revista UniNorte DE MEDICINA Y CIENCIAS DE LA SALUD

ISSN: 3134-7835
**Vol. 15 (2026)**
Núm. 3

ARTÍCULO DE REVISIÓN

Medicina

# Eficacia y seguridad de las terapias basadas en incretinas para el manejo del peso en adultos sin diabetes mellitus tipo 2 con sobrepeso u obesidad: Una revisión sistemática y metaanálisis

Lohamy Stefanni Pereira de Quadros[1], Taiula Lázara Mariana Bernardes de Lima[1], Luana Cristina de Oliveira e Souza[1], Gabrielle Luz Ribeiro de Jesus[1], Sarah Medeiros Ramos Pereira[1], Tania Matos da Costa[1], Victor Begnini[1], Thamires Ribeiro Dias[1], Fernanda Aparecida Marchesi[1], Isabella Camargo de Macedo[1], Paola Britos[1,2]

[1]Facultad de Ciencias Médicas, Universidad Privada del Este, filial Ciudad del Este, Paraguay; [2]ChauxLab Institute, Asunción, Paraguay.

**Enviado:** 1 de agosto de 2026
**Aceptado:** 12 de octubre de 2026
**Publicado:** 15 de octubre de 2026

**DOI:** 10.5281/zenodo.23393271

Licencia Creative Commons Atribución 4.0 Internacional (CC-BY)

## RESUMEN

**Antecedentes:** Los ensayos pivotales de las terapias basadas en incretinas muestran reducciones de peso sin precedentes en adultos sin diabetes mellitus tipo 2, pero su síntesis cuantitativa exige distinguir el diseño de los ensayos, la clase de fármaco y la seguridad. **Objetivo:** Estimar el efecto de estas terapias sobre el peso corporal, la probabilidad de perder ≥ 5 % del peso y los eventos adversos graves en adultos con sobrepeso u obesidad sin diabetes tipo 2. **Métodos:** Revisión sistemática y metaanálisis registrados en PROSPERO (CRD420261427544) y reportados según PRISMA 2020, con búsquedas en PubMed/MEDLINE, Scite y ClinicalTrials.gov. El riesgo de sesgo se evaluó con RoB 2 y la certeza, con GRADE. Se usó un modelo de efectos aleatorios (DerSimonian-Laird) con intervalo de predicción, análisis de REML con Hartung-Knapp y estratificación por clase de fármaco. Los ensayos de inicio de tratamiento se analizaron aparte de los de mantenimiento o retirada aleatorizada. **Resultados:** De 889 registros se incluyeron 101 informes, correspondientes a 52 ensayos aleatorizados con informe primario, 36 de ellos con placebo. En los ensayos de inicio de ≥ 24 semanas (k = 13), el cambio de peso frente a placebo fue de −9,88 puntos porcentuales (IC 95 %: −12,94 a −6,82; $I^2$ = 98 %; intervalo de predicción

**Autor corresponsal:** Paola Britos. Facultad de Ciencias Médicas, Universidad Privada del Este, filial Ciudad del Este, Paraguay; ChauxLab Institute, Asunción, Paraguay
Correo electrónico: paola.britos@chauxlab.com

**Financiación:** Los autores declaran que no recibieron fondos externos para este estudio.

**Conflictos de interés:** Los autores declaran no tener conflictos de interés financieros, personales, profesionales, institucionales ni de otra índole relacionados con este trabajo.

−22,5 a 2,8). La diferencia entre clases explicó la mayor parte de la heterogeneidad: liraglutida −3,50, semaglutida −11,51 y tirzepatida −16,79 puntos porcentuales ($I^2$ dentro de cada clase, de 26 a 65 %). El riesgo relativo de perder ≥ 5 % fue de 3,11 (IC 95 %: 2,45 a 3,94; k = 10) y el de eventos adversos graves, de 1,18 (IC 95 %: 0,79 a 1,75; k = 12; recuentos derivados). La certeza fue moderada para las estimaciones de peso por clase, baja para la estimación global y para la pérdida ≥ 5 %, y muy baja para los eventos adversos graves. **Conclusiones:** Las terapias basadas en incretinas reducen el peso frente a placebo en adultos sin diabetes tipo 2, con una magnitud que depende de la clase de fármaco. La estimación global combina intervenciones demasiado distintas para interpretarse como un efecto único. La seguridad, evaluada con datos derivados, no mostró un aumento claro de los eventos adversos graves, con certeza muy baja.

*Palabras clave:* obesidad, agonistas del receptor de GLP-1, incretinas, pérdida de peso, seguridad, metaanálisis

## INTRODUCCIÓN

La obesidad constituye uno de los mayores desafíos de salud pública contemporáneos. El análisis conjunto más reciente de tendencias mundiales, que combinó 3663 estudios poblacionales con mediciones antropométricas de 222 millones de niños, adolescentes y adultos en 200 países, documenta que la prevalencia de obesidad en adultos se ha más que triplicado desde 1975 y que, hacia 2022, más de 880 millones de adultos vivían con esta condición (1). Este incremento sostenido conlleva una carga creciente de enfermedad cardiometabólica, osteoarticular y oncológica asociada al exceso de adiposidad, y ha desplazado a la obesidad del terreno estético al de una enfermedad crónica recurrente que requiere abordaje médico estructurado y de largo plazo, más allá de la recomendación genérica de reducir la ingesta calórica y aumentar la actividad física.

Las intervenciones exclusivamente conductuales —dieta, actividad física y modificación del estilo de vida— logran reducciones de peso modestas y, salvo en programas intensivos y prolongados, difíciles de sostener en el tiempo, lo que ha impulsado el desarrollo de un arsenal farmacológico específico. La guía de práctica clínica de la Endocrine Society sobre manejo farmacológico de la obesidad ya señalaba la necesidad de incorporar tratamiento médico como complemento de la modificación conductual en pacientes con obesidad o con sobrepeso y comorbilidad asociada (2). En este contexto surgieron los agonistas del receptor de GLP-1 y, más recientemente, los agonistas duales y otros agentes basados en incretinas, que actúan tanto a nivel central —modulando circuitos hipotalámicos y del tronco encefálico que regulan el apetito y la saciedad— como periférico, enlenteciendo el vaciamiento gástrico y modificando la señalización de otras hormonas intestinales relacionadas con la homeostasis energética (3).

La evidencia individual de ensayos pivotales como STEP-1, con semaglutida subcutánea semanal (4), y SURMOUNT-1, con el agonista dual tirzepatida (5), ha mostrado reducciones de peso sin precedentes para agentes farmacológicos en población sin diabetes mellitus tipo 2. Sin embargo, el corpus de ensayos clínicos de agentes incretínicos en esta población —que excluye el efecto atenuante que suele observarse en personas con diabetes mellitus tipo 2— se ha expandido rápidamente e incluye moléculas de distintas clases (agonistas simples, duales y en desarrollo clínico), poblaciones diversas (obesidad no complicada, apnea obstructiva del sueño, síndrome de ovario poliquístico, insuficiencia cardíaca con fracción de eyección preservada, entre otras), diseños que responden a preguntas distintas (inicio de tratamiento frente a mantenimiento o retirada aleatorizada) y duraciones de seguimiento heterogéneas. Persiste, por tanto, la necesidad de una síntesis cuantitativa que integre esta evidencia dispersa sin ocultar esas diferencias, y que informe a la vez la seguridad.

El objetivo de esta revisión sistemática y metaanálisis fue estimar el efecto de las terapias basadas en incretinas frente a placebo sobre el cambio de peso corporal, en términos porcentuales y absolutos, la probabilidad de alcanzar una pérdida de peso ≥ 5 % y los eventos adversos graves, en adultos con sobrepeso u obesidad sin diabetes mellitus tipo 2, distinguiendo los ensayos de inicio de tratamiento de los de mantenimiento o retirada aleatorizada y

explorando las fuentes de heterogeneidad por clase de fármaco y duración.

# MÉTODOS

## Diseño, registro y desviaciones del protocolo

Se condujo una revisión sistemática con metaanálisis de ensayos clínicos aleatorizados de terapias basadas en incretinas para el manejo del peso en adultos con sobrepeso u obesidad sin diabetes mellitus tipo 2, registrada de manera prospectiva en PROSPERO con el número CRD420261427544. El reporte siguió la declaración PRISMA 2020 (6). Respecto del protocolo registrado, se introdujeron las siguientes desviaciones, que se reconocen explícitamente:

1. El protocolo incluía estudios aleatorizados y no aleatorizados; esta revisión analiza únicamente ensayos clínicos aleatorizados. 2. El protocolo contemplaba CENTRAL, Embase, Web of Science, LILACS, SciELO y la plataforma ICTRP, que no se buscaron; en su lugar, se incluyó Scite como fuente suplementaria (véase más adelante). 3. El protocolo preveía el IMC como desenlace primario y desenlaces cardiometabólicos secundarios; no se analizaron. Los eventos adversos graves, desenlace primario registrado, se analizaron con los datos disponibles en la extracción, con recuentos derivados cuando el artículo reportaba porcentajes. 4. El protocolo preveía metaanálisis solo con $I^2 < 50$ %, y con $I^2 \geq 50$ % análisis de subgrupos con la heterogeneidad como hallazgo. Como esa condición no se cumplió, el estimador global se presenta como exploratorio, y la lectura principal es la estratificada por clase de fármaco, dosis y duración. 5. El protocolo especificaba el riesgo relativo como medida de los desenlaces dicotómicos; la versión inicial de este manuscrito empleó la razón de momios, y en esta versión el riesgo relativo es la medida principal. 6. La síntesis cuantitativa se restringió a comparaciones con placebo. No se documentó el seguimiento de las listas de referencias de los estudios incluidos ni se calculó el acuerdo entre revisores.

## Criterios de elegibilidad

Se definieron los criterios de elegibilidad según el marco PICO. **Población:** adultos con sobrepeso u obesidad, sin diagnóstico de diabetes mellitus tipo 2, según la definición empleada en cada ensayo de origen (cohortes explícitamente no diabéticas o subgrupos separables de cohortes mixtas). **Intervención:** agonistas del receptor de GLP-1 u otros agentes basados en incretinas, incluidos los agonistas duales, en cualquier dosis o vía de administración aprobada o en investigación clínica destinada al manejo del peso. **Comparador:** placebo y/o control activo, según lo reportado en cada estudio; la síntesis cuantitativa se restringió a comparaciones controladas con placebo. **Desenlaces:** cambio del peso corporal en términos porcentuales y en kilogramos, proporción de participantes con pérdida de peso $\geq 5$ % y eventos adversos graves. **Diseño de estudio:** ensayos clínicos aleatorizados de cualquier fase con texto completo disponible. Los ensayos se clasificaron como de **inicio de tratamiento**, cuando la aleatorización ocurre al comenzar el fármaco y el efecto se mide desde el valor basal, o de **mantenimiento o retirada aleatorizada**, cuando la aleatorización ocurre tras una fase de introducción abierta o una pérdida de peso previa inducida por dieta y el efecto se mide desde la aleatorización (STEP-4, SURMOUNT-4 y dos ensayos de liraglutida tras una pérdida de peso inducida por dieta).

## Fuentes de información y estrategia de búsqueda

Los registros se identificaron a partir de búsquedas en PubMed/MEDLINE (ecuaciones 1 a 4 del protocolo), ClinicalTrials.gov (ecuación 5) y Scite (ecuaciones 1 a 4 y búsquedas complementarias), entre enero de 2010 y mayo de 2026 y sin restricción de idioma, con exportación a un gestor bibliográfico el 2 de agosto de 2026 (889 registros). Cada registro conserva la etiqueta de su fuente y de la ecuación que lo recuperó. Los textos completos de las ecuaciones y el número de registros por fuente están disponibles a petición de la autora corresponsal. La ecuación 6 del protocolo no tiene registros etiquetados y no puede darse por ejecutada. Los registros de repositorios de ensayos clínicos sin informe primario publicado no se trataron como fuente de datos elegible, dado que no permiten extraer desenlaces verificables ni evaluar el riesgo de sesgo.

## Proceso de selección

Los registros se depuraron por duplicados mediante identificadores únicos y se cribaron por título y resumen por dos revisores del equipo de investigación de forma independiente, con resolución de las discrepancias por consenso; el mismo procedimiento se aplicó a la evaluación a texto completo. Los textos completos que no pudieron recuperarse pese a las gestiones institucionales y manuales realizadas se clasificaron como no recuperados y no se computaron como exclusiones por criterios de elegibilidad. Tras la revisión por pares se repitió el intento de recuperación por vías legales de acceso abierto, sin acceso institucional ni contacto con autores: de los 80 informes inicialmente no recuperados, se obtuvieron 9. Los informes se

vincularon con su ensayo por número de registro y nombre del programa. Solo el informe primario de cada ensayo aportó datos a la síntesis; los informes secundarios, los posanálisis y las extensiones se usaron para identificar el ensayo y no se contaron como estudios adicionales, de modo que no hay participantes ni comparaciones duplicados (detalle disponible a petición). En los ensayos con varias dosis se usó el brazo de dosis más alta frente a placebo; en los ensayos con varios fármacos y un placebo compartido (STEP-8), el tamaño del placebo se dividió entre las comparaciones.

### Extracción de datos y riesgo de sesgo

Las características de los estudios y los desenlaces se extrajeron mediante un formulario estructurado que capturó, para cada informe, el diseño y el programa del ensayo, el entorno y el tamaño de la muestra aleatorizada, el estado de diabetes de la población, el índice de masa corporal basal, el fármaco, la dosis y la vía de administración, la duración del seguimiento, el comparador, el desenlace primario declarado, el cambio de peso en kilogramos y en términos porcentuales por brazo, la proporción de participantes con pérdida ≥ 5 % y ≥ 10 %, la discontinuación y los eventos adversos graves. La extracción fue realizada por revisores del equipo de investigación y cada extracción se sometió a un control de calidad documentado antes de su incorporación a la síntesis. Los datos de los ensayos recuperados tras la revisión por pares se extrajeron con el mismo formulario.

El riesgo de sesgo del desenlace de peso se evaluó con la herramienta RoB 2 (7), en sus cinco dominios —proceso de aleatorización, desviaciones de las intervenciones previstas, datos de desenlace faltantes, medición del desenlace y selección del resultado reportado—, con un juicio global por estudio derivado del algoritmo de la herramienta («algunas preocupaciones» cuando al menos un dominio lo presentaba y ninguno era de alto riesgo). La justificación de cada dominio, por ensayo, está disponible a petición de la autora corresponsal. Los ensayos con riesgo de sesgo global alto se excluyeron de los análisis primarios y se incorporaron en un análisis de sensibilidad cuando existían datos con varianza verificable.

### Síntesis y metaanálisis

El desenlace continuo primario fue la diferencia de medias del cambio porcentual del peso corporal frente a placebo, restringido a comparaciones con varianza muestral verificable —desviación estándar, error estándar o intervalo de confianza del 95 % a nivel de brazo, o diferencia de tratamiento estimada publicada con su intervalo de confianza —; no se emplearon desviaciones estándar asumidas. El análisis primario se restringió a ensayos de inicio de tratamiento con duración ≥ 24 semanas; los ensayos de mantenimiento o retirada aleatorizada se analizaron por separado. Los análisis secundarios incluyeron el cambio de peso en kilogramos, el riesgo relativo de alcanzar una pérdida ≥ 5 % (con la razón de momios como medida secundaria) y el riesgo relativo de eventos adversos graves. En los ensayos con celdas de cero o con la totalidad de los participantes con evento se añadió 0,5 a cada celda (8), y se repitió el análisis con esa corrección aplicada a todos los ensayos y sin los ensayos con celdas extremas. Los recuentos de eventos adversos graves se derivaron del porcentaje reportado y del tamaño del brazo cuando el artículo no daba el recuento; esos ensayos se señalan en las tablas.

Todos los análisis agrupados emplearon un modelo de efectos aleatorios con ponderación por varianza inversa y el estimador de DerSimonian-Laird (9), como se registró en el protocolo. Se informaron el estadístico $I^2$, la prueba Q, la varianza entre estudios ($\tau^2$) y el intervalo de predicción del 95 % (10). Como análisis de sensibilidad se estimó el modelo de máxima verosimilitud restringida (REML) con el ajuste de Hartung-Knapp-Sidik-Jonkman (11,12) y se repitió el análisis excluyendo un ensayo cada vez, restringido a ensayos con riesgo de sesgo bajo y sin los ensayos incorporados tras la revisión por pares. La heterogeneidad se exploró mediante subgrupos por clase de fármaco y por duración (24 a 51 frente a ≥ 52 semanas), con una prueba Q entre subgrupos; no se hicieron metarregresiones por el número reducido de ensayos. Los subgrupos con un solo ensayo no se acompañan de medidas de heterogeneidad. El sesgo de publicación se evaluó con gráficos de embudo y la prueba de Egger (13) en los análisis con diez o más ensayos, considerando las limitaciones de esas pruebas ante heterogeneidad importante y pocos estudios (14). La certeza de la evidencia se calificó con el enfoque GRADE (15) para cada desenlace y clase. Los análisis se realizaron con código propio en Python 3 (SciPy), verificado al reproducir exactamente las estimaciones de la versión inicial de este manuscrito.

## RESULTADOS

### Selección de estudios

De 889 registros identificados se eliminaron 8 duplicados y se cribaron 881 títulos y resúmenes. La decisión de consenso a nivel de título y resumen incluyó 173 registros, clasificó 44 como dudosos y

excluyó 664. De los 217 informes buscados para recuperación a texto completo —la suma de los incluidos y los dudosos— se recuperaron 146 (137 inicialmente y 9 tras la revisión por pares) y no fue posible recuperar 71, lo que constituye una limitación importante que se discute más adelante. Se excluyeron 45 informes a texto completo, con las siguientes razones: diseño de estudio incorrecto (n = 20), población incorrecta (n = 16), desenlace incorrecto (n = 7) y tipo de publicación incorrecto (n = 2); el listado completo está disponible a petición de la autora corresponsal. Se incluyeron 101 informes: 52 informes primarios de ensayos aleatorizados y 49 informes secundarios, posanálisis, extensiones o informes primarios no elegibles para la síntesis (estudios no aleatorizados o con diabetes no separable), que no aportan datos cuantitativos (Figura 1). Siete programas de ensayos grandes —entre ellos SELECT, STEP-5, STEP-6, SURMOUNT-3, SURMOUNT-5, SURMOUNT-J y REDEFINE-1— están representados en la cohorte solo por informes secundarios, porque su informe primario no se recuperó o no estaba en el corpus.

Los materiales suplementarios —ecuaciones de búsqueda y registros por fuente; listados de informes no recuperados y excluidos; correspondencia entre informes y ensayos; evaluación RoB 2 por dominio con justificación; figuras de sensibilidad, embudo y mantenimiento; tabla completa de características de los 52 ensayos; y tablas de datos por ensayo de los metaanálisis— están disponibles a petición de la autora corresponsal.

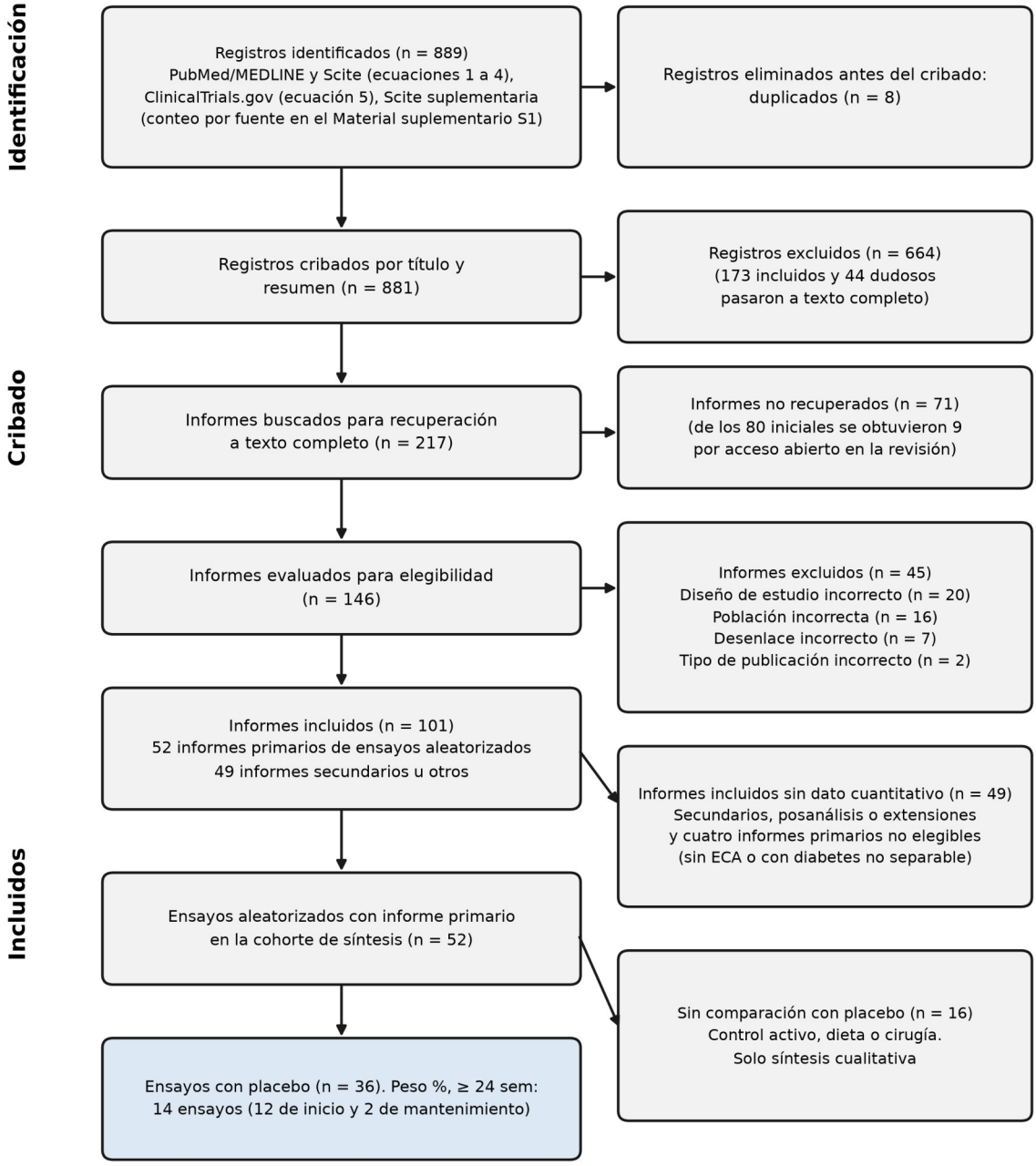



**Figura 1. Diagrama de flujo PRISMA 2020 del proceso de identificación, cribado e inclusión de estudios.**

## Características de los estudios incluidos

La cohorte de síntesis comprendió 52 ensayos clínicos aleatorizados con informe primario, de los cuales 36 correspondieron a comparaciones controladas con placebo y 16 a comparadores activos, como metformina, terapia conductual intensiva sin fármaco, dieta, sitagliptina (16) o combinaciones farmacológicas alternativas. Las clases farmacológicas representadas fueron liraglutida (n = 22), exenatida (n = 9), semaglutida (n = 10), tirzepatida (n = 5), mazdutida (n = 1), dulaglutida (n = 1), efpeglenatida (n = 1), danuglipron oral (n = 1) y otros agonistas incretínicos en desarrollo clínico (n = 3; STEP-8 evaluó dos fármacos y se cuenta en ambas clases), lo que evidencia el peso predominante de los agentes de primera generación en el volumen total de evidencia disponible. El tamaño de la muestra aleatorizada osciló entre 30 y 2539 participantes por ensayo y la duración del seguimiento, entre 5 y 104 semanas. Las poblaciones incluyeron obesidad no complicada y contextos clínicos específicos —apnea obstructiva del sueño, síndrome de ovario poliquístico, insuficiencia cardíaca con fracción de eyección preservada, prediabetes y periodos posteriores a cirugía bariátrica o a pérdida de peso inducida por dieta—, lo que amplía la aplicabilidad de la evidencia a costa de heterogeneidad clínica adicional. Las características individuales de los 52 ensayos (diseño, dosis, comparador, IMC y juicio RoB 2) figuran en una tabla extendida disponible a petición de la autora corresponsal.

## Riesgo de sesgo

Del total de 52 ensayos evaluados con RoB 2, 15 recibieron un juicio global de riesgo bajo, 34 de algunas preocupaciones y 3 de riesgo alto, de modo que la mayoría se ubicó en la categoría intermedia. Los ensayos de riesgo alto fueron: un análisis del desenlace de peso restringido a un subgrupo posaleatorización de respondedores tempranos, que plantea una preocupación seria de sesgo de selección del resultado reportado (17); un ensayo con atrición muy elevada, con cerca de 60 % de abandono del tratamiento (18); y el estudio de Basolo *et al.* (19), en el que solo un subconjunto de los aleatorizados tiene peso a 24 semanas. Los dominios con más preocupaciones fueron la falta de datos de desenlace (26 de 52 ensayos) y las desviaciones de las intervenciones previstas (20 de 52, sobre todo en ensayos abiertos). La justificación por dominio y por ensayo está disponible a petición de la autora corresponsal. Los ensayos con riesgo alto se excluyeron de los análisis primarios, con la excepción de la sensibilidad descrita más adelante.

## Cambio porcentual del peso corporal frente a placebo: ensayos de inicio de tratamiento

En el análisis primario, restringido a ensayos de inicio de tratamiento con duración ≥ 24 semanas, varianza muestral verificable y riesgo de sesgo no alto, contribuyeron 12 ensayos y 13 comparaciones (STEP-8 aporta dos). La diferencia de medias combinada por efectos aleatorios fue de −9,88 puntos porcentuales (IC 95 %: −12,94 a −6,82) a favor del tratamiento, con heterogeneidad extremadamente elevada ($I^2$ = 98 %; $\tau^2$ = 30,5; intervalo de predicción del 95 %: −22,51 a 2,75), de modo que en un ensayo futuro de este conjunto el efecto podría ir desde una reducción de más de 22 puntos hasta un leve aumento. Este estimador global debe leerse como exploratorio. El modelo REML con Hartung-Knapp dio −9,88 (IC 95 %: −13,11 a −6,64). Al excluir un ensayo cada vez, la estimación varió entre −9,20 y −10,49 puntos y su intervalo excluyó siempre el cero (Figura 2). Los datos numéricos por ensayo del metaanálisis de peso porcentual están disponibles a petición de la autora corresponsal.

Las diferencias entre clases explicaron la mayor parte de la heterogeneidad (prueba Q entre subgrupos, $p < 0{,}001$). Las estimaciones por clase fueron: liraglutida −3,50 puntos porcentuales (k = 5; IC 95 %: −4,25 a −2,75; $I^2$ = 26 %), semaglutida −11,51 (k = 4; IC 95 %: −12,82 a −10,20; $I^2$ = 65 %), tirzepatida −16,79 (k = 2; IC 95 %: −19,35 a −14,22; $I^2$ = 58 %), mazdutida −14,31 (k = 1; IC 95 %: −15,88 a −12,74) y otra incretina −14,94 (k = 1; IC 95 %: −17,45 a −12,43). En los subgrupos de un solo ensayo no se informa heterogeneidad ni intervalo de predicción. Las estimaciones de las clases se obtuvieron con distinta dosis, duración y población, por lo que su comparación directa no está ajustada; la única comparación aleatorizada directa recuperada, STEP-8, mostró una diferencia entre semaglutida 2,4 mg y liraglutida 3,0 mg de −9,4 puntos porcentuales (IC 95 %: −12,0 a −6,8) (20). La duración no explicó la heterogeneidad de forma significativa (24 a 51 semanas: −8,07, k = 5; ≥ 52 semanas: −11,04, k = 8; p = 0,29).

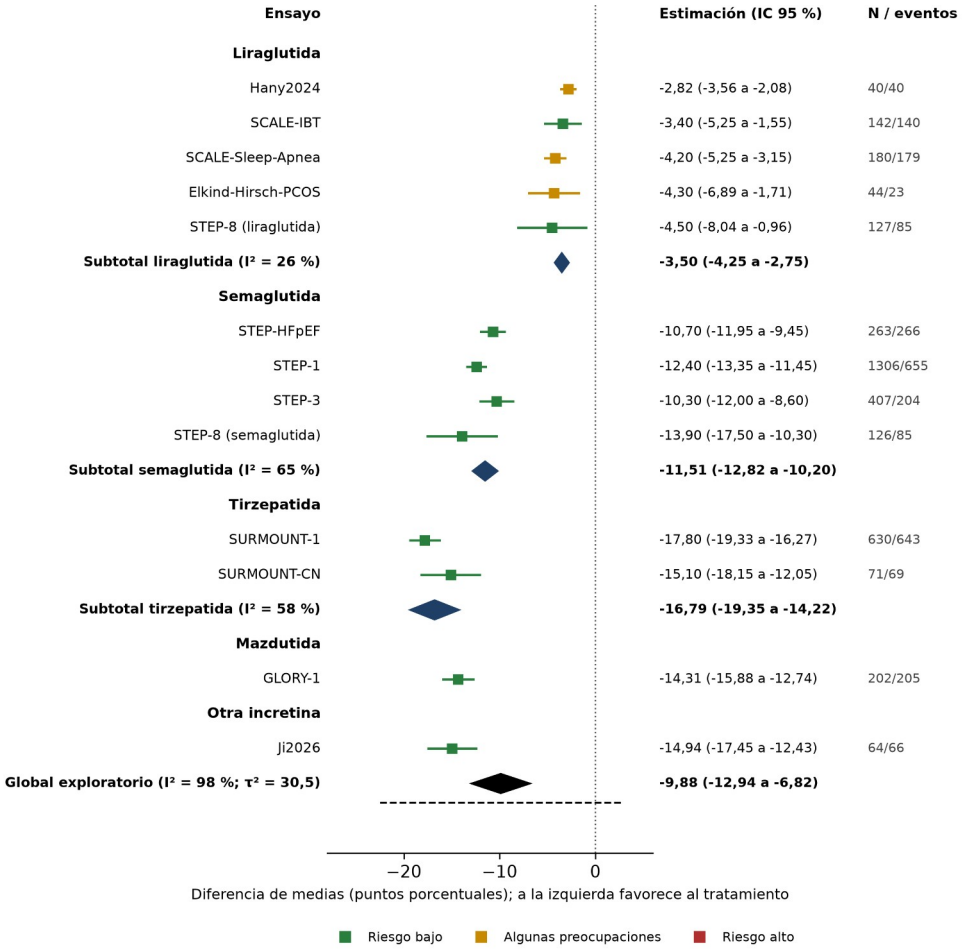



**Figura 2. Cambio porcentual del peso corporal frente a placebo, ensayos de inicio de tratamiento de ≥ 24 semanas, por clase de fármaco.** La línea discontinua bajo el global es el intervalo de predicción del 95 %. En STEP-8 el placebo compartido se dividió entre los dos brazos activos.

El análisis restringido a los 10 ensayos de riesgo bajo dio −11,77 puntos (IC 95 %: −14,27 a −9,27; $I^2$ = 95 %; intervalo de predicción −21,15 a −2,39), consistente en dirección con el análisis primario y de magnitud algo mayor (material suplementario a petición); sin los ensayos incorporados tras la revisión por pares (k = 9) fue de −9,42 (IC 95 %: −13,21 a −5,64), y sin STEP-8, de −10,00 (IC 95 %: −13,33 a −6,66). Los ensayos STEP-3 (21), STEP-8 (20) y SURMOUNT-CN (22), recuperados por vías de acceso abierto, contribuyeron a estas estimaciones. La Tabla 1 resume los análisis agrupados principales.

## Tabla 1. Resumen de los resultados agrupados principales

| Desenlace | Análisis | k | Estimación (IC 95 %) | $I^2$ (%) |
|---|---|---|---|---|
| Peso (pp) | Inicio, ≥ 24 sem (primario) | 13 | −9,88 (−12,94 a −6,82) | 98 |
| Peso (pp) | Inicio, liraglutida | 5 | −3,50 (−4,25 a −2,75) | 26 |
| Peso (pp) | Inicio, semaglutida | 4 | −11,51 (−12,82 a −10,20) | 65 |
| Peso (pp) | Inicio, tirzepatida | 2 | −16,79 (−19,35 a −14,22) | 58 |
| Peso (pp) | Mantenimiento o retirada | 2 | −17,09 (−21,69 a −12,48) | 95 |
| Pérdida ≥ 5 % (RR) | Inicio | 10 | 3,11 (2,45 a 3,94) | 81 |
| Eventos adversos graves (RR) | Inicio | 12 | 1,18 (0,79 a 1,75) | 66 |

*Nota:* pp = puntos porcentuales; RR = riesgo relativo. Intervalo de predicción del análisis primario de peso: −22,51 a 2,75. Los análisis de sensibilidad (bajo riesgo de sesgo, peso en kilogramos, razón de momios y mantenimiento de ≥ 5 % o eventos adversos graves) y los datos por ensayo están disponibles a petición de la autora corresponsal.

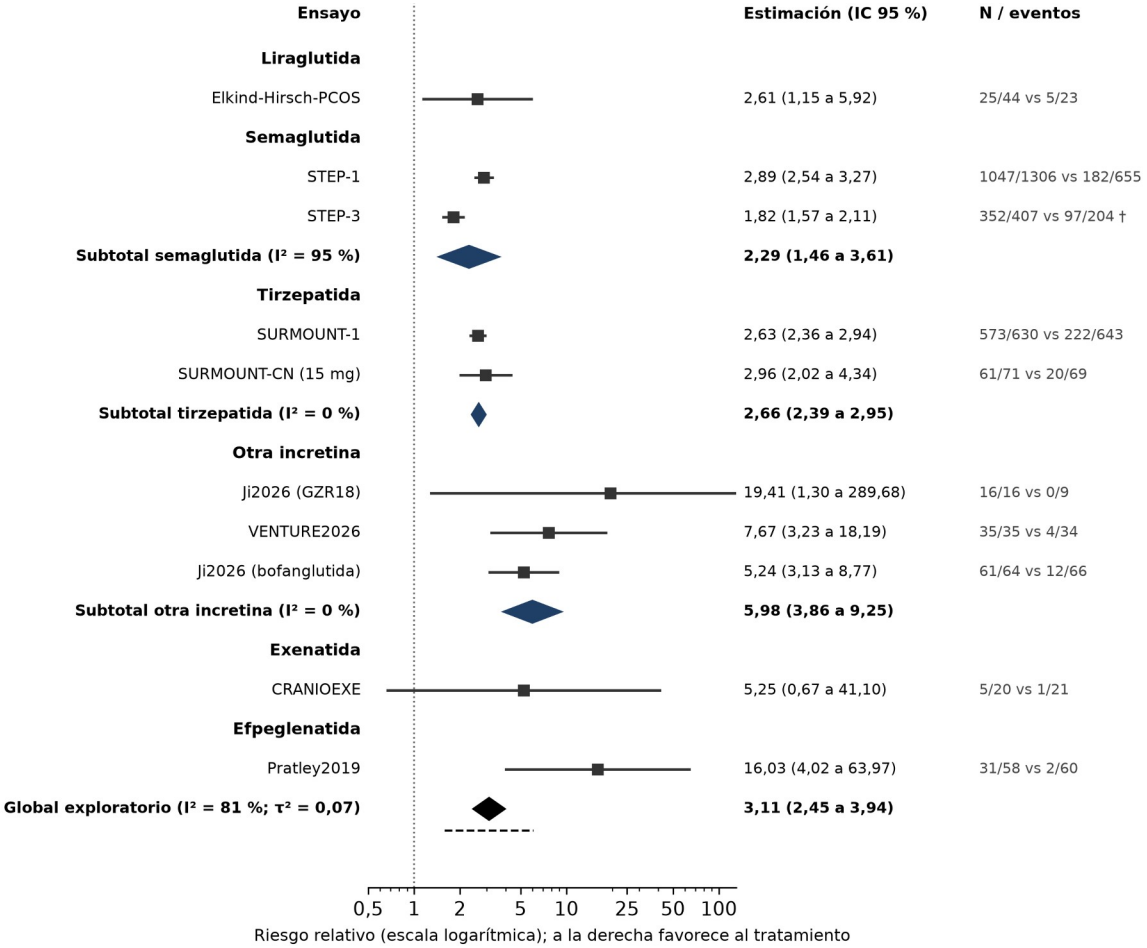



**Figura 3. Pérdida de peso ≥ 5 %: riesgo relativo frente a placebo, ensayos de inicio, por clase.** † Recuentos derivados del porcentaje reportado y del número de aleatorizados. Corrección de 0,5 solo en ensayos con celdas de cero o totales.

## Ensayos de mantenimiento o retirada aleatorizada

En los dos ensayos de mantenimiento con duración ≥ 24 semanas (STEP-4 y SURMOUNT-4), la diferencia de medias combinada fue de −17,09 puntos porcentuales (IC 95 %: −21,69 a −12,48; $I^2$ = 95 %). Este valor se mide desde la aleatorización, tras una fase de introducción abierta que enriquece la muestra con quienes toleran y responden al fármaco, y no es comparable con el efecto de los ensayos de inicio; por eso no se combinó con ellos. Con k = 2, el intervalo con el ajuste de Hartung-Knapp fue amplio (−46,93 a 12,76).

## Cambio absoluto de peso

En los ensayos de inicio con dispersión reportada a nivel de brazo para el peso en kilogramos (k = 13, de cualquier duración), la diferencia de medias combinada fue de −8,01 kg (IC 95 %: −10,00 a −6,02; $I^2$ = 94 %; intervalo de predicción −16,04 a 0,02); restringida a ≥ 24 semanas (k = 6), fue de −8,66 kg (IC 95 %: −12,81 a −4,50). La inclusión del ensayo de riesgo alto de Basolo *et al.* (−1,72 kg; IC 95 %: −5,77 a 2,30; n = 13 y 20) redujo la estimación a −7,64 kg (IC 95 %: −9,59 a −5,70) sin modificar la conclusión. En los tres ensayos de mantenimiento fue de −11,39 kg (IC 95 %: −15,76 a −7,03). Estos análisis en kilogramos parten de un conjunto distinto de ensayos que el de puntos porcentuales y no son comparables punto a punto.

## Pérdida de peso ≥ 5 %

Para el desenlace de respondedores contribuyeron 10 ensayos de inicio con recuentos de eventos disponibles. El riesgo relativo combinado de alcanzar una pérdida ≥ 5 % frente a placebo fue de 3,11 (IC 95 %: 2,45 a 3,94; $I^2$ = 81 %; $\tau^2$ = 0,07; intervalo de predicción 1,59 a 6,09); con REML y Hartung-Knapp fue de 3,45 (IC 95 %: 2,25 a 5,29). La razón de momios secundaria fue de 15,57 (IC 95 %: 9,58 a 25,31), mayor que el riesgo relativo porque la proporción de respondedores con el fármaco supera el 80 % en varios ensayos; la razón de momios no equivale a una razón de probabilidades de la misma magnitud, y por eso el riesgo relativo es la medida interpretable (Figura 3). Dos ensayos tuvieron celdas de cero o con todos los participantes como respondedores; sin ellos (k = 8), el riesgo relativo fue de 2,87 (IC 95 %: 2,28 a 3,61) y la razón de momios, de 13,37 (IC 95 %: 8,54 a 20,94); con la corrección de 0,5 en todos los ensayos fue de 3,07 (IC 95 %: 2,43 a 3,88). Los ensayos de mantenimiento (k = 3) dieron un riesgo relativo de 1,46 (IC

95 %: 1,26 a 1,69), como era esperable cuando todos los participantes ya han perdido peso en una fase previa. Las estimaciones por clase fueron: semaglutida 2,29 (k = 2), tirzepatida 2,66 (k = 2) y otras incretinas 5,98 (k = 3), con diferencias entre clases (p = 0,001), pero basadas en pocos ensayos y con intervalos amplios.

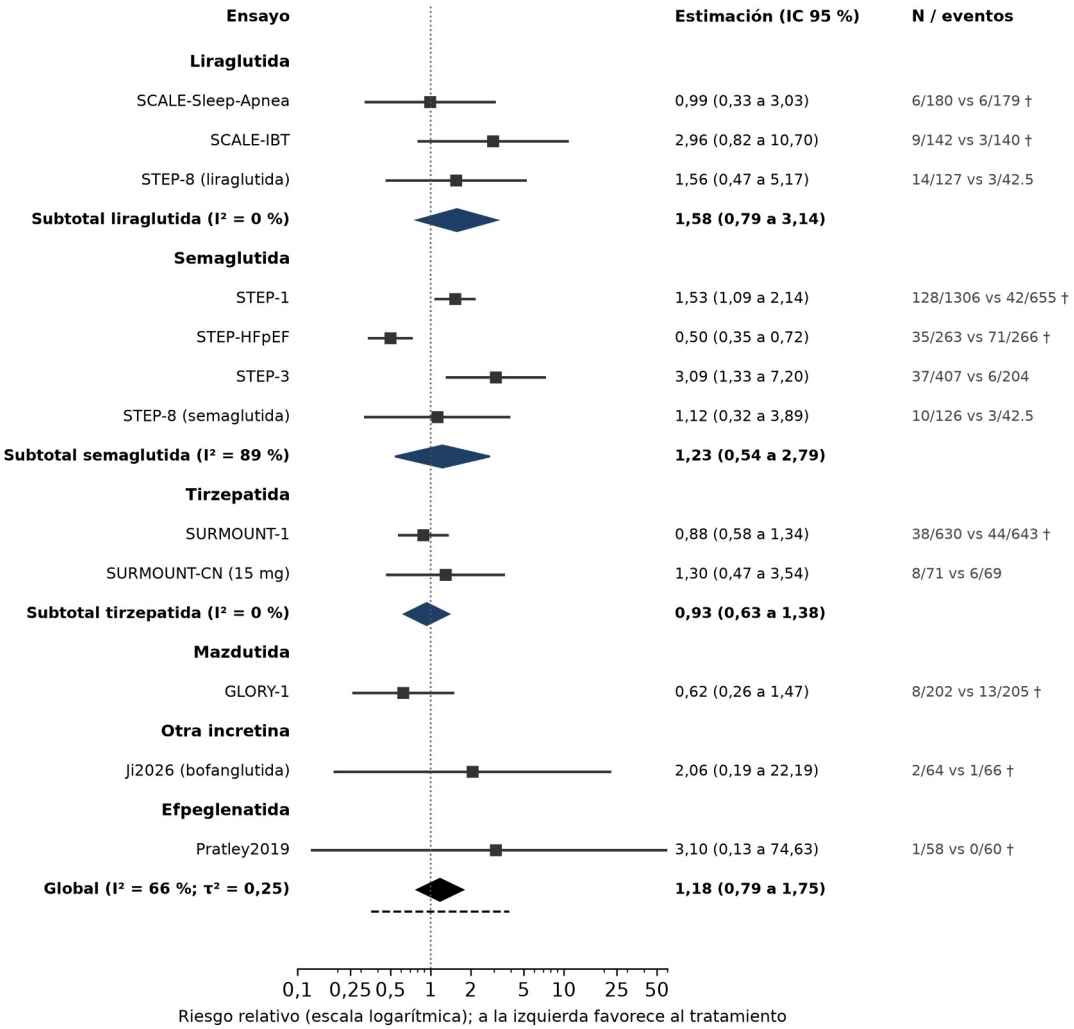



**Figura 4. Eventos adversos graves: riesgo relativo frente a placebo, ensayos de inicio, por clase.** † Recuentos derivados del porcentaje reportado y del tamaño del brazo (aproximación). En STEP-8 el placebo compartido se dividió entre los dos brazos activos.

## Seguridad: eventos adversos graves

Los eventos adversos graves se pudieron agrupar en 12 ensayos de inicio. El riesgo relativo combinado frente a placebo fue de 1,18 (IC 95 %: 0,79 a 1,75; $I^2$ = 66 %; $\tau^2$ = 0,25; intervalo de predicción 0,36 a 3,90), sin diferencia clara; con REML y Hartung-Knapp fue de 1,17 (IC 95 %: 0,76 a 1,81). Por clase, las estimaciones fueron de 1,58 para liraglutida (k = 3), 1,23 para semaglutida (k = 4) y 0,93 para tirzepatida (k = 2), todas con intervalos que incluían el valor nulo (Figura 4). En los tres ensayos de mantenimiento el riesgo relativo fue de 0,90 (IC 95 %: 0,33 a 2,46). La heterogeneidad se concentró en STEP-HFpEF, con menos eventos graves con semaglutida (35 de 263 frente a 71 de 266), población con insuficiencia cardíaca en la que los eventos graves reflejan en gran medida la enfermedad de base.

En STEP-3 los eventos graves fueron más frecuentes con semaglutida (37 de 407 frente a 6 de 204), con más trastornos de la vesícula biliar (4,9 % frente a 1,5 %). Los eventos gastrointestinales fueron los efectos adversos más comunicados en los ensayos, pero se reportaron con definiciones y formatos tan distintos que no se agruparon; la proporción de participantes con eventos gastrointestinales fue mayor con el fármaco que con placebo en los ensayos pivotales y en la mayoría de los que la reportaron (por ejemplo, 82,8 % frente a 63,2 % en STEP-3 y 32,4 % frente a 5,8 % con náuseas en SURMOUNT-CN con la dosis de 15 mg). Estos análisis de seguridad se basan, en la mayoría de los ensayos, en recuentos derivados de porcentajes y no pueden descartar diferencias menores en eventos graves.

### Sesgo de publicación

En los análisis con diez o más ensayos, la prueba de Egger no mostró asimetría significativa del embudo: cambio porcentual del peso (k = 13; intercepto −6,15; p = 0,19), peso en kilogramos (k = 13; p = 0,26), riesgo relativo de pérdida ≥ 5 % (k = 10; p = 0,11) y eventos adversos graves (k = 12; p = 0,28). Estos resultados deben interpretarse con cautela por la heterogeneidad importante, el número reducido de ensayos y la potencia limitada de la prueba, y no descartan un sesgo de recuperación por los informes no obtenidos.

### Certeza de la evidencia

La Tabla 2 presenta la certeza de la evidencia (GRADE) por desenlace y clase. La certeza fue moderada para las estimaciones de peso de liraglutida, semaglutida y tirzepatida; baja para la estimación global del peso porcentual, para el peso en kilogramos, para la pérdida ≥ 5 % y para los ensayos de mantenimiento; y muy baja para los eventos adversos graves.

**Tabla 2. Resumen de hallazgos y certeza de la evidencia (GRADE)**

| Desenlace | Estrato (k) | Efecto (IC 95 %) | Certeza |
|---|---|---|---|
| Peso (pp) | Global, inicio (13) | −9,88 (−12,94 a −6,82) | Baja |
| Peso (pp) | Liraglutida (5) | −3,50 (−4,25 a −2,75) | Moderada |
| Peso (pp) | Semaglutida (4) | −11,51 (−12,82 a −10,20) | Moderada |
| Peso (pp) | Tirzepatida (2) | −16,79 (−19,35 a −14,22) | Moderada |
| Pérdida ≥ 5 % (RR) | Inicio (10) | 3,11 (2,45 a 3,94) | Baja |
| Eventos adversos graves (RR) | Inicio (12) | 1,18 (0,79 a 1,75) | Muy baja |
| Peso (pp) | Mantenimiento (2) | −17,09 (−21,69 a −12,48) | Baja |

*Nota:* pp = puntos porcentuales; RR = riesgo relativo. La degradación de la certeza global de peso se debió a inconsistencia muy seria ($I^2$ = 98 %; el intervalo de predicción incluye el cero); en las clases, a inconsistencia o riesgo de sesgo serios. En ≥ 5 % y eventos adversos graves pesaron el uso de recuentos derivados, la inconsistencia y, en seguridad, la imprecisión. El detalle por dominio GRADE y el análisis de peso en kilogramos están disponibles a petición de la autora corresponsal.

## DISCUSIÓN

Esta revisión sistemática y metaanálisis, que incorporó 52 ensayos clínicos aleatorizados con informe primario de terapias basadas en incretinas en adultos sin diabetes mellitus tipo 2, encontró una reducción del peso corporal frente a placebo cuya magnitud depende de la clase de fármaco: unos 3,5 puntos porcentuales con liraglutida, 11,5 con semaglutida y 16,8 con tirzepatida en los ensayos de inicio de tratamiento de 24 semanas o más. La diferencia entre clases, y no la inconsistencia dentro de una misma clase, explicó la mayor parte de la heterogeneidad. La estimación global combinada, de −9,9 puntos porcentuales, tiene un intervalo de predicción que incluye el cero y no debe interpretarse como el efecto esperable de un fármaco o ensayo individual: resume intervenciones, dosis, duraciones y poblaciones demasiado diferentes. La dirección del efecto fue consistente en todos los ensayos de inicio, mientras que su magnitud es incierta fuera de cada clase.

Los resultados son coherentes con el mecanismo descrito para estos agentes, que combina efectos centrales sobre los circuitos hipotalámicos de regulación del apetito con efectos periféricos sobre el vaciamiento gástrico y la señalización de saciedad (3), y con la eficacia observada en los ensayos pivotales (4,5). Esta revisión no estima la superioridad de una clase sobre otra: las estimaciones por clase provienen de ensayos con dosis, duraciones y poblaciones distintas y no ajustadas, y una comparación indirecta no sustituye a ensayos de comparación directa. La única comparación aleatorizada directa recuperada (STEP-8) favoreció a la semaglutida sobre la liraglutida, y la comparación directa de tirzepatida con semaglutida (SURMOUNT-5) figura entre los informes no recuperados.

La separación de los ensayos de inicio y de mantenimiento fue necesaria para interpretar las magnitudes. En STEP-4 y SURMOUNT-4 se aleatoriza tras una fase de introducción abierta, de modo que el efecto frente a placebo refleja el beneficio de continuar el tratamiento en quienes ya lo toleran y han respondido, y no el efecto de iniciarlo. Su estimación combinada (−17,1 puntos porcentuales) no es comparable con la de los ensayos de inicio, y combinarlas habría sobrestimado el efecto esperable al comenzar un tratamiento. En la pérdida de peso ≥ 5 %, el riesgo relativo (3,1) es la medida interpretable: la razón de momios (15,6) crece al aumentar la proporción de respondedores y puede

interpretarse erróneamente como una probabilidad quince veces mayor, cuando en realidad la proporción de quienes alcanzan esa pérdida se triplica aproximadamente.

En cuanto a la seguridad, los eventos adversos graves no mostraron un aumento claro con las incretinas (riesgo relativo 1,18; IC 95 %: 0,79 a 1,75), pero la certeza fue muy baja: los recuentos se derivaron de porcentajes, el desenlace era secundario en la mayoría de los ensayos y el intervalo es compatible con un beneficio o con un aumento relevante del riesgo. La heterogeneidad se concentró en un ensayo en insuficiencia cardíaca, donde los eventos graves reflejan la enfermedad de base. Los eventos gastrointestinales, los más comunicados, fueron más frecuentes con el fármaco en los ensayos pivotales y en la mayoría de los que los reportaron, aunque no pudieron agruparse. Estos hallazgos son compatibles con el perfil conocido de la clase, pero este metaanálisis no sustituye a la información de seguridad de los ensayos pivotales ni a los estudios de farmacovigilancia.

Esta revisión tiene limitaciones que deben considerarse al interpretar sus hallazgos. Primero, la búsqueda se restringió a PubMed/MEDLINE, Scite y ClinicalTrials.gov, sin las bases adicionales del protocolo (Embase, CENTRAL, Web of Science, LILACS, SciELO) ni el seguimiento de listas de referencias; la exportación original no conservó un registro del procedimiento de búsqueda más allá de las etiquetas de fuente de cada registro, y no se calculó el acuerdo entre revisores. Segundo, 71 de los 217 informes buscados no se recuperaron, entre ellos publicaciones primarias de ensayos grandes contra placebo y de comparación directa (STEP-7, STEP-9, STEP-10, OASIS-1 y OASIS-4, SURMOUNT-5, entre otros), y siete programas de ensayos están representados solo por informes secundarios. Aunque los ensayos pivotales recuperados muestran todos reducciones grandes de peso, la ausencia de estos estudios puede alterar la composición del conjunto y la estimación de cada clase de manera que no puede cuantificarse.

Tercero, la heterogeneidad estadística fue muy alta en el análisis global; se limitó explorándola por clase y duración, pero no pudo resolverse completamente ni analizarse por dosis o población por el número de ensayos. Cuarto, los análisis de seguridad usaron recuentos derivados de porcentajes, no cubrieron los desenlaces no graves y no incluyeron IMC ni desenlaces cardiometabólicos, a diferencia de lo previsto en el protocolo. Quinto, las pruebas de asimetría tienen poca potencia con heterogeneidad importante y menos de veinte ensayos. Sexto, el análisis de sensibilidad con riesgo alto se limitó a un ensayo con datos verificables. Séptimo, esta revisión se centró en el efecto del tratamiento activo y no evaluó la sostenibilidad de la pérdida de peso tras su discontinuación; un estudio observacional de seguimiento a dos años en mujeres con síndrome de ovario poliquístico tratadas con semaglutida y metformina documentó una recuperación parcial y progresiva del peso perdido tras la retirada del agonista (23), de modo que el efecto cuantificado corresponde a un tratamiento en curso.

Como fortalezas, cabe destacar el uso exclusivo de varianza muestral verificable —sin desviaciones estándar asumidas—, la separación de los ensayos de inicio y de mantenimiento, la presentación de la heterogeneidad con $\tau^2$ e intervalos de predicción en lugar de un único estimador global, la evaluación individual del riesgo de sesgo y la declaración explícita de las desviaciones del protocolo.

Los estudios futuros deberían evaluar de forma prospectiva y con comparación directa las distintas clases de fármaco, reportar los eventos adversos graves y gastrointestinales con definiciones homogéneas, y analizar la pérdida de peso a largo plazo y tras la discontinuación del tratamiento.

En conclusión, las terapias basadas en incretinas reducen el peso corporal frente a placebo en adultos con sobrepeso u obesidad sin diabetes tipo 2, con una magnitud que depende de la clase de fármaco y una dirección de efecto consistente. La estimación global combina intervenciones demasiado heterogéneas para interpretarse como un efecto único, y la comparación entre clases no está ajustada. Los ensayos de mantenimiento o retirada aleatorizada responden a otra pregunta y no deben combinarse con los de inicio. Los eventos adversos graves no aumentaron de forma clara, con certeza muy baja, y la falta de 71 informes, incluidos ensayos grandes, limita la generalización de las estimaciones por clase.

## REFERENCIAS

1. NCD Risk Factor Collaboration (NCD-RisC). Worldwide trends in underweight and obesity from 1990 to 2022: a pooled analysis of 3663 population-representative studies with 222 million children, adolescents, and adults. Lancet. 2024;403(10431):1027-1050. https://doi.org/10.1016/S0140-6736(23)02750-2
2. Apovian CM, Aronne LJ, Bessesen DH, McDonnell ME, Murad MH, Pagotto U, et al. Pharmacological management of obesity: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2015;100(2):342-362. https://doi.org/10.1210/jc.2014-3415

3. Moiz A, Filion KB, Tsoukas MA, Yu OHY, Peters TM, Eisenberg MJ. Mechanisms of GLP-1 receptor agonist-induced weight loss: a review of central and peripheral pathways in appetite and energy regulation. Am J Med. 2025;138(6):934-940. https://doi.org/10.1016/j.amjmed.2025.01.021
4. Wilding JPH, Batterham RL, Calanna S, Davies M, Van Gaal LF, Lingvay I, et al. Once-weekly semaglutide in adults with overweight or obesity. N Engl J Med. 2021;384(11):989-1002. https://doi.org/10.1056/NEJMoa2032183
5. Jastreboff AM, Aronne LJ, Ahmad NN, Wharton S, Connery L, Alves B, et al. Tirzepatide once weekly for the treatment of obesity. N Engl J Med. 2022;387(3):205-216. https://doi.org/10.1056/NEJMoa2206038
6. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. https://doi.org/10.1136/bmj.n71
7. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019;366:l4898. https://doi.org/10.1136/bmj.l4898
8. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. Stat Med. 2004;23(9):1351-1375. https://doi.org/10.1002/sim.1761
9. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177-188. https://doi.org/10.1016/0197-2456(86)90046-2
10. IntHout J, Ioannidis JPA, Rovers MM, Goeman JJ. Plea for routinely presenting prediction intervals in meta-analysis. BMJ Open. 2016;6(7):e010247. https://doi.org/10.1136/bmjopen-2015-010247
11. IntHout J, Ioannidis JPA, Borm GF. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. BMC Med Res Methodol. 2014;14:25. https://doi.org/10.1186/1471-2288-14-25
12. Deeks JJ, Higgins JPT, Altman DG, McKenzie JE, Veroniki AA. Chapter 10: Analysing data and undertaking meta-analyses. En: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editores. Cochrane Handbook for Systematic Reviews of Interventions. Versión 6.5. Cochrane; 2024. Disponible en: https://www.cochrane.org/authors/handbooks-and-manuals/handbook/current/chapter-10
13. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629-634. https://doi.org/10.1136/bmj.315.7109.629
14. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002. https://doi.org/10.1136/bmj.d4002
15. Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. J Clin Epidemiol. 2011;64(4):383-394. https://doi.org/10.1016/j.jclinepi.2010.04.026
16. Mashayekhi M, Beckman JA, Nian H, Garner EM, Mayfield D, Devin JK, et al. Comparative effects of weight loss and incretin-based therapies on vascular endothelial function, fibrinolysis and inflammation in individuals with obesity and prediabetes: a randomized controlled trial. Diabetes Obes Metab. 2023;25(2):570-580. https://doi.org/10.1111/dom.14903
17. Rodgers M, Migdal AL, Rodríguez TG, et al. Weight loss outcomes among early high responders to exenatide treatment. Front Endocrinol (Lausanne). 2021. https://doi.org/10.3389/fendo.2021.742873
18. Buckeridge C, Cobain S, Bays HE, Matsuoka N, et al. Efficacy and safety of danuglipron (PF-06882961) in adults with obesity. Diabetes Obes Metab. 2025. https://doi.org/10.1111/dom.16534
19. Basolo A, Burkholder J, Osgood K, Graham A, Bundrick S, Frankl J, et al. Exenatide has a pronounced effect on energy intake but not energy expenditure in non-diabetic subjects with obesity: a randomized, double-blind, placebo-controlled trial. Metabolism. 2018;85:116-125. https://doi.org/10.1016/j.metabol.2018.03.017
20. Rubino DM, Greenway FL, Khalid U, O'Neil PM, Rosenstock J, Sørrig R, et al. Effect of weekly subcutaneous semaglutide vs daily liraglutide on body weight in adults with overweight or obesity without diabetes: the STEP 8 randomized clinical trial. JAMA. 2022;327(2):138-150. https://doi.org/10.1001/jama.2021.23619
21. Wadden TA, Bailey TS, Billings LK, Davies M, Frias JP, Koroleva A, et al. Effect of subcutaneous semaglutide vs placebo as an adjunct to intensive behavioral therapy on body weight in adults with overweight or obesity: the STEP 3 randomized clinical trial. JAMA. 2021;325(14):1403-1413. https://doi.org/10.1001/jama.2021.1831
22. Zhao L, Cheng Z, Lu Y, Liu M, Chen H, Zhang M, et al. Tirzepatide for weight reduction in Chinese adults with obesity: the SURMOUNT-CN randomized clinical trial. JAMA. 2024;332(7):551-560. https://doi.org/10.1001/jama.2024.9217
23. Jensterle M, Ferjan S, Janez A. The maintenance of long-term weight loss after semaglutide withdrawal in obese women with PCOS treated with metformin: a 2-year observational study. Front Endocrinol (Lausanne). 2024. https://doi.org/10.3389/fendo.2024.1366940