ello, aunque la estimación puntual se orientaba hacia una reducción de la velocidad, los autores calificaron esta evidencia como muy incierta. Para exactitud, la estimación fue g = 0,228 (IC95 % −0,01 a 0,46; $I^2$ = 0 %; 109 participantes), compatible con poca o ninguna diferencia y sin evidencia de reducción de la exactitud.

En formación de conceptos y razonamiento, la subclase de razonamiento presentó g = 0,010 (IC95 % −0,11 a 0,13; $I^2$ = 0 %; 263 participantes) para velocidad y g = 0,051 (IC95 % −0,14 a 0,25; $I^2$ = 0 %; 100 participantes) para exactitud. Para procedimientos matemáticos, los valores fueron g = 0,033 (IC95 % −0,12 a 0,18; $I^2$ = 0 %; 168 participantes) en velocidad y g = 0,232 (IC95 % −0,12 a 0,59; $I^2$ = 86 %; 253 participantes) en exactitud. En este último análisis, aunque la estimación puntual fue positiva, la heterogeneidad fue elevada y el intervalo incluyó el valor nulo.

Un resultado relevante por ausencia de evidencia es que no se identificó ningún estudio elegible para el dominio de funciones ejecutivas. Por tanto, la revisión no permite extraer conclusiones experimentales sobre los efectos agudos de los CEM-RF en este dominio cognitivo.

### *Heterogeneidad y análisis de subgrupos*

La heterogeneidad fue muy variable entre los distintos metaanálisis. Numerosas síntesis presentaron $I^2$ = 0 % o valores bajos, particularmente cuando las tareas agrupadas eran muy semejantes, como ocurrió en varios análisis de *n-back*, tiempos de reacción y razonamiento. En otros desenlaces se observó heterogeneidad considerable, como en exactitud de memoria verbal y visual ($I^2$ = 85 %), exactitud en procedimientos matemáticos ($I^2$ = 86 %) y velocidad en habilidades motoras ($I^2$ = 96 %).

Los autores exploraron posibles fuentes de heterogeneidad mediante los análisis de subgrupos previstos y mediante una estratificación adicional según tamaño muestral, inferior a 30 frente a 30 participantes o más. Sin embargo, no se identificó una explicación única y consistente de la heterogeneidad. Parte de ella podía relacionarse con la diversidad de pruebas agrupadas dentro de una misma subclase y con determinados estudios que producían estimaciones alejadas de las restantes.

La revisión no realizó finalmente análisis exposición-respuesta ni los análisis de subgrupos previstos para modulación, duración de la exposición, campo cercano frente a campo lejano y frecuencia. Los autores justificaron esta decisión por el predominio general de resultados nulos independientemente del nivel de exposición y, especialmente, porque muchos estudios no proporcionaban información espacial suficientemente detallada sobre la exposición en las distintas regiones cerebrales.

Además, gran parte de los estudios y de los valores atípicos se concentraba en una misma categoría, exposición no intermitente con señales moduladas GSM, de campo cercano y frecuencia inferior a 1 GHz, lo que limitaba la capacidad de esas comparaciones para explicar la heterogeneidad.

### *Sesgo de publicación y análisis de sensibilidad*

El posible sesgo de publicación se evaluó mediante *funnel plots* y, cuando se disponía de al menos diez estudios, mediante la prueba de Egger. a inspección visual de los *funnel plots* no sugirió sesgo de publicación en ninguno de los dominios o subclases. Únicamente para la velocidad en la tarea 1-*back* la prueba de Egger alcanzó el umbral preestablecido de significación (p = 0,071), pero, dada la simetría del *funnel plot*, los autores no consideraron que existiera evidencia de un sesgo de publicación relevante.

Los autores examinaron asimismo resultados atípicos que contribuían de forma importante a la heterogeneidad. En algunos casos, la exclusión de un valor extremo redujo considerablemente $I^2$ y desplazó la estimación agrupada hacia valores próximos a cero. Estas comprobaciones reforzaron la interpretación de que determinadas heterogeneidades estaban impulsadas por estudios individuales más que por un patrón reproducible de efecto de la exposición.

No obstante, los autores mantuvieron una aproximación conservadora y consideraron el posible sesgo de publicación dentro de la evaluación GRADE cuando coincidían la evidencia estadística de asimetría y la inspección visual del *funnel plot*.

### *Evaluación del riesgo de sesgo*

La evaluación OHAT mostró una calidad metodológica heterogénea entre los 76 estudios. Los autores prestaron especial atención al cegamiento y a la evaluación del desenlace, considerados dominios clave para la clasificación global del riesgo de sesgo. Los estudios se distribuyeron entre los tres *tiers* de OHAT, desde estudios con bajo riesgo global hasta trabajos con limitaciones metodológicas importantes.

La aleatorización fue un requisito de inclusión, por lo que este aspecto estaba razonablemente protegido por el propio diseño de la revisión. El cegamiento, sin embargo, fue más variable. Los estudios en los que únicamente los participantes estaban cegados, pero no el personal investigador, fueron clasificados en *tier* 3 debido a la importancia atribuida a este dominio. También se consideraron posibles problemas específicos de los diseños *cross-over*, incluidos efectos de periodo y *carry-over*.

La notificación de los resultados fue desigual. El 42 % de los estudios proporcionó valores numéricos y medidas de dispersión para todos los desenlaces evaluados. Otro 51 % proporcionó datos numéricos sólo para algunos desenlaces, presentó determinados resultados únicamente mediante figuras o comunicó información insuficientemente detallada, como valores de significación estadística sin los datos necesarios para calcular medidas de efecto.

Sólo tres estudios, el 4 %, parecían presentar notificación selectiva de desenlaces y otros dos, el 3 %, no proporcionaron resultados específicos y se limitaron a descripciones verbales generales.

La caracterización de la exposición merece una consideración específica. Aunque todos los estudios incluidos presentaban contraste suficiente entre exposición y *sham* para asegurar que los participantes habían sido realmente expuestos, en algunos estudios existía un riesgo elevado de sesgo en la caracterización de la exposición porque no podían determinarse con precisión su magnitud absoluta o su distribución espacial en la cabeza.

Este problema no invalida necesariamente la comparación experimental exposición-*sham*, pero sí limita considerablemente la posibilidad de establecer relaciones exposición-respuesta o vincular un eventual efecto a la dosis recibida por regiones cerebrales concretas.

### *Certeza de la evidencia (GRADE)*

La valoración GRADE constituye un elemento esencial para interpretar los resultados, porque permite distinguir entre ausencia de una asociación estadísticamente significativa y evidencia con suficiente certeza de que la exposición produce poca o ninguna diferencia. En conjunto, los resultados de los diferentes dominios y subclases proporcionaron evidencia con certeza entre alta y baja de que la exposición a corto plazo a CEM-RF no reduce el rendimiento cognitivo en estudios experimentales humanos.

La certeza no fue uniforme. De las 35 medidas de desenlace evaluadas mediante GRADE, en 16 persistió algún grado de incertidumbre, debido fundamentalmente a limitaciones metodológicas de los estudios, inconsistencia entre resultados o imprecisión de las estimaciones agrupadas.

Por el contrario, para una parte importante de los desenlaces, especialmente aquellos con mayor número de estudios, estimaciones próximas al valor nulo, baja heterogeneidad y riesgo de sesgo más favorable, la evidencia alcanzó niveles moderados o altos.

Esta gradación se refleja en el lenguaje empleado por los propios autores. Cuando la certeza era alta, concluyeron que la exposición produce poca o ninguna diferencia en el rendimiento; con certeza moderada, que probablemente produce poca o ninguna diferencia; y con certeza baja, que puede producir poca o ninguna diferencia.

El caso más incierto fue la velocidad de las habilidades motoras, donde la estimación puntual sugería una posible reducción, pero se basaba en sólo 42 participantes, presentaba un IC95 % extremadamente amplio y una heterogeneidad del 96 %, por lo que no permite establecer de manera fiable ni la existencia ni la magnitud de un efecto.

En conjunto, Pophof et al. consideran que la evidencia experimental disponible no muestra que la exposición a corto plazo a CEM-RF reduzca el rendimiento cognitivo humano. Esta conclusión es particularmente consistente para diversas funciones de orientación y atención, velocidad de procesamiento y memoria de trabajo, para las que existe un cuerpo de evidencia relativamente amplio y varias estimaciones presentan certeza moderada o alta. Sin embargo, la conclusión no debe generalizarse con igual confianza a todas las funciones cognitivas: persiste incertidumbre en 16 de los 35 desenlaces evaluados, la evidencia es muy limitada para construcción y rendimiento motor y funciones verbales y lingüísticas, y no existe evidencia experimental elegible para funciones ejecutivas.

### 4.1.5. Fortalezas y debilidades de la revisión

#### *Fortalezas*

Una de las principales fortalezas de Pophof et al. (2024) es el carácter prospectivo y preespecificado de la revisión. El trabajo se desarrolló a partir de un protocolo previamente publicado y registrado en PROSPERO (CRD42021236168), en el que se definieron la pregunta PECO, los criterios de elegibilidad, la estrategia de búsqueda, la evaluación del riesgo de sesgo y el plan general de síntesis y valoración de la certeza.

Las modificaciones introducidas posteriormente respecto al protocolo fueron descritas y justificadas por los autores, lo que permite distinguir las decisiones metodológicas previstas *a priori* de aquellas adoptadas tras conocer las características de la evidencia disponible.

La búsqueda bibliográfica fue amplia y utilizó cinco bases de datos, PubMed, Embase, Scopus, Web of Science y EMF-Portal, sin restricciones por fecha o idioma, complementándose con la revisión de las referencias de revisiones relevantes. La selección de los estudios se realizó independientemente por dos revisores en las fases de título/resumen y texto completo, con resolución de discrepancias mediante discusión y, cuando era necesario, intervención de un tercer revisor.

Aunque la extracción inicial fue realizada por un único revisor, toda la información fue posteriormente comprobada por un segundo frente a las publicaciones originales.

Además, cuando los autores de la revisión habían participado en algún estudio primario incluido, se abstuvieron de extraer los datos y de evaluar el riesgo de sesgo de sus propios trabajos. Estas medidas reducen el riesgo de errores y de conflictos de interés durante el proceso de revisión.

También constituye una fortaleza importante la exigencia de un diseño experimental capaz de proporcionar una comparación causal razonablemente robusta. Sólo fueron elegibles estudios aleatorizados, de grupos paralelos o *cross-over*, con un contraste suficiente entre condiciones de exposición y con los participantes cegados respecto al nivel aplicado.

Se excluyeron los experimentos claramente no aleatorizados, aquellos en los que los participantes conocían la condición de exposición y los estudios con caracterización insuficiente de los CEM-RF. De este modo, la revisión se concentró deliberadamente en diseños experimentales con mayor capacidad, en principio, para separar un eventual efecto agudo de la exposición de la confusión, el sesgo de selección y las expectativas de los participantes.

La clasificación detallada del rendimiento cognitivo representa otra fortaleza metodológica. En lugar de agrupar todas las pruebas bajo una única categoría de «cognición», los autores organizaron los desenlaces conforme a siete dominios neuropsicológicos y a subclases más específicas, diferenciando además las medidas de velocidad y exactitud.

Esta estrategia reduce el riesgo de combinar indiscriminadamente resultados que representan componentes diferentes del rendimiento cognitivo y permite detectar, al menos en principio, efectos que pudieran limitarse a una determinada función o dimensión. La utilización de la g de Hedges facilita asimismo la combinación de pruebas que evalúan una capacidad semejante mediante escalas diferentes.

La revisión realiza además una evaluación del riesgo de sesgo particularmente detallada. La herramienta OHAT fue adaptada a las características de los estudios experimentales humanos y, específicamente, a los diseños *cross-over*, incorporando cuestiones relativas a efectos de periodo y *carry-over*.

El cegamiento y la evaluación del desenlace fueron considerados dominios especialmente importantes para la clasificación global en *tiers*. Esta aproximación permite que la valoración de la evidencia no dependa únicamente de si un experimento comunica o no una diferencia estadísticamente significativa, sino también de la credibilidad metodológica del resultado.

La estrategia de síntesis incorpora igualmente varios mecanismos para examinar la robustez de los resultados. Los autores cuantificaron la heterogeneidad mediante $I^2$, exploraron posibles fuentes mediante análisis de subgrupos, inspeccionaron *funnel plots* y aplicaron la prueba de Egger cuando el número de estudios lo permitía. También examinaron determinados valores atípicos que podían explicar una parte considerable de la heterogeneidad.

Finalmente, la aplicación de GRADE a cada cuerpo de evidencia permite traducir conjuntamente riesgo de sesgo, inconsistencia, imprecisión y posible sesgo de publicación en un nivel explícito de certeza. Por ello, la conclusión de que la exposición a corto plazo no reduce el rendimiento cognitivo no se basa simplemente en que los metaanálisis no alcancen significación estadística, sino en la evaluación graduada del conjunto de la evidencia.

Por último, la amplitud de la revisión constituye una fortaleza respecto de síntesis anteriores. Se incluyeron 76 estudios con 3.846 participantes de 19 países y se pudieron incorporar 50 estudios con 2.433 participantes a los metaanálisis.

La revisión abarca numerosos componentes del rendimiento cognitivo y representa una actualización sustancial de revisiones previas, algunas de las cuales se habían restringido a poblaciones, dominios cognitivos o tipos de efectos concretos.

## *Limitaciones*

La principal limitación procede de la distribución muy desigual de la evidencia entre funciones cognitivas. Aunque el número total de estudios es relativamente elevado, se concentra de forma clara en determinados componentes de orientación y atención, velocidad de procesamiento y memoria de trabajo. Por el contrario, la evidencia es escasa para funciones verbales y habilidades lingüísticas y para construcción y rendimiento motor, y no se identificó ningún estudio elegible sobre funciones ejecutivas.

Por tanto, la ausencia general de efectos observada en los metaanálisis no puede extrapolarse con el mismo grado de confianza a todo el espectro de la función cognitiva humana. De hecho, 16 de los 35 desenlaces evaluados mediante GRADE conservaron algún grado relevante de incertidumbre por limitaciones metodológicas, inconsistencia o imprecisión.

Existe también una representación insuficiente de determinados grupos poblacionales. La gran mayoría de los experimentos se realizó en adultos, mientras que la evidencia en niños y adolescentes fue mucho menor y la correspondiente a personas mayores especialmente escasa.

Dado que edad y estado de desarrollo pueden influir tanto en el rendimiento cognitivo como en la susceptibilidad a determinados factores ambientales, esta distribución limita la generalización de las conclusiones a todo el ciclo vital. Los propios autores identifican a las personas mayores como una prioridad para futuras investigaciones y recomiendan también una representación más adecuada de ambos sexos.

Otra limitación es el reducido tamaño muestral de numerosos estudios primarios. Aunque el número agregado de participantes es considerable, muchos experimentos individuales incluyeron pocas decenas de sujetos. Los tamaños pequeños disminuyen la precisión de las estimaciones, aumentan la sensibilidad a valores extremos y pueden contribuir a resultados inestables entre experimentos.

Este problema es especialmente evidente en algunos desenlaces con muy pocos participantes, como la velocidad de las habilidades motoras, cuya estimación agrupada se acompañó de un intervalo de confianza extremadamente amplio y una heterogeneidad del 96 %. Por ello, un resultado puntual alejado de cero no puede interpretarse de forma fiable cuando se sustenta en un número tan reducido de observaciones.

La heterogeneidad de las pruebas cognitivas constituye otra dificultad importante. Aunque la clasificación en subclases permitió aumentar la homogeneidad conceptual, dentro de algunas de ellas se combinaron tareas con características diferentes. Los autores reconocen un compromiso inevitable entre utilizar categorías muy específicas, que pueden terminar representadas por pocos estudios o por varios trabajos de un mismo grupo de investigación, y utilizar categorías más amplias, que permiten incrementar la potencia estadística, pero introducen mayor heterogeneidad entre las tareas.

Esta diversidad contribuyó probablemente a los valores elevados de $I^2$ observados en algunos análisis, como la exactitud en memoria verbal y visual, procedimientos matemáticos o determinadas medidas atencionales.

La caracterización de la exposición es una limitación particularmente relevante desde el punto de vista biofísico. Todos los experimentos incluidos proporcionaban un contraste suficiente entre exposición y *sham* para estudiar si existía un efecto agudo, pero en una parte de ellos no era posible determinar con precisión la magnitud absoluta de la exposición y, especialmente, su distribución espacial dentro de la cabeza. Esto es importante porque, si existiese un efecto de los CEM-RF sobre una función cognitiva concreta, podría depender no sólo de la SAR global o máxima, sino de la energía absorbida en las regiones cerebrales implicadas en la tarea. Esta deficiencia limitó de forma importante la interpretación dosimétrica de los resultados.

Precisamente por estas limitaciones dosimétricas, la revisión no pudo realizar el análisis exposición-respuesta inicialmente previsto. Tampoco se efectuaron finalmente algunos análisis de subgrupos planificados según frecuencia, modulación, duración de la exposición o campo cercano frente a campo lejano.

Los autores explican que una gran parte de los estudios y varios valores atípicos se concentraban en una misma configuración, fundamentalmente señales moduladas GSM, exposición de campo cercano, frecuencia inferior a 1 GHz y exposición de la cabeza. En estas condiciones, comparar categorías de exposición habría tenido una capacidad limitada para explicar la heterogeneidad. La ausencia de un análisis dosis-respuesta impide determinar si los tamaños de efecto varían sistemáticamente con la intensidad de la exposición.

La concentración de la evidencia en determinadas tecnologías y escenarios de exposición limita también su aplicabilidad. La mayoría de los estudios utilizó exposición localizada de la cabeza mediante GSM de 900 MHz en *uplink*. Por tanto, aunque la revisión comprende formalmente el intervalo de RF definido en el PECO, su evidencia empírica no representa de forma uniforme todas las frecuencias, modulaciones y configuraciones actualmente relevantes. Las conclusiones son más directamente informativas para las condiciones experimentales realmente estudiadas que para cualquier posible exposición a CEM-RF.

Los diseños *cross-over*, predominantes en esta literatura, proporcionan ventajas importantes porque cada participante puede actuar como su propio control, pero presentan problemas específicos si no se gestionan adecuadamente los efectos de periodo, aprendizaje y *carry-over*.

La revisión adaptó la herramienta OHAT para valorar estos aspectos, pero la información proporcionada en las publicaciones primarias fue variable. También existió heterogeneidad en el entrenamiento previo de los participantes, el contrabalanceo del orden de las tareas y las condiciones, la separación entre sesiones y la duración de las exposiciones. Estas diferencias pueden introducir variabilidad independiente de la exposición a CEM-RF y dificultar la comparación entre experimentos.

Asimismo, aunque la aleatorización y el cegamiento de los participantes fueron criterios de inclusión, el cegamiento del personal investigador no fue universal. Los autores consideraron este aspecto suficientemente importante como para clasificar en *tier* 3 los estudios en los que únicamente los participantes estaban cegados.

La documentación de los resultados tampoco fue completa en todos los trabajos: parte de los estudios no proporcionaba valores numéricos y medidas de dispersión para todos los desenlaces, lo que explica que sólo 50 de los 76 estudios pudieran incorporarse a los metaanálisis.

Esta pérdida de información reduce la capacidad de síntesis cuantitativa y puede introducir incertidumbre adicional.

Finalmente, la propia naturaleza del diseño experimental determina el alcance de las conclusiones. Los estudios incluidos evalúan efectos agudos o a corto plazo de exposiciones controladas. Su elevada validez interna para esta pregunta no permite inferir directamente qué ocurre después de exposiciones repetidas o mantenidas durante meses o años. Por ello, la evidencia de Pophof et al. (2024) debe interpretarse como complementaria, y no sustitutiva, de la revisión observacional de efectos a más largo plazo. La conclusión de que la exposición a corto plazo a CEM-RF no reduce el rendimiento cognitivo, con certeza que varía entre alta y baja según el desenlace, no equivale a demostrar ausencia de cualquier posible efecto cognitivo bajo cualquier patrón temporal o condición de exposición.

En conjunto, Pophof et al. (2024) proporcionan una revisión amplia y metodológicamente estructurada de la evidencia experimental humana, con criterios de inclusión exigentes, evaluación detallada del riesgo de sesgo, síntesis separada por componentes cognitivos y valoración formal de la certeza. Sus principales limitaciones derivan menos de la estrategia general de la revisión que de la evidencia primaria disponible: muestras pequeñas en numerosos estudios, cobertura desigual de dominios y poblaciones, heterogeneidad de las pruebas, información dosimétrica espacial insuficiente y predominio de un conjunto relativamente restringido de configuraciones de exposición. Estas limitaciones son especialmente importantes para definir hasta dónde puede extenderse la conclusión global de ausencia de deterioro agudo y qué preguntas permanecen abiertas.

### 4.1.6. Conclusiones

Pophof et al. (2024) concluyen que el conjunto de la evidencia experimental disponible indica que la exposición a corto plazo a CEM-RF, a niveles de SAR dentro de los límites recomendados, no afecta negativamente al rendimiento cognitivo en los dominios estudiados. Ninguno de los metaanálisis realizados mostró un efecto estadísticamente significativo de la exposición frente a *sham*, tanto para las medidas relacionadas con la velocidad de procesamiento como para las relacionadas con la exactitud. La certeza de esta conclusión varió entre alta y baja según el dominio o subclase cognitiva analizada, siendo predominantemente moderada o alta.

Esta conclusión debe interpretarse atendiendo a la estructura de la evidencia disponible. os autores evaluaron separadamente múltiples componentes de orientación y atención, percepción, memoria, funciones verbales y habilidades lingüísticas, construcción y rendimiento motor y formación de conceptos y razonamiento.

Sin embargo, la cantidad y calidad de la información no fueron uniformes entre ellos y no se identificaron estudios elegibles para el dominio de funciones ejecutivas. Por tanto, la conclusión general no implica que todos los componentes posibles de la cognición hayan sido investigados con el mismo grado de profundidad y certeza.

De los 35 desenlaces o subdominios evaluados mediante GRADE, en 16 persistió algún grado de incertidumbre debido principalmente a limitaciones metodológicas de los estudios, inconsistencia entre resultados o imprecisión de las estimaciones combinadas.

Este aspecto es relevante para interpretar correctamente la ausencia de diferencias estadísticamente significativas: en aquellos desenlaces con evidencia de alta o moderada certeza, los resultados proporcionan un respaldo relativamente sólido a la ausencia de un deterioro cognitivo agudo; allí donde la certeza es baja, o muy baja en el caso excepcional de la velocidad de las habilidades motoras, la interpretación debe ser necesariamente más prudente.

La principal excepción que requiere investigación adicional corresponde a la velocidad en el dominio de construcción y rendimiento motor. En este caso, el cuerpo de evidencia procedía únicamente de dos estudios pequeños, con un total de 42 participantes, presentaba una heterogeneidad extrema ($I^2$ = 96 %) y sugería un efecto negativo de gran magnitud, pero con certeza muy baja.

Los autores consideran que la información disponible es insuficiente para determinar si este resultado representa un efecto real o deriva de la escasez, heterogeneidad e imprecisión de los datos. Por ello, este hallazgo no contradice la conclusión general de la revisión, pero tampoco puede descartarse sobre la base de la evidencia actualmente disponible.

Desde la perspectiva de evaluación del riesgo, la revisión proporciona, por tanto, evidencia predominantemente de certeza moderada a alta de que las exposiciones agudas a CEM-RF en las condiciones experimentales estudiadas y a niveles de SAR dentro de los límites recomendados no producen un deterioro detectable del rendimiento cognitivo. Esta formulación es importante porque delimita correctamente el alcance de los resultados: se refiere a exposiciones de corta duración y a los dominios y condiciones efectivamente investigados, y no permite extrapolar directamente las conclusiones a posibles efectos derivados de exposiciones mantenidas durante meses o años.

### *Implicaciones para investigación futura*

Pophof et al. (2024) no plantean de forma genérica la necesidad de repetir los numerosos experimentos de provocación ya realizados, sino que identifican lagunas concretas de conocimiento en las que nuevos estudios pueden aportar información relevante. La prioridad más clara corresponde a construcción y rendimiento motor, especialmente a sus medidas de velocidad.

Dado que el resultado actualmente disponible se basa únicamente en dos estudios pequeños y heterogéneos, los autores consideran necesarios nuevos experimentos que permitan confirmar o refutar el posible efecto observado. Recomiendan además que estos trabajos comuniquen numéricamente todos los desenlaces evaluados, evitando una notificación incompleta que dificulte posteriores síntesis cuantitativas.

Otra prioridad es mejorar la representación de los grupos de edad insuficientemente estudiados. Sólo cinco de los estudios incluidos se realizaron en niños y únicamente uno en personas mayores.

Los autores consideran pertinente ampliar la investigación en niños y adolescentes y, particularmente, en personas de edad avanzada, dado que el rendimiento cognitivo cambia a lo largo de la vida y estos grupos podrían responder de manera diferente a la exposición. Los estudios futuros deberían permitir determinar específicamente si estos grupos presentan una susceptibilidad distinta de la observada en adultos jóvenes.

También se recomienda prestar mayor atención al sexo de los participantes. Los análisis de subgrupos proporcionaron indicios débiles de posibles diferencias según sexo, insuficientes para establecer conclusiones, pero suficientes para justificar que los futuros experimentos incluyan adecuadamente a ambos sexos y presenten análisis estratificados cuando el tamaño muestral lo permita.

Esta recomendación debe entenderse como una necesidad de resolver una incertidumbre existente, no como evidencia de que actualmente se haya demostrado una respuesta diferencial entre hombres y mujeres.

Los autores subrayan asimismo la necesidad de mejorar la potencia estadística de los nuevos experimentos. Muchos de los estudios disponibles presentan muestras pequeñas, por lo que recomiendan realizar cálculos de potencia antes de establecer el diseño definitivo y utilizar tamaños muestrales suficientes para detectar incluso efectos de pequeña magnitud.

Los nuevos estudios deberían diseñarse además para minimizar sistemáticamente el riesgo de sesgo, atendiendo a los criterios establecidos en el protocolo de la revisión para alcanzar una clasificación de riesgo definitivamente bajo.

Esto implica una adecuada aleatorización, cegamiento, control de los problemas específicos de los diseños *cross-over*, evaluación objetiva de los desenlaces y descripción suficientemente detallada de los procedimientos experimentales. En este contexto, los autores remiten también a las recomendaciones metodológicas formuladas específicamente para la investigación de rendimiento cognitivo en bioelectromagnetismo.

Un aspecto adicional que se desprende de la revisión es la conveniencia de mejorar la caracterización dosimétrica de la exposición. La información disponible no permitió realizar el análisis exposición-respuesta inicialmente previsto ni determinar de forma suficientemente homogénea la distribución espacial de la energía absorbida en el cerebro.

Puesto que un eventual efecto sobre una función cognitiva específica podría depender de la exposición de las regiones cerebrales implicadas y no únicamente de una medida global de SAR, una dosimetría más completa y comparable entre estudios aumentaría la capacidad interpretativa de futuros experimentos. La concentración de muchos estudios en determinadas configuraciones de exposición constituye, además, una limitación para evaluar si características como la frecuencia, la modulación o la configuración de la exposición modifican los resultados.

En conjunto, las implicaciones para la investigación futura apuntan menos a la mera acumulación de nuevos estudios de provocación similares a los ya existentes que a cubrir las lagunas concretas identificadas y aumentar la calidad informativa de cada experimento: estudiar los dominios cognitivos insuficientemente representados, particularmente construcción y rendimiento motor; ampliar la evidencia en niños, adolescentes y personas mayores; representar adecuadamente ambos sexos; aumentar el tamaño y la potencia de las muestras; minimizar el riesgo de sesgo; comunicar de forma completa los resultados y mejorar la caracterización dosimétrica.

Con la evidencia actualmente disponible, Pophof et al. (2024) consideran que existe un respaldo predominantemente de certeza moderada a alta para concluir que la exposición a corto plazo a CEM-RF, a niveles de SAR dentro de los límites recomendados, no deteriora los dominios de rendimiento cognitivo investigados; las prioridades de investigación deben dirigirse, por tanto, hacia las áreas en las que esa conclusión todavía no puede establecerse con suficiente certeza.

## 4.2. Revisión de estudios observacionales

### 4.2.1. Referencia completa del artículo

**Título:** The effects of radiofrequency exposure on cognition: A systematic review and meta-analysis of human observational studies. Benke G., Abramson M.J., Brzozek C., McDonald S., Kelsall H., Sanagou M., Zeleke B.M., Kaufman J., Brennan S., Verbeek J., Karipidis K. **Revista y datos bibliográficos:** Environment International, 188, 108779 (2024). DOI: 10.1016/j.envint.2024.108779. **Registro de protocolo:** PROSPERO CRD42021257548.

### 4.2.2. Objetivos y preguntas de investigación

El objetivo principal de Benke et al. (2024) fue evaluar los posibles efectos a largo plazo de la exposición a CEM-RF sobre la función cognitiva en estudios observacionales humanos. Los autores definieron expresamente la exposición a largo plazo como aquella mantenida durante meses o años, planteamiento que complementa la revisión experimental de Pophof et al. (2024), centrada en los posibles efectos agudos o a corto plazo de exposiciones controladas. La revisión pretendía determinar si una mayor exposición a CEM-RF se relaciona con diferencias en la función cognitiva respecto de niveles de exposición menores o nulos en la población general y trabajadora. Como objetivo secundario, se planteó específicamente investigar la existencia de una relación exposición-respuesta entre los niveles de exposición a CEM-RF y la función cognitiva en hombres, mujeres y niños.

La pregunta de investigación se estructuró mediante el marco PECO. La población (P) comprendió personas de cualquier edad y sexo, tanto de la población general como de poblaciones ocupacionalmente expuestas. La exposición (E) incluyó CEM-RF procedentes de fuentes personales o de campo cercano, ambientales o de campo lejano y ocupacionales, estas últimas potencialmente de campo cercano o lejano.

Para ser elegibles, los estudios debían comparar al menos dos niveles diferentes de intensidad, frecuencia o duración de la exposición, incluyendo como referencia un grupo no expuesto o con un nivel de exposición inferior. El comparador (C) fue, por tanto, la ausencia de exposición o una exposición de menor nivel que la del grupo de interés. Además, para preservar la temporalidad necesaria para estudiar posibles efectos a largo plazo, se exigió un seguimiento mínimo de seis meses desde la evaluación basal.

La exposición se concibió idealmente en términos de dosis total de energía absorbida en el órgano o tejido relevante, dependiente de la tasa de absorción específica (SAR), de la duración y de las características de la exposición. Sin embargo, los autores reconocen que esta magnitud resulta difícil de medir o modelizar adecuadamente en estudios observacionales.

Por ello, admitieron diferentes aproximaciones indirectas, entre ellas intensidad del campo eléctrico en V/m, densidad de potencia en W/m², uso autoinformado de teléfonos móviles o inalámbricos e información ocupacional. Para las exposiciones laborales se contemplaron mediciones de CEM-RF, puesto o tarea de trabajo, matrices de exposición laboral o de fuentes y evaluación por expertos.

En cambio, se excluyeron métricas consideradas sustitutos inadecuados de la exposición, como la distancia autoinformada a antenas o la distancia calculada a estaciones base de telefonía móvil, aunque sí se admitió la distancia calculada a transmisores de radiodifusión situados en una localización única.

El desenlace (O) fue la función o rendimiento cognitivo. Los autores distinguieron entre función cognitiva global y funciones específicas por dominio. Estas últimas comprendieron atención compleja, función ejecutiva, aprendizaje y memoria, lenguaje, habilidad perceptivo-motora y cognición social.

Se aceptaron diferentes definiciones e instrumentos debido a la heterogeneidad esperable entre estudios, siempre que las pruebas o criterios utilizados contaran con evidencia de validez y fiabilidad para evaluar la función cognitiva y no estuvieran orientados específicamente a trastornos neurodegenerativos. Los resultados podían expresarse mediante una puntuación global que integrase diferentes capacidades cognitivas, mediante puntuaciones correspondientes a dominios específicos o mediante ambas modalidades.

La revisión se restringió a estudios observacionales longitudinales de cohortes o de casos y controles. Se excluyeron expresamente los estudios transversales, dado que la evaluación simultánea de exposición y desenlace impide establecer adecuadamente su secuencia temporal y dificulta la inferencia sobre posibles efectos causales.

También se excluyeron los estudios sobre exposición materna durante el embarazo y posteriores efectos sobre el neurodesarrollo infantil, los centrados en trastornos del neurodesarrollo o enfermedades neurodegenerativas, aquellos con menos de seis meses de seguimiento y los que carecían de un grupo comparador con menor exposición. De este modo, la revisión se orientó específicamente a cambios de la función cognitiva potencialmente relacionados con una exposición prolongada, y no a enfermedades neurológicas o alteraciones del neurodesarrollo consideradas como entidades clínicas.

Un aspecto particularmente relevante de la pregunta de investigación fue la consideración explícita de la relación exposición-respuesta. Los autores no pretendían limitarse a comparar grupos definidos como expuestos y no expuestos, sino evaluar, cuando los datos lo permitieran, si los cambios en la función cognitiva variaban con niveles crecientes de exposición a CEM-RF.

Esta cuestión está estrechamente condicionada por la calidad de la evaluación de la exposición, ya que las medidas disponibles en estudios observacionales no necesariamente representan la energía realmente absorbida por las regiones cerebrales implicadas en cada función cognitiva.

De hecho, los propios autores señalan que una evaluación de la exposición que no sea específica de las regiones cerebrales relacionadas con cada dominio limita la interpretación de los resultados a escala regional. Como veremos posteriormente, tanto la escasez de estudios elegibles como las limitaciones de las métricas de exposición condicionaron de manera importante la posibilidad de establecer relaciones exposición-respuesta fiables.

## 4.2.3. Metodología

La revisión se desarrolló a partir de un protocolo metodológico previamente establecido y publicado por Benke et al. (2022), y fue registrada prospectivamente en PROSPERO (CRD42021257548).

La revisión se informó de acuerdo con las directrices PRISMA. El protocolo definió previamente la pregunta PECO, los criterios de elegibilidad, las fuentes y estrategia de búsqueda, los procedimientos de selección y extracción de datos, la evaluación del riesgo de sesgo, las medidas de efecto, el plan de síntesis y análisis exposición-respuesta y la valoración de la certeza de la evidencia. El planteamiento metodológico estuvo específicamente diseñado para evaluar posibles efectos a largo plazo de los CEM-RF sobre la función cognitiva mediante estudios observacionales longitudinales. Los autores documentaron asimismo las modificaciones introducidas respecto al protocolo, entre ellas determinadas decisiones relativas al modelo de metaanálisis finalmente utilizado y a la operacionalización de la clasificación por *tiers* del riesgo de sesgo.

### *Búsqueda y fuentes*

La búsqueda bibliográfica se realizó en cuatro bases de datos: PubMed, Embase, PsycInfo y EMF-Portal, esta última especializada en literatura científica sobre campos electromagnéticos. Las búsquedas se efectuaron inicialmente el 22 de septiembre de 2021 y se repitieron el 30 de septiembre de 2022, sin restricciones por fecha ni idioma de publicación. La estrategia desarrollada para PubMed combinó términos MeSH y términos de texto libre relacionados con la exposición, los desenlaces cognitivos y los estudios observacionales humanos. Posteriormente se adaptó a Embase y PsycInfo incorporando los correspondientes términos de los tesauros específicos de estas bases de datos.

La búsqueda electrónica se complementó mediante el examen de siete revisiones de la literatura publicadas entre 2004 y 2015 por organismos nacionales e internacionales relevantes que habían evaluado los efectos sobre la salud de la exposición a CEM-RF. Los estudios relativos a efectos cognitivos identificados en estas revisiones se extrajeron y sometieron a los mismos criterios de elegibilidad que los registros recuperados mediante las bases bibliográficas. Este procedimiento pretendía aumentar la sensibilidad de la búsqueda y reducir la posibilidad de omitir estudios epidemiológicos relevantes.

### *Criterios de elegibilidad*

Se incluyeron estudios observacionales longitudinales de cohortes o de casos y controles realizados en población general o trabajadora, sin restricciones por edad o sexo. Los estudios debían evaluar exposición personal o de campo cercano, ambiental o de campo lejano, u ocupacional, que podía corresponder a cualquiera de las dos configuraciones, y comparar al menos dos niveles diferentes de intensidad, frecuencia o duración de la exposición, incluyendo un grupo no expuesto o con una exposición inferior. Para garantizar una temporalidad mínima entre exposición y desenlace, se exigió un seguimiento de al menos seis meses desde la evaluación basal.

Se excluyeron los diseños distintos de cohortes o casos y controles y, específicamente, los estudios transversales, dado que la ausencia de una secuencia temporal clara entre exposición y desenlace dificulta establecer posibles efectos causales.

También se excluyeron los estudios sobre exposición materna a CEM-RF durante el embarazo y efectos posteriores sobre el neurodesarrollo infantil, los centrados en trastornos del neurodesarrollo o enfermedades neurodegenerativas, aquellos con un seguimiento inferior a seis meses y los que carecían de un comparador con un nivel inferior de exposición.

La métrica ideal de exposición se definió como la dosis total de energía absorbida en el órgano o tejido relevante, determinada por la tasa de absorción específica (SAR), la duración y las características de la exposición. Los autores reconocieron, sin embargo, que esta magnitud difícilmente puede medirse o modelizarse adecuadamente en estudios observacionales.

Por ello se admitieron aproximaciones indirectas, entre ellas estimaciones de intensidad de campo eléctrico en V/m o densidad de potencia en W/m², uso autoinformado de teléfonos móviles o inalámbricos e información ocupacional. Para las exposiciones laborales se aceptaron mediciones de CEM-RF, puesto o tarea de trabajo, matrices de exposición laboral o de fuentes y evaluación por expertos. Se excluyeron la distancia autoinformada a antenas y la distancia calculada a estaciones base de telefonía móvil por considerarse *proxies* inadecuados de exposición, aunque se admitió la distancia calculada a transmisores de radiodifusión situados en una localización única.

Los estudios debían proporcionar al menos una medida de función o rendimiento cognitivo. Los desenlaces primarios comprendieron la función cognitiva global y diferentes dominios específicos: atención compleja, función ejecutiva, aprendizaje y memoria, lenguaje, habilidad perceptivo-motora y cognición social.

Dada la diversidad esperable de instrumentos, se admitieron distintas definiciones y medidas, siempre que las pruebas o criterios empleados dispusieran de evidencia de validez y fiabilidad para evaluar la función cognitiva y no estuvieran específicamente relacionados con enfermedades neurodegenerativas. Los resultados podían expresarse mediante una puntuación cognitiva global, puntuaciones de dominios específicos o ambas.

### *Selección y extracción de datos*

Los registros recuperados se gestionaron mediante Covidence y la selección de los estudios se realizó por duplicado. Al menos dos revisores examinaron independientemente los títulos y resúmenes recuperados y los potencialmente elegibles pasaron posteriormente a evaluación del texto completo, también por dos revisores independientes.

Los desacuerdos se resolvieron mediante consenso entre los miembros del equipo. Ningún miembro del equipo participó en el cribado o la extracción de datos de un estudio del que fuera autor.

La extracción de datos se realizó mediante un formulario previamente pilotado e incluyó los identificadores y características de los estudios, las características del diseño, de los grupos de exposición y comparación y de los participantes, así como los resultados necesarios para los análisis.

Para los modelos multivariables se priorizaron las estimaciones con el ajuste más completo, excepto cuando el modelo incorporaba una variable que pudiera actuar como intermediaria en la vía causal. También se recopiló la información necesaria para la evaluación posterior del riesgo de sesgo.

### *Medidas de efecto y síntesis cuantitativa*

Las medidas de efecto se seleccionaron en función de la forma en que se habían evaluado los desenlaces cognitivos. Cuando los estudios utilizaban escalas continuas comparables, se contempló la diferencia de medias (MD); cuando un mismo constructo cognitivo se medía mediante escalas diferentes, se utilizó la diferencia de medias estandarizada (SMD).

Para desenlaces ordinales dicotomizados, el protocolo había previsto transformar las *odds ratios* (OR) en SMD, pero esta transformación no se realizó finalmente porque sólo un estudio comunicó resultados mediante OR.

Para interpretar la magnitud de MD y SMD se utilizaron como referencia las convenciones de Cohen: valores comprendidos aproximadamente entre −0,2 y 0,2 se consideraron compatibles con un efecto trivial o poca o ninguna diferencia, mientras que 0,2, 0,5 y 0,8 se utilizaron como referencias para efectos pequeños, moderados y grandes, respectivamente.

Para las OR se consideró potencialmente importante un efecto relativo superior a 1,25 o inferior a 0,80, interpretación que se contrastó además mediante la diferencia absoluta de riesgos, considerando un 5 % como umbral de importancia.

La estrategia de síntesis contempló dos aproximaciones complementarias. En primer lugar, se comparó el grupo con mayor exposición con el grupo de menor exposición. Posteriormente se evaluó el cambio incremental del riesgo o del rendimiento cognitivo asociado a un incremento unitario de la exposición, con el objetivo de investigar posibles relaciones exposición-respuesta.

El protocolo preveía realizar metaanálisis de efectos aleatorios cuando existieran al menos dos estudios suficientemente comparables para una combinación concreta de exposición y desenlace.

Sin embargo, la evidencia finalmente disponible fue mucho más escasa de lo previsto y los únicos dos estudios combinables procedían del mismo grupo de investigación. Siguiendo las recomendaciones metodológicas para metaanálisis con muy pocos estudios, los autores optaron por utilizar un modelo de efectos fijos. Los análisis se realizaron mediante StataSE, versión 18, y las estimaciones agrupadas se representaron mediante *forest plots*. Los estudios de niños o adolescentes se consideraron separadamente de los realizados en adultos.

La heterogeneidad se examinó mediante inspección visual de las curvas exposición-respuesta específicas de los estudios. Además, se utilizó la prueba $\chi^2$, con un nivel de significación de $\alpha = 0,1$, y se cuantificó la inconsistencia mediante $I^2$.

La escasez del cuerpo de evidencia limitó de forma importante los análisis adicionales y la capacidad de investigar modificadores del efecto previstos inicialmente en el protocolo.

### *Evaluación del riesgo de sesgo*

El riesgo de sesgo se evaluó mediante la herramienta *Risk of Bias Rating Tool for Human and Animal Studies* de la Office of Health Assessment and Translation del National Toxicology Program (NTP/OHAT), adaptada a las características de los estudios observacionales incluidos.

La evaluación consideró sesgo de selección, confusión y variables modificadoras, pérdidas o exclusiones de participantes, sesgo de detección en la caracterización de la exposición, sesgo de detección en la evaluación de los desenlaces y notificación selectiva de resultados, además de otras amenazas potenciales para la validez.

Cada dominio se clasificó como riesgo definitivamente bajo, probablemente bajo, probablemente alto o definitivamente alto. A partir de estas valoraciones, los estudios se clasificaron también por *tiers* de calidad, desde *tier* 1, correspondiente a menor preocupación por riesgo de sesgo, hasta *tier* 3, correspondiente a mayor preocupación.

La evaluación de la confusión recibió especial atención debido a la naturaleza observacional de la evidencia. Los autores consideraron la posibilidad de que características sociodemográficas, educativas, conductuales y otras exposiciones pudieran relacionarse simultáneamente con el uso de dispositivos emisores de CEM-RF y con el rendimiento cognitivo.

Del mismo modo, la caracterización de la exposición constituyó un dominio metodológicamente crítico, dado que gran parte de la evidencia dependía de aproximaciones indirectas de la dosis realmente absorbida y, por tanto, estaba sujeta a posible clasificación errónea de la exposición.

El protocolo contemplaba también la evaluación del sesgo debido a resultados ausentes o sesgo de publicación siguiendo el marco del *Cochrane Handbook*. Sin embargo, esta evaluación no pudo realizarse de forma completa. No se disponía de la información necesaria, como registros o protocolos suficientemente detallados, para determinar si existían desenlaces recogidos pero no publicados, y el número de estudios fue demasiado reducido para elaborar *funnel plots* o aplicar pruebas estadísticas de asimetría.

### *Certeza de la evidencia*

La certeza del cuerpo de evidencia se evaluó mediante GRADE, incorporando las modificaciones propuestas por OHAT para estudios de exposiciones ambientales. La valoración se realizó para los resultados de los análisis exposición-respuesta y consideró los dominios de riesgo de sesgo, inconsistencia no explicada, imprecisión, indirectitud y posible sesgo de publicación o notificación.

El enfoque contempló asimismo los criterios que pueden aumentar la confianza en evidencia observacional, como una magnitud elevada del efecto o la existencia de sesgos plausibles que tenderían a desplazar la estimación hacia el valor nulo.

Siguiendo las reglas de decisión de OHAT, los cuerpos de evidencia observacional partieron de una certeza inicial baja o moderada. La calificación inicial fue moderada cuando se cumplían tres características clave del diseño: que la evaluación de la exposición representase exposiciones anteriores al desarrollo del desenlace, que el desenlace se evaluase a nivel individual y que existiera un grupo comparador apropiado.

La certeza final dependió, por tanto, no sólo de la magnitud y precisión de las estimaciones, sino también de la calidad de la caracterización de la exposición, el control de la confusión, la consistencia entre estudios, la aplicabilidad directa de la evidencia a la pregunta PECO y la posibilidad de sesgos de publicación o notificación.

La escasez de estudios elegibles y la imposibilidad de combinar cuantitativamente buena parte de los resultados fueron, como veremos, determinantes para la certeza finalmente asignada a los diferentes desenlaces cognitivos. Los juicios sobre la certeza y las estimaciones correspondientes se sintetizaron mediante tablas *Summary of Findings* y perfiles de evidencia para los distintos desenlaces evaluados.

## 4.2.4. Resultados

### *Flujo PRISMA*

La búsqueda bibliográfica identificó 3.945 registros únicos, de los cuales 3.625 fueron excluidos durante el cribado por título y resumen. Se evaluaron posteriormente 320 publicaciones a texto completo y, finalmente, cinco estudios, publicados en diez informes, cumplieron los criterios de inclusión de la revisión. Los cinco estudios correspondían a cuatro cohortes y comprendían en conjunto 4.639 participantes, 1.831 niños o adolescentes y 2.808 adultos.

La diferencia entre el número de estudios y el de cohortes se debe a que dos de los estudios incluidos, Roser et al. (2016) y Foerster et al. (2018), utilizaron participantes de la cohorte suiza HERMES. Roser et al. fue mantenido como estudio independiente porque evaluó capacidad de concentración, mientras que Foerster et al., que incorporó además una segunda ola de participantes, analizó memoria verbal y figural. Los autores evitaron así contabilizar como evidencia independiente resultados redundantes procedentes de la misma cohorte para un mismo desenlace.

### *Características de los estudios incluidos*

Los cinco estudios incluidos fueron cohortes prospectivas. Cuatro se realizaron en niños o adolescentes de Australia y Suiza, con un seguimiento aproximado de un año, mientras que el quinto correspondió a adultos de 55 años o más pertenecientes al Singapore Longitudinal Ageing Studies (SLAS), con un seguimiento medio de 3,88 años. Las cohortes representadas fueron HERMES en Suiza, MoRPhEUS y ExPOSURE en Australia y SLAS en Singapur.

En HERMES, Roser et al. (2016) incluyeron 493 adolescentes en situación basal, mientras que Foerster et al. (2018), tras incorporar una segunda ola de participantes, incluyeron 895. MoRPhEUS comprendió 317 adolescentes y ExPOSURE 619 niños. La cohorte SLAS comprendía 2.808 adultos mayores, de los cuales 871 aportaron datos basales y de seguimiento, y el número efectivamente disponible fue todavía menor para determinados análisis cognitivos.

Todos los estudios investigaron exposición personal a CEM-RF, fundamentalmente asociada al uso de teléfonos móviles y, en algunos casos, teléfonos inalámbricos. Thomas et al. (2010), Bhatt et al. (2017) y Ng et al. utilizaron principalmente información autodeclarada sobre el uso del teléfono.

Los estudios HERMES incorporaron una caracterización más elaborada de la exposición, combinando cuestionarios con registros de operadores y estimaciones de dosis cerebral acumulada de CEM-RF. Roser et al. (2016) evaluó además exposición ambiental mediante mediciones personales. No se identificó ningún estudio elegible sobre exposición ocupacional.

Los principales dominios cognitivos estudiados fueron atención compleja, función ejecutiva y aprendizaje y memoria. La función cognitiva global se evaluó únicamente en adultos mayores. La cobertura de otros dominios fue muy limitada y no existió evidencia susceptible de evaluación para lenguaje o cognición social. Esta distribución desigual de los desenlaces constituye un aspecto importante para interpretar los resultados, ya que las conclusiones de la revisión no pueden extrapolarse al conjunto de las funciones cognitivas.

### *Exposición personal y función cognitiva en niños y adolescentes*

La mayor parte de la evidencia disponible correspondió a exposición personal asociada al uso de teléfonos móviles en niños y adolescentes. Para algunos desenlaces fue posible combinar cuantitativamente los resultados de MoRPhEUS y ExPOSURE, comparando participantes cuyo número de llamadas semanales aumentó durante el seguimiento con aquellos en los que disminuyó o permaneció estable.

En atención compleja, el metaanálisis de dos estudios con 615 participantes no mostró diferencias relevantes en las tareas de detección e identificación. Para la exactitud en la tarea de detección, la MD fue 0,02 (IC95 % −0,04 a 0,08; $I^2$ = 0 %), y para la exactitud en identificación, 0,00 (IC95 % −0,04 a 0,05; $I^2$ = 0 %). Para el tiempo de respuesta, la MD fue 0,02 (IC95 % 0,01 a 0,03; $I^2$ = 63 %) en detección y 0,00 (IC95 % −0,01 a 0,02; $I^2$ = 8 %) en identificación.

Aunque el intervalo de confianza de la estimación correspondiente al tiempo de respuesta en la tarea de detección no incluyó el valor nulo, la magnitud del efecto fue considerada trivial y los autores interpretaron globalmente los resultados como indicativos de poca o ninguna diferencia en atención compleja.

Roser et al. (2016), que empleó el FAKT-II y no pudo incorporarse al metaanálisis por utilizar una prueba diferente, tampoco proporcionó evidencia clara de un efecto. La SMD para la exactitud de concentración por incremento de un rango intercuartílico en los minutos diarios autodeclarados de uso del teléfono móvil fue −0,40 (IC95 % −0,95 a 0,16; 290 participantes), mientras que para la dosis cerebral acumulada de CEM-RF fue −0,13 (IC95 % −0,68 a 0,42). Ambos intervalos fueron amplios e incluyeron el valor nulo.

Para función ejecutiva fue posible combinar dos estudios, con 612 participantes, utilizando medidas derivadas de la prueba de Stroop. La MD para el cociente temporal (B-A)/A fue 0,02 (IC95 % −0,01 a 0,04; $I^2$ = 88 %), mientras que para (D-C)/C fue 0,00 (IC95 % −0,06 a 0,05; $I^2$ = 59 %).

El primer análisis mostró una inconsistencia importante entre los dos estudios, con asociaciones en direcciones opuestas y ausencia de solapamiento entre sus intervalos de confianza, mientras que para la segunda medida la inconsistencia fue considerada menos relevante. En conjunto, los autores interpretaron estos resultados como indicativos de poca o ninguna diferencia en función ejecutiva, aunque con certeza muy baja.

Para aprendizaje y memoria, el metaanálisis de los mismos dos estudios y 615 participantes mostró igualmente efectos pequeños. En la tarea *one-back*, la MD fue −0,03 (IC95 % −0,07 a 0,02; $I^2$ = 66 %) para exactitud y −0,01 (IC95 % −0,04 a 0,03; $I^2$ = 70 %) para tiempo de respuesta. En la tarea *one-card learning*, la MD fue −0,02 (IC95 % −0,04 a 0,00; $I^2$ = 0 %) para exactitud y −0,01 (IC95 % −0,04 a 0,03; $I^2$ = 70 %) para tiempo de respuesta. Los autores consideraron que estas magnitudes eran compatibles con poca o ninguna diferencia en aprendizaje y memoria.

Foerster et al. (2018), no incluido en estos metaanálisis, analizó memoria verbal y figural mediante el *Intelligenz-Struktur-Test* y diferentes métricas de exposición. Para un incremento de un rango intercuartílico en los minutos diarios autodeclarados de uso del teléfono móvil, la SMD fue −0,01 (IC95 % −0,29 a 0,27; 676 participantes) para memoria verbal y −0,21 (IC95 % −0,51 a 0,09; 670 participantes) para memoria figural. No obstante, el estudio encontró disminuciones estadísticamente significativas de la memoria figural asociadas al aumento de llamadas con teléfono inalámbrico y a una mayor dosis cerebral acumulada estimada de CEM-RF.

Estas asociaciones no se reprodujeron para el uso autodeclarado del teléfono móvil y no se observaron efectos sobre memoria verbal con ninguna de las métricas de exposición, por lo que los resultados no configuraron un patrón uniforme entre las diferentes medidas de exposición.

### *Exposición ambiental y función cognitiva en niños y adolescentes*

La evidencia sobre exposición ambiental fue especialmente limitada. Únicamente Roser et al. (2016) evaluó este tipo de exposición mediante mediciones personales. «Para una exposición ambiental total estimada mediante mediciones personales, con una media de 66,8 μW/m², la SMD correspondiente a exactitud de concentración en el FAKT-II fue −0,09 (IC95 % −0,76 a 0,57; 79 participantes).

El intervalo de confianza fue muy amplio y compatible tanto con una disminución como con un aumento relevante del rendimiento, por lo que la evidencia fue considerada de certeza baja debido a imprecisión. No se dispuso de evidencia sobre exposición ambiental para los restantes dominios cognitivos en niños.

### *Exposición personal y función cognitiva en adultos mayores*

La evidencia en adultos procedió exclusivamente de la cohorte SLAS de Singapur, que estudió personas de 55 años o más durante un seguimiento próximo a cuatro años. La exposición se clasificó a partir del uso autodeclarado del teléfono móvil como nunca o raramente, algunas veces, o frecuentemente.

Los resultados se expresaron como odds ratios de deterioro cognitivo y presentaron intervalos de confianza amplios.

Para la comparación de mayor contraste, correspondiente al uso frecuente, definido como al menos siete llamadas semanales, frente a uso nunca o raro, la OR de deterioro de la función cognitiva global fue 0,81 (IC95 % 0,42 a 1,58; 649 participantes).

Para atención compleja fue 0,67 (IC95 % 0,27 a 1,68; 159 participantes), para función ejecutiva 1,07 (IC95 % 0,37 a 3,05; 146 participantes) y para aprendizaje y memoria 0,75 (IC95 % 0,29 a 1,99; 159 participantes). Los intervalos de confianza fueron compatibles con asociaciones en ambas direcciones y la certeza fue muy baja para todos estos desenlaces.

Aunque algunas estimaciones puntuales fueron inferiores a uno, especialmente para atención compleja y aprendizaje y memoria, los autores interpretaron estas posibles reducciones del riesgo como probablemente poco importantes al considerar también los efectos absolutos; además, las estimaciones fueron muy imprecisas. tanto, estos resultados no constituyen evidencia convincente de un efecto beneficioso del uso del teléfono móvil sobre la cognición.

### *Relación exposición-respuesta y ausencia de evidencia para determinadas exposiciones y dominios*

Uno de los objetivos secundarios de la revisión era determinar si existía una relación exposición-respuesta entre la exposición prolongada a CEM-RF y la función cognitiva. Sin embargo, no se identificó una relación exposición-respuesta para ninguno de los desenlaces cognitivos evaluados.

La escasez de estudios, la heterogeneidad de las métricas de exposición y el predominio de indicadores indirectos limitaron considerablemente la posibilidad de establecer gradientes consistentes.

Las lagunas de evidencia fueron importantes. No se identificaron estudios elegibles sobre exposición ambiental y función cognitiva en adultos o personas mayores, ni sobre exposición ocupacional y cognición.

En niños tampoco se evaluó la función cognitiva global y la evidencia quedó concentrada en atención compleja, función ejecutiva y aprendizaje y memoria. Estas ausencias delimitan de forma importante el alcance de las conclusiones de la revisión.

### *Sesgo de publicación*

El posible sesgo de publicación no pudo evaluarse formalmente. Los autores habían previsto valorar tanto la ausencia de resultados de desenlaces elegibles dentro de los estudios incluidos como la posible existencia de estudios no publicados. Sin embargo, no dispusieron de suficiente información procedente de registros o protocolos detallados y, sobre todo, el número de estudios fue demasiado reducido para elaborar *funnel plots* o aplicar pruebas de asimetría. En consecuencia, la ausencia de una evaluación estadística del sesgo de publicación no debe interpretarse como evidencia de ausencia de este sesgo.

### *Evaluación del riesgo de sesgo*

La evaluación mediante OHAT mostró diferencias relevantes entre los cinco estudios. Tres fueron clasificados globalmente con riesgo de sesgo probablemente alto y dos con riesgo probablemente bajo.

Thomas et al. (2010) y Bhatt et al. (2017) fueron considerados de riesgo probablemente alto, fundamentalmente porque la exposición a CEM-RF se estimó indirectamente a partir del uso autodeclarado de teléfonos móviles y otros teléfonos inalámbricos.

Roser et al. (2016) y Foerster et al. (2018), ambos de HERMES, fueron clasificados con riesgo global probablemente bajo y en *tier* 1. Aunque también utilizaron información autodeclarada, incorporaron métodos adicionales para caracterizar la exposición, incluidos registros de operadores y estimaciones de dosis cerebral de CEM-RF.

Estas aproximaciones redujeron la preocupación por errores en la caracterización de la exposición respecto de los estudios basados exclusivamente en autoinforme.

El estudio de Ng et al. fue clasificado con riesgo global probablemente alto y en *tier* 2. La principal preocupación fue el sesgo por pérdidas o exclusiones, ya que sólo alrededor del 31 % de la cohorte inicial aportó los datos necesarios para ser incluido en los análisis y, además, no todos estos participantes completaron las diferentes pruebas neuropsicológicas.

A ello se añadió una caracterización de la exposición basada en uso autodeclarado del teléfono móvil que no estimaba adecuadamente la intensidad y duración de la exposición real a CEM-RF.

### *Certeza de la evidencia*

La certeza de la evidencia, evaluada mediante GRADE con las adaptaciones de OHAT para estudios de exposición, fue en todos los casos baja o muy baja. Por tanto, aunque el conjunto de los resultados mostró predominantemente poca o ninguna asociación entre la exposición a CEM-RF y la función cognitiva, la confianza que puede depositarse en estas estimaciones es limitada.

En niños y adolescentes, la evidencia sobre exposición personal y atención compleja fue de certeza baja. La certeza fue igualmente baja para aprendizaje y memoria. Para función ejecutiva, la medida principal de Stroop, (B−A)/A, se calificó como de certeza muy baja debido al riesgo de sesgo y a la inconsistencia entre los estudios; para la segunda medida, (D−C)/C, la certeza fue baja.

Para exposición ambiental y atención compleja, la certeza fue baja, fundamentalmente por la marcada imprecisión de la estimación procedente de un único estudio con 79 participantes.

En adultos mayores, la certeza fue muy baja para función cognitiva global, atención compleja, función ejecutiva y aprendizaje y memoria. En todos estos casos la evidencia procedía de un único estudio con riesgo de sesgo probablemente alto y las estimaciones presentaban intervalos de confianza amplios, compatibles tanto con disminuciones como con aumentos del deterioro cognitivo.

En conjunto, Benke et al. consideran que existe una falta consistente de evidencia de asociación entre la exposición a CEM-RF y la función cognitiva en los dominios estudiados. Sin embargo, esta conclusión debe interpretarse atendiendo a la baja o muy baja certeza de la evidencia.

Para niños y adolescentes, los resultados sugieren que la exposición asociada al uso de teléfonos móviles puede tener poca o ninguna repercusión a largo plazo, durante periodos de meses a pocos años, sobre atención compleja, función ejecutiva y aprendizaje y memoria.

Para adultos mayores, los datos disponibles sugieren igualmente poca o ninguna asociación, pero la evidencia es de certeza muy baja. La escasez de estudios, las limitaciones en la evaluación de la exposición y las importantes lagunas relativas a poblaciones, tipos de exposición y dominios cognitivos impiden formular conclusiones más firmes.

Una precisión importante respecto a tus notas: he evitado decir que «no hubo estudios de habilidad perceptivo-motora» de forma absoluta. La propia Tabla 1 de Benke identifica *perceptual motor ability* en Thomas et al., pero la sección de resultados afirma después que ninguno de los estudios incluidos evaluó efectos en ese dominio. Hay, por tanto, una inconsistencia interna en el artículo. Para nuestro informe es más seguro decir que no hubo evidencia susceptible de evaluación o síntesis para ese dominio, en lugar de afirmar categóricamente que ningún estudio incluyó una prueba relacionada con él.

## 4.2.5. Fortalezas y debilidades de la revisión

### *Fortalezas*

La revisión de Benke et al. (2024) presenta varias fortalezas metodológicas relevantes. En primer lugar, fue desarrollada a partir de un protocolo publicado previamente y registrado prospectivamente en PROSPERO, y se informó de acuerdo con PRISMA. La búsqueda bibliográfica fue amplia, sin restricciones de fecha ni idioma, e incluyó cuatro bases de datos, EMF-Portal, revisiones previas de organismos nacionales e internacionales, páginas web institucionales, registros proporcionados por la OMS y las referencias de los estudios elegibles.

La selección de los estudios, la extracción de datos y la evaluación del riesgo de sesgo se realizaron de forma independiente por dos revisores, y los miembros del equipo no participaron en la extracción ni en la evaluación del riesgo de sesgo de estudios de los que fueran autores.

Otra fortaleza importante es la definición prospectiva y relativamente estricta de los criterios de elegibilidad. La exclusión de estudios transversales permitió preservar la temporalidad entre exposición y desenlace, un aspecto esencial cuando se pretenden investigar posibles efectos a largo plazo. Asimismo, se exigió un seguimiento mínimo de seis meses y un contraste entre diferentes niveles de exposición.

La revisión estableció además criterios específicos para evitar indicadores de exposición considerados particularmente inadecuados, como la distancia autodeclarada a antenas o la distancia calculada a estaciones base de telefonía móvil. Para la evaluación de los desenlaces se exigieron pruebas o criterios diagnósticos con evidencia de validez y fiabilidad.

También constituye una fortaleza el tratamiento explícito de la calidad de la evidencia. El riesgo de sesgo se evaluó mediante OHAT, considerando de forma diferenciada selección, confusión, pérdidas o exclusiones, caracterización de la exposición, evaluación de los desenlaces y notificación selectiva. Posteriormente, la certeza de la evidencia se valoró mediante GRADE con las adaptaciones de OHAT para estudios de exposición.

Esta aproximación permite interpretar los resultados teniendo en cuenta no sólo la dirección y magnitud de las estimaciones, sino también las limitaciones metodológicas del cuerpo de evidencia, evitando equiparar automáticamente la ausencia de asociaciones estadísticamente significativas con evidencia concluyente de ausencia de efecto.

La revisión presta además una atención especial a uno de los problemas metodológicos centrales de la epidemiología de los CEM-RF: la caracterización de la exposición. Los autores distinguen entre diferentes fuentes y métricas de exposición y valoran explícitamente su capacidad para representar la exposición realmente relevante para el tejido de interés.

Esta consideración permitió diferenciar estudios basados fundamentalmente en autoinforme de aquellos, como los correspondientes a HERMES, que incorporaron registros de operadores y estimaciones de dosis cerebral acumulada. Esta diferenciación se reflejó posteriormente en la evaluación del riesgo de sesgo y en la certeza de la evidencia.

Finalmente, la separación de los resultados por población, tipo de exposición y dominio cognitivo evita combinar evidencias que responden a preguntas diferentes. Los autores diferencian la exposición personal de la ambiental, los resultados obtenidos en niños y adolescentes de los correspondientes a adultos mayores y los distintos dominios cognitivos. Esta estructura facilita delimitar con precisión qué combinaciones de población, exposición y desenlace cuentan realmente con evidencia y cuáles constituyen todavía lagunas de conocimiento.

### *Limitaciones*

La principal limitación de la revisión procede del propio cuerpo de evidencia disponible. A pesar de la amplitud de la búsqueda, sólo se identificaron cinco estudios correspondientes a cuatro cohortes independientes. Además, cuatro estudios se realizaron en niños o adolescentes y el único estudio en adultos correspondió a personas de 55 años o más. Por tanto, prácticamente no existe evidencia longitudinal elegible para adultos jóvenes o de mediana edad. Tampoco se identificaron estudios elegibles sobre exposición ocupacional ni sobre exposición ambiental en adultos o personas mayores.

En niños, los estudios tampoco abarcaron todos los dominios cognitivos considerados en el protocolo. En consecuencia, las conclusiones sólo pueden aplicarse directamente a las poblaciones, exposiciones y funciones cognitivas efectivamente estudiadas.

Una segunda limitación fundamental afecta a la caracterización de la exposición. La métrica ideal sería la dosis total de energía de RF absorbida por el órgano o tejido de interés, dependiente de la SAR, la duración y las características de la exposición, pero esta información resulta difícil de obtener en estudios epidemiológicos. Por ello, buena parte de la evidencia se basó en indicadores indirectos, especialmente el uso autodeclarado del teléfono móvil o inalámbrico.

Estos indicadores pueden introducir errores de clasificación y, además, el tiempo o la frecuencia de utilización de un dispositivo no equivalen directamente a la dosis de CEM-RF absorbida por el cerebro. Los autores señalan que la clasificación errónea derivada de las diferencias entre información objetiva y autodeclarada tendería generalmente a desplazar las estimaciones hacia el valor nulo, aunque no puede descartarse un sesgo diferencial.

Esta limitación no afectó por igual a todos los estudios. Los trabajos de HERMES utilizaron una evaluación más elaborada de la exposición, combinando información autodeclarada con registros de operadores y estimaciones de dosis cerebral acumulada, y fueron clasificados con riesgo de sesgo probablemente bajo. En cambio, otros estudios se basaron fundamentalmente en el uso autodeclarado de dispositivos inalámbricos y presentaron mayor preocupación por sesgo en la caracterización de la exposición.

En el estudio de adultos mayores, además, la información sobre utilización del teléfono móvil tampoco permitía estimar adecuadamente la intensidad y duración de la exposición real a CEM-RF.

El uso del teléfono móvil como indicador de exposición plantea además un problema específico de interpretación causal. Las actividades asociadas al uso del dispositivo pueden relacionarse con el rendimiento cognitivo independientemente de la exposición a CEM-RF. Por tanto, una eventual asociación entre utilización del teléfono y rendimiento cognitivo no podría atribuirse automáticamente a las radiofrecuencias.

Esta dificultad refuerza la necesidad de utilizar estimaciones dosimétricas más específicas cuando el objetivo es aislar el posible efecto de los CEM-RF de otros factores asociados al comportamiento y al uso de los dispositivos.

La síntesis cuantitativa estuvo asimismo muy condicionada por el escaso número de estudios. Los metaanálisis disponibles combinaron únicamente dos estudios y, además, procedentes del mismo grupo de investigación, circunstancia que llevó a los autores a utilizar modelos de efectos fijos en lugar de los modelos de efectos aleatorios inicialmente previstos.

Esta escasez limita tanto la precisión de las estimaciones agrupadas como la posibilidad de explorar adecuadamente la heterogeneidad y determinar si los resultados son reproducibles en poblaciones y contextos independientes.

La heterogeneidad de las herramientas utilizadas para evaluar la función cognitiva constituye otra limitación. Diferentes estudios emplearon distintas pruebas y métricas para representar dominios conceptualmente semejantes, lo que redujo considerablemente las posibilidades de síntesis cuantitativa.

Esta diversidad metodológica, unida al pequeño número de cohortes, explica que parte de la evidencia tuviera que evaluarse individualmente y que algunos dominios estuvieran representados por un único estudio o carecieran de evidencia evaluable. Los autores destacan, por ello, la conveniencia de una mayor estandarización de las pruebas utilizadas para evaluar dominios cognitivos específicos.

El estudio de adultos mayores presenta limitaciones adicionales. Sólo el 31 % de los participantes aportó datos basales utilizables y no todos completaron las diferentes pruebas neuropsicológicas, lo que determinó una valoración de riesgo probablemente alto de sesgo por pérdidas o exclusiones.

A ello se sumaron la evaluación autodeclarada de la exposición y los tamaños efectivos relativamente pequeños para varios dominios cognitivos, que produjeron intervalos de confianza amplios. Como consecuencia, toda la evidencia correspondiente a adultos mayores fue calificada como de certeza muy baja.

Finalmente, el escaso número de estudios impidió evaluar formalmente el posible sesgo de publicación mediante *funnel plots* o pruebas de asimetría. Tampoco se identificó una relación exposición-respuesta para ninguno de los desenlaces cognitivos.

Esta ausencia de gradiente no debe interpretarse por sí sola como evidencia contra una relación causal, puesto que la escasez de estudios y las limitaciones de las métricas de exposición reducen considerablemente la capacidad para identificarla.

En conjunto, las principales limitaciones no derivan tanto del procedimiento seguido por la revisión sistemática como de la escasez y reducida diversidad de la evidencia primaria disponible, de las dificultades para caracterizar de manera válida la exposición individual a CEM-RF y de la heterogeneidad de las pruebas cognitivas. Estos factores explican que, pese al predominio de resultados compatibles con poca o ninguna asociación, la certeza de la evidencia se mantenga entre baja y muy baja.

# 7. Conclusiones

Benke et al. (2024) concluyen que la evidencia epidemiológica disponible sobre los posibles efectos a largo plazo de la exposición a CEM-RF sobre la función cognitiva es escasa y presenta importantes limitaciones, pero que los resultados disponibles muestran de forma predominantemente consistente poca o ninguna asociación entre la exposición y los dominios cognitivos estudiados.

Esta interpretación debe realizarse teniendo en cuenta que la certeza de la evidencia fue baja o muy baja para todos los desenlaces evaluados, por lo que las conclusiones no deben entenderse como demostración de ausencia de cualquier posible efecto de los CEM-RF sobre la cognición.

En niños y adolescentes, la exposición personal asociada al uso de teléfonos móviles puede tener poca o ninguna repercusión a largo plazo, durante periodos de meses a pocos años, sobre la atención compleja, la función ejecutiva y el aprendizaje y la memoria. La certeza fue baja para atención compleja y aprendizaje y memoria, y muy baja para función ejecutiva.

Aunque algunos estudios individuales mostraron asociaciones puntuales entre determinadas métricas de exposición y resultados cognitivos, especialmente para memoria figural, estas asociaciones no fueron consistentes entre las diferentes medidas de exposición ni configuraron un patrón global que apoyase un deterioro cognitivo relacionado con los CEM-RF.

La evidencia sobre exposición ambiental en niños y adolescentes fue mucho más limitada. Procedió de un único estudio y únicamente permitió evaluar atención compleja, con una estimación muy imprecisa y certeza baja. Por tanto, no existe un cuerpo de evidencia suficiente para establecer conclusiones generales sobre los posibles efectos cognitivos a largo plazo de la exposición ambiental a CEM-RF durante estas etapas de la vida.

En adultos mayores, los datos disponibles sugieren igualmente poca o ninguna asociación entre el uso del teléfono móvil y el deterioro de la función cognitiva global, la atención compleja, la función ejecutiva o el aprendizaje y la memoria. Sin embargo, toda esta evidencia procede de una única cohorte y fue calificada como de certeza muy baja debido, entre otros factores, al riesgo de sesgo por pérdidas o exclusiones, a las limitaciones de la caracterización de la exposición y a la elevada imprecisión de las estimaciones. En consecuencia, la evidencia disponible en adultos mayores es insuficiente para establecer conclusiones firmes.

La revisión tampoco encontró evidencia de una relación exposición-respuesta para ninguno de los dominios cognitivos evaluados. No obstante, este resultado debe interpretarse con cautela, ya que el pequeño número de estudios y las limitaciones de las métricas utilizadas para estimar la exposición reducen la capacidad para identificar un posible gradiente.

Además, existen importantes lagunas de evidencia: no se identificaron estudios elegibles sobre exposición ocupacional, la exposición ambiental apenas fue investigada y determinadas poblaciones y dominios cognitivos están insuficientemente representados o carecen de evidencia evaluable.

En conjunto, la revisión muestra una ausencia consistente de evidencia de asociación entre la exposición a CEM-RF y las funciones cognitivas estudiadas, pero la baja o muy baja certeza de la evidencia no permite excluir posibles efectos con un grado elevado de confianza. Esta distinción es importante: el predominio de estimaciones próximas al valor nulo debe interpretarse conjuntamente con la baja o muy baja certeza de la evidencia y con las importantes lagunas existentes. Por ello, la conclusión más ajustada a los resultados de Benke et al. es que la evidencia disponible sugiere poca o ninguna asociación para las combinaciones de exposición, población y dominio cognitivo estudiadas, pero la confianza en estas conclusiones continúa siendo limitada.

### *Implicaciones para investigación futura*

Benke et al. (2024) consideran necesaria nueva investigación epidemiológica longitudinal que permita superar las principales limitaciones de la evidencia disponible. Una prioridad es ampliar considerablemente la diversidad de las poblaciones estudiadas.

La evidencia actual está concentrada en niños y adolescentes y en una única cohorte de adultos mayores, por lo que se necesitan estudios en adultos de diferentes edades y contextos y una representación más amplia de las poblaciones actualmente insuficientemente estudiadas También es necesario cubrir las importantes lagunas existentes en relación con determinados dominios cognitivos y con exposiciones ambientales y ocupacionales.

Una cuestión especialmente importante es mejorar la evaluación de la exposición. Los futuros estudios deberían reducir la dependencia del uso autodeclarado del teléfono móvil o inalámbrico como sustituto de la exposición a CEM-RF y utilizar, siempre que sea posible, métodos objetivos que permitan estimar de manera más válida la exposición relevante para los órganos y tejidos de interés. Los autores destacan la utilidad de combinar información sobre utilización de dispositivos con registros objetivos y modelos dosimétricos, siguiendo aproximaciones como las empleadas en HERMES.

La mejora de la caracterización de la exposición resulta esencial tanto para reducir el error de clasificación como para poder investigar adecuadamente posibles relaciones exposición-respuesta.

Los autores señalan además que la evolución de las tecnologías y de los patrones de utilización de los dispositivos, incluida la introducción de 5G y el creciente predominio de aplicaciones frente a las llamadas de voz, modifica los perfiles de exposición, lo que refuerza la necesidad de desarrollar modelos capaces de estimar la deposición de energía y la duración de la exposición.

Los estudios futuros deberían asimismo separar, en la medida de lo posible, los efectos atribuibles a la exposición a CEM-RF de aquellos relacionados con el propio uso de los dispositivos. El tiempo dedicado al teléfono móvil puede asociarse con factores conductuales, sociales o de estilo de vida que también pueden influir sobre la cognición.

Por ello, utilizar exclusivamente indicadores de uso dificulta distinguir un eventual efecto de las radiofrecuencias de otros efectos vinculados al comportamiento asociado al dispositivo. Una mejor dosimetría, acompañada de un control adecuado de los factores de confusión, permitiría abordar esta cuestión con mayor validez causal.

Otra prioridad es mejorar la comparabilidad de la evaluación cognitiva. La diversidad de pruebas y métricas empleadas en los estudios incluidos redujo las posibilidades de realizar síntesis cuantitativas y de comparar resultados entre cohortes. Benke et al. señalan la necesidad de una mayor estandarización de las pruebas utilizadas para medir dominios cognitivos específicos.

El empleo de instrumentos validados, comparables entre estudios y suficientemente sensibles a cambios longitudinales facilitaría la acumulación de evidencia y permitiría metaanálisis más informativos. En particular, los autores proponen considerar la prueba de Stroop de palabras y colores para la evaluación de la función ejecutiva, por tratarse de un instrumento ampliamente validado, relativamente sencillo de administrar en muestras grandes y susceptible de comparación entre diferentes contextos culturales.

Finalmente, son necesarias cohortes prospectivas de tamaño suficiente, con seguimientos prolongados y pérdidas reducidas, que incorporen diferentes niveles de exposición y permitan estudiar relaciones exposición-respuesta. Además, el reducido número de cohortes disponibles pone de manifiesto la conveniencia de disponer de evidencia procedente de nuevas cohortes y contextos de investigación. Una evidencia más amplia, independiente y metodológicamente homogénea permitiría aumentar la precisión de las estimaciones, evaluar mejor la heterogeneidad y elevar la certeza actualmente baja o muy baja de las conclusiones.

En este sentido, la principal necesidad identificada por la revisión no es simplemente aumentar el número de estudios, sino generar evidencia longitudinal de mayor calidad, con mejor caracterización dosimétrica de la exposición, evaluación cognitiva estandarizada, control adecuado de la confusión y representación más amplia de poblaciones, fuentes de exposición y dominios cognitivos. Sólo mediante este tipo de estudios será posible determinar con mayor certeza si la ausencia predominante de asociaciones observada hasta ahora refleja realmente una ausencia de efectos cognitivos relevantes de los CEM-RF a largo plazo o está condicionada, al menos en parte, por las limitaciones de la evidencia epidemiológica disponible.

## Bibliografía


Benke, G., Abramson, M.J., Brzozek, C., McDonald, S., Kelsall, H., Sanagou, M., Zeleke, B.M., Kaufman, J., Brennan, S., Verbeek, J., Karipidis, K., 2024. The effects of radiofrequency exposure on cognition: A systematic review and meta-analysis of human observational studies. Environ Int 188, 108779. https://doi.org/10.1016/j.envint.2024.108779

Benke, G., Abramson, M.J., Zeleke, B.M., Kaufman, J., Karipidis, K., Kelsall, H., McDonald, S., Brzozek, C., Feychting, M., Brennan, S., 2022. The effect of long-term radiofrequency exposure on cognition in human observational studies: A protocol for a systematic review. Environ. Int. 159, 106972. https://doi.org/10.1016/j.envint.2021.106972

Pophof, B., Burns, J., Danker-Hopfe, H., Dorn, H., Egblomassé-Roidl, C., Eggert, T., Fuks, K., Henschenmacher, B., Kuhne, J., Sauter, C., Schmid, G., 2021. The effect of exposure to radiofrequency electromagnetic fields on cognitive performance in human experimental studies: A protocol for a systematic review. Environment International 157, 106783. https://doi.org/10.1016/j.envint.2021.106783

Pophof, B., Kuhne, J., Schmid, G., Weiser, E., Dorn, H., Henschenmacher, B., Burns, J., Danker-Hopfe, H., Sauter, C., 2024. The effect of exposure to radiofrequency electromagnetic fields on cognitive performance in human experimental studies: Systematic review and *meta*-analyses. Environ. Int. 191, 108899. https://doi.org/10.1016/j.envint.2024.108899

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