REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

TERAPIAS BASEADAS EM INCRETINAS E DEMÊNCIA: CONCILIANDO SINAIS EPIDEMIOLÓGICOS DE PREVENÇÃO COM ENSAIOS NEGATIVOS DE MODIFICAÇÃO DA DOENÇA NA DOENÇA DE ALZHEIMER SINTOMÁTICA INICIAL

INCRETIN-BASED THERAPIES AND DEMENTIA: RECONCILING EPIDEMIOLOGICAL PREVENTION SIGNALS WITH NEGATIVE DISEASE-MODIFYING TRIALS IN EARLY SYMPTOMATIC ALZHEIMER'S DISEASE

TERAPIAS BASADAS EN INCRETINAS Y DEMENCIA: CONCILIANDO SEÑALES EPIDEMIOLÓGICAS DE PREVENCIÓN CON ENSAYOS NEGATIVOS DE MODIFICACIÓN DE LA ENFERMEDAD EN LA ENFERMEDAD DE ALZHEIMER SINTOMÁTICA TEMPRANA

João Fernandes Floriano
Farmacêutico Clínico em Endocrinologia e Metabologia
CRF-SC 9130 | RQE 8459/06
Doutorando em Ciências da Saúde, Farmácia, Centro Internacional de Pesquisas Integralize (CIPI), Brasil
Mestrando em Gestão em Saúde, University Seal, Estados Unidos
Especialista em Epidemiologia
ORCID: 0009-0000-5791-029X
E-mail: joaofernandesfloriano@gmail.com

Maria Elizabeth Menezes
Doutora em Microbiologia, Coordenação de Ensino e Pesquisa (COEPE), Secretaria de Estado da Saúde de Santa Catarina, Brasil
ORCID: 0000-0003-3573-3346

Resumo

Os agonistas do receptor do peptídeo semelhante ao glucagon-1 (AR GLP-1) e o agonismo duplo dos receptores do polipeptídeo insulinotrópico dependente de glicose (GIP) e do GLP-1 têm sido propostos como intervenções contra o declínio cognitivo. Esta revisão crítica estruturada avalia evidências clínicas, epidemiológicas, farmacológicas e de biomarcadores até 23 de setembro de 2026 e examina por que sinais favoráveis de prevenção divergem de ensaios negativos de modificação da doença. Dados agrupados de ensaios cardiovasculares e uma metanálise posterior de ensaios randomizados, parcialmente sobrepostos, sugerem menos eventos de demência registrados como desfechos secundários ou de segurança com AR GLP-1, enquanto emulações de ensaios-alvo permanecem vulneráveis a canalização, viés de vigilância e confundimento residual. Uma análise de sensibilidade por E-value mostra variação acentuada na robustez das principais estimativas observacionais, desde baixa robustez nos limites dos intervalos de confiança em várias comparações

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

até estimativas de efeito tão grandes que se tornam causalmente implausíveis, compatíveis com viés estrutural. Em EVOKE e EVOKE+ (n=3.808), a semaglutida oral não retardou a progressão no Clinical Dementia Rating-Sum of Boxes em 104 semanas, apesar de alterações modestas em biomarcadores liquóricos relacionados à doença de Alzheimer. O ensaio de fase 2b com liraglutida na doença de Alzheimer e o ensaio de fase 3 com exenatida na doença de Parkinson também foram negativos nos desfechos primários. Comparamos quatro explicações concorrentes (estágio da doença, exposição central insuficiente, epidemiologia confundida e mediação vascular indireta) e as evidências capazes de discriminá-las. Formalizamos uma Hipótese da Janela Metabólica (Metabolic Window Hypothesis), geradora de hipóteses e dependente do estágio, com predições e condições de refutação explícitas, e propomos um Framework de Precisão Incretina-Neurodegeneração como estrutura provisória para pesquisas estratificadas por biomarcadores. Nem a semaglutida nem a tirzepatida são terapias estabelecidas para prevenir ou tratar a doença de Alzheimer ou a demência.

Palavras-chave: Doença de Alzheimer; agonistas do receptor de GLP-1; semaglutida; tirzepatida; E-value.

Abstract

Glucagon-like peptide-1 receptor agonists (GLP-1RAs) and dual glucose-dependent insulinotropic polypeptide (GIP)/GLP-1 receptor agonism have been proposed as candidate interventions against cognitive decline. This structured critical review appraises clinical, epidemiological, pharmacological and biomarker evidence through 23 September 2026 and examines why favourable prevention signals diverge from negative disease-modification trials. Pooled cardiovascular-trial data and a later, partly overlapping randomized-trial meta-analysis suggest fewer dementia events, recorded as secondary or safety outcomes, with GLP-1RAs, whereas target-trial emulations remain vulnerable to channelling, surveillance and residual confounding. A quantitative E-value sensitivity analysis shows marked variation in robustness across key observational estimates, from weak confidence-bound robustness in several comparisons to effect estimates so large as to be causally implausible and compatible with structural bias. In EVOKE and EVOKE+ (n=3,808), oral semaglutide did not slow Clinical Dementia Rating-Sum of Boxes progression at 104 weeks despite modest changes in cerebrospinal-fluid Alzheimer-related biomarkers. The phase 2b liraglutide trial in AD and the phase 3 exenatide trial in Parkinson's disease were also negative on their primary outcomes. We compare four competing explanations (disease stage, insufficient central exposure, confounded epidemiology and indirect vascular mediation) and specify evidence capable of discriminating among them. We further formalize a hypothesis-generating, stage-dependent Metabolic Window Hypothesis with explicit predictions and refutation conditions, and propose a Precision Incretin-Neurodegeneration Framework as a provisional operational scaffold for biomarker-stratified research. Neither semaglutide nor tirzepatide is an established therapy for preventing or treating Alzheimer's disease or dementia.

Keywords: Alzheimer disease; glucagon-like peptide-1 receptor agonists; semaglutide; tirzepatide; E-value.

Resumen

Los agonistas del receptor del péptido similar al glucagón tipo 1 (AR GLP-1) y el agonismo dual GIP/GLP-1 (GIP: polipéptido insulinotrópico dependiente de glucosa) se han propuesto como intervenciones contra el deterioro cognitivo. Esta revisión crítica estructurada evalúa evidencia clínica, epidemiológica, farmacológica y de biomarcadores hasta el 23 de septiembre de 2026 y examina por qué las señales favorables de prevención divergen de ensayos negativos de modificación de la enfermedad. Datos agrupados de ensayos cardiovasculares y un metaanálisis posterior de ensayos aleatorizados, parcialmente superpuestos, sugieren menos eventos de demencia registrados como desenlaces secundarios o de seguridad con AR GLP-1, y las emulaciones de ensayos objetivo siguen expuestas a canalización, sesgo de vigilancia y confusión residual. Un análisis de sensibilidad mediante E-value muestra marcada variación en la robustez de las principales estimaciones observacionales, desde baja robustez en los límites de confianza de varias comparaciones hasta efectos tan grandes que resultan causalmente inverosímiles y compatibles con sesgo estructural. En

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

EVOKE y EVOKE+ (n=3.808), la semaglutida oral no enlenteció la progresión del Clinical Dementia Rating-Sum of Boxes a las 104 semanas, pese a cambios modestos en biomarcadores de Alzheimer en líquido cefalorraquídeo. El ensayo de fase 2b con liraglutida en la enfermedad de Alzheimer y el de fase 3 con exenatida en Parkinson también fueron negativos en los desenlaces primarios. Comparamos cuatro explicaciones alternativas (estadio de la enfermedad, exposición central insuficiente, epidemiología confundida y mediación vascular indirecta) y la evidencia capaz de discriminarlas. Formalizamos una Hipótesis de la Ventana Metabólica (Metabolic Window Hypothesis), generadora de hipótesis y dependiente del estadio, con predicciones y condiciones de refutación explícitas, y proponemos un Marco de Precisión Incretina-Neurodegeneración como estructura provisional para investigación estratificada por biomarcadores. Ni la semaglutida ni la tirzepatida son terapias establecidas para prevenir o tratar la enfermedad de Alzheimer o la demencia.

Palabras clave: Enfermedad de Alzheimer; agonistas del receptor de GLP-1; semaglutida; tirzepatida; E-value.

1. Introduction

Alzheimer's disease (AD) is defined biologically by amyloid-β deposition and tau aggregation, but these processes unfold within an environment shaped by vascular injury, neuroinflammation, impaired cerebral energy metabolism and systemic metabolic disease. The clustering of type 2 diabetes, obesity, hypertension and chronic kidney disease with cognitive impairment has motivated the question of whether pharmacological correction of systemic metabolism can alter the trajectory of brain ageing.

Incretin-based therapies occupy the centre of that question. Semaglutide and tirzepatide produce glycaemic, ponderal and cardiovascular effects of a magnitude not previously achieved pharmacologically, and observational datasets and secondary analyses of cardiovascular trials have associated GLP-1RA exposure with fewer dementia diagnoses (Nørgaard et al.; Wang et al.; Seminer et al.). The phase 3 EVOKE and EVOKE+ programme, which randomized 3,808 people with amyloid-confirmed early symptomatic AD to oral semaglutide or placebo, provided the first large phase 3 direct test of oral semaglutide for disease modification in early symptomatic AD and found no effect on clinical progression (Cummings et al.). A pharmacological class associated with lower future dementia incidence thus failed, in its largest direct test, to modify early symptomatic disease.

Stage dependence is not new in AD therapeutics; it was articulated for anti-amyloid strategies as the principle of testing the right drug, at the right target, at the right stage (Sperling et al.). Recent post-EVOKE analyses have already emphasized metabolic

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

heterogeneity, translational limitations and the need to reconsider GLP-1 receptor agonism beyond treatment of symptomatic AD. Gupta et al. highlighted the intersection between metabolic syndrome and disease stage, whereas Stringhi et al. emphasized the translational gap between strong mechanistic rationale and negative randomized trials. Commentary in the British Journal of Clinical Pharmacology has likewise underscored the growing, but inherently confounded, real-world evidence linking GLP-1RA exposure to lower dementia risk (Lam et al.).

The present review builds on this emerging literature by moving beyond narrative reconciliation. It quantitatively examines the sensitivity of key observational associations to unmeasured confounding using E-values; it explicitly compares four competing explanations for the prevention–treatment divergence, each linked to a discriminating test; and it formalizes a hypothesis-generating, stage-dependent metabolic hypothesis of incretin neuroprotection, the Metabolic Window Hypothesis, with explicit predictions and refutation conditions, together with a provisional operational scaffold for research.

The principal contribution of this review is not the introduction of stage dependence as a new concept, but its integration with quantitative confounding sensitivity analysis, central drug-exposure constraints and competing causal explanations. This approach transforms the prevention–treatment discrepancy into a set of empirically distinguishable research questions, rather than attributing the negative trials to disease stage alone.

2. Methods

2.1 Design, questions and reporting

This is a structured critical review, not a systematic review; no pooled estimate is produced. The preparation and critical appraisal of the review were informed by the quality domains of the Scale for the Assessment of Narrative Review Articles (SANRA) (Baethge et al.); SANRA is a quality-assessment instrument rather than a formal reporting guideline, and no claim of adherence to a reporting standard is made. The review addressed four questions: (Q1) Is GLP-1RA exposure associated with

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

lower incidence of dementia or AD? (Q2) Do these agents modify progression of established cognitive impairment or AD? (Q3) Does dual GIP/GLP-1 agonism have neurological effects distinguishable from selective GLP-1 agonism? (Q4) Which explanations best account for divergence between the answers to Q1 and Q2?

2.2 Search and eligibility

PubMed/MEDLINE was searched from inception to 23 September 2026 combining incretin drug terms with dementia, AD, mild cognitive impairment (MCI), cognition, neurodegeneration and Parkinson's disease, supplemented by mechanistic terms (insulin resistance, amyloid, tau, neuroinflammation, blood–brain barrier, cerebrovascular disease). Reference lists of pivotal trials and recent reviews were screened, and post-EVOKE articles retrieved by the PubMed search or identified through reference lists were examined. Pivotal studies were defined as phase 3 or other large randomized trials, randomized-trial meta-analyses, and the largest target-trial emulations or registry analyses addressing Q1–Q3. Full strings are given in Supplementary Table S1. Records were screened by title and abstract by the first author alone, and duplicates between the core and mechanistic searches were removed during screening. Because this is a structured critical review rather than a systematic review, no PRISMA-type flow diagram is provided; Supplementary Table S1 reports the records retrieved by each search and the works included in the synthesis, and full-text screening counts are not reported. The two PubMed strategies retrieved 3,818 records before deduplication (core search, n = 843; mechanistic search, n = 2,975); the synthesis cites 27 works, 11 studies are tabulated in Table 2 and 6 observational studies contribute estimates to the E-value analysis (Table 1). These are retrieval and inclusion counts, not counts of articles assessed in full text. The documented PubMed strategies were rerun on 8 October 2026 to check reproducibility, and, to identify registered trials not yet published, ClinicalTrials.gov was additionally searched on the same date (Supplementary Table S1). Randomized trials, systematic reviews and meta-analyses, target-trial emulations, large cohorts and human biomarker studies were prioritized. Studies were entered in Table 2 when they reported a human cognitive or dementia outcome for an incretin-based therapy,

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

were a randomized trial, a meta-analysis of randomized trials, a target-trial emulation or a large cohort or registry analysis, and contributed an estimate needed to answer Q1–Q3. Commentaries and editorials were used only to contextualize the post-EVOKE debate and did not contribute data. Preclinical studies were used only to establish biological plausibility. Retracted reports were excluded. Parkinson's disease trials were included as indirect evidence because they test the same pharmacological hypothesis.

2.3 Evidence domains

Evidence was interpreted across four domains to prevent observational associations from being read as efficacy: (A) biological plausibility; (B) observational prevention; (C) randomized prevention signals, that is, dementia or cognitive outcomes recorded as secondary or safety outcomes in metabolic and cardiovascular trials; and (D) randomized disease modification in people with cognitive impairment or neurodegenerative disease.

2.4 Appraisal of observational evidence

Observational studies were appraised qualitatively against six bias domains: confounding by indication and channelling; healthy-adherer and healthy-initiator effects; immortal-time and prevalent-user bias; detection and surveillance bias; misclassification of dementia subtype in electronic health records; and competing risk of death. The domain-level judgements for each observational study, made qualitatively by a single reviewer, are reported in Supplementary Table S3; they are not ROBINS-I ratings. Effects too large to be causally plausible (a judgement of biological and pharmacological plausibility, not a statistical criterion), or comparable reductions in non-neurological outcomes such as all-cause mortality, were treated as signals that warrant concern about residual confounding.

Robustness to unmeasured confounding was quantified with E-values (VanderWeele and Ding), defined as the minimum strength of association, on the risk-ratio scale, that an unmeasured confounder would need with both exposure and outcome, conditional on measured covariates, to explain away an observed association. The

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

analysis is an exploratory E-value sensitivity analysis of key observational studies, not an exhaustive analysis of every observational estimate identified by the search. To limit selective presentation, inclusion rules were established for this exploratory analysis during manuscript preparation (no protocol was registered): estimates from the observational studies in Table 2 that compared an incretin-based therapy with a comparator, whose 95% confidence interval excluded the null and that could reasonably be expressed on a relative-risk scale; all-cause mortality was included where reported alongside a cognitive outcome, as an indicator of residual population differences. Where a study reported several eligible comparisons, each was retained, except that for a study reporting a series of comparator arms against the same exposure (Wang et al.) the insulin and other GLP-1RA contrasts reported in the study abstract were analysed as illustrative contrasts; they are not the numerical extremes across the seven comparators, all of which are listed in Supplementary Table S2. During revision, one active-comparator cohort in which a GLP-1RA was the reference arm (Pai et al.) was added; this extension was not part of the original selection, and its reported direction (SGLT2i relative to GLP-1RA) is preserved. Supplementary Table S2 lists every estimate considered, with the decision to include or exclude it and the reason. For protective estimates the reciprocal was used. Hazard ratios were treated as approximate risk ratios, which is reasonable given the low cumulative incidence of dementia over the reported follow-up periods; E-values derived from hazard ratios are therefore approximate and exploratory, not exact transformations of cumulative incidence. E-values were computed for the point estimate and for the confidence limit closest to the null. The formula and worked calculations are given in Supplementary Material S4.

3. Biological Rationale and the Central Exposure Problem

3.1 Metabolic dysfunction and GLP-1 signalling

Brain insulin resistance provides the principal bridge between systemic metabolism and AD. Systematic reviews have identified abnormalities in glucose utilization, ATP production and synaptic function in AD, while concluding that the relationship remains

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

incompletely established and that tools to measure cerebral insulin signalling in living humans are lacking (Andrade et al.; Rhea et al.). These data justify metabolic investigation but not reductionist labels such as "type 3 diabetes"; metabolic dysfunction is better understood as one modifiable node in a heterogeneous pathological network whose contribution varies across individuals and disease stages.

GLP-1 receptors are expressed in several brain regions, and experimental activation has been associated with neuronal survival, synaptic signalling, mitochondrial function, attenuated microglial activation and modulation of amyloid and tau pathways. A systematic review found a relatively consistent preclinical signal for reduced amyloid and hyperphosphorylated tau pathology, contrasting with markedly inconsistent clinical findings (Corcoran et al.). Biological plausibility is a necessary condition for therapeutic development; it is not evidence of clinical efficacy.

3.2 Does the drug reach the target?

An explanation for the negative trials independent of disease stage is insufficient central exposure. The most direct human evidence comes from EVOKE/EVOKE+ itself: in the cerebrospinal fluid substudy, the geometric mean CSF semaglutide concentration was 0.08 nmol/L (coefficient of variation 109.7%; n=98), against a model-estimated geometric mean steady-state systemic concentration of 17.21 nmol/L (n=1,863) (Cummings et al.). These data demonstrate measurable but very low and highly variable CSF exposure relative to systemic concentrations; whether this corresponds to adequate receptor engagement in disease-relevant parenchymal regions remains unknown, because CSF concentration is a marker of central exposure, not of parenchymal concentration, receptor occupancy or neuronal pharmacodynamic effect. Pharmacological data are consistent with this picture. Semaglutide is a large, albumin-bound peptide. In rodents, labelled semaglutide accessed circumventricular organs and periventricular regions but did not broadly cross the blood–brain barrier into deeper parenchyma (Gabery et al.). The formulation tested in EVOKE adds a pharmacokinetic consideration: oral semaglutide is co-formulated with the absorption enhancer sodium N-(8-[2-hydroxybenzoyl]amino)

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

caprylate, and under recommended dosing conditions its estimated mean oral bioavailability is approximately 0.8%, with substantial within-subject variability and dependence on post-dose fasting time, water volume and body weight (Overgaard et al.). Low oral bioavailability is not, by itself, evidence of low brain exposure, because the oral regimen is dosed to achieve therapeutic systemic concentrations; for this question the human CSF data are more informative than bioavailability. If central exposure does not produce sufficient pharmacological target engagement, a negative trial in symptomatic AD cannot distinguish an inappropriate disease stage from an inadequate exposure strategy, and a pre-symptomatic trial with the same agent could fail for the same reason.

4. Observational Evidence of Dementia Prevention

4.1 Principal studies

Nørgaard et al. combined three cardiovascular outcome trials (15,820 participants with type 2 diabetes) with a Danish nationwide cohort of 120,054 individuals. Dementia occurred less frequently with GLP-1RAs than placebo in the pooled trials (HR 0.47; 95% CI 0.25–0.86), and each additional year of exposure in the registry was associated with lower dementia incidence (HR 0.89; 95% CI 0.86–0.93). Because dementia was not a trial outcome and events were few, the randomized component is a prevention signal rather than a test of prevention. Its dementia events were later incorporated into the meta-analysis of Seminer et al. (Section 5), so the two randomized signals are not independent replications.

Wang et al. emulated target trials in more than one million people with type 2 diabetes using United States electronic health records. Semaglutide was associated with a lower incidence of first AD diagnosis compared with insulin (HR 0.33; 95% CI 0.21–0.51) and with other GLP-1RAs (HR 0.59; 95% CI 0.37–0.95). The contrast between the two comparators is informative: the association weakens considerably when the comparator shares the treatment class and much of the indication.

Tang et al. conducted parallel emulations of GLP-1RAs, SGLT2 inhibitors and other second-line agents. Both GLP-1RAs (HR 0.67; 95% CI 0.47–0.96) and SGLT2

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

inhibitors (HR 0.57; 95% CI 0.43–0.75) were associated with lower risk of AD and related dementias relative to other agents, but the two classes did not differ from each other (HR 0.97; 95% CI 0.72–1.32). That two pharmacologically unrelated classes sharing cardiometabolic benefits produce indistinguishable associations suggests that much of the signal may reflect shared cardiometabolic effects, or shared prescribing patterns, rather than a GLP-1-specific action on AD pathology. This does not exclude a direct neural effect, but it limits the extent to which the observational literature can be taken as support for one.

A complementary active-comparator cohort (Pai et al.), using the TriNetX US network, matched 165,566 patients per arm and found slightly lower incidence of dementia with SGLT2 inhibitors than with GLP-1RAs (HR 0.92; 95% CI 0.87–0.98), together with lower all-cause mortality (HR 0.88; 95% CI 0.85–0.91). Differences in contrast direction, population and outcome definition preclude reading this as a direct replication or refutation of the semaglutide-specific emulations, but it does not suggest a GLP-1-specific advantage over another cardioprotective class.

4.2 Qualitative bias appraisal

Newer, costlier incretin drugs are preferentially prescribed to patients with better access to care, higher adherence and lower frailty, whereas insulin is channelled towards patients with advanced disease. Patients who start and persist on injectable therapies with gastrointestinal adverse effects are selected for motivation and functional capacity. Frequent specialist contact can increase dementia recording, while prescribers may avoid initiating new agents in patients with emerging cognitive complaints. Dementia codes have limited specificity for AD, so outcomes labelled as AD plausibly include vascular and mixed dementias, the forms most likely to respond to cardiometabolic improvement. Differential mortality introduces competing-risk distortion when follow-up is short.

A large apparent effect on outcomes a drug is not expected to influence to that degree is a useful warning sign. In the target-trial emulations of Younis et al., tirzepatide was associated with lower dementia risk than semaglutide (HR 0.69; 95% CI 0.48–0.99) and also with lower all-cause mortality (HR 0.72; 95% CI 0.58–0.90) over two years.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Against SGLT2 inhibitors, the same analyses reported a dementia HR of 0.66 (95% CI 0.47–0.93) alongside an all-cause mortality HR of 0.29 (95% CI 0.23–0.37) for tirzepatide, and a dementia HR of 0.71 (95% CI 0.63–0.78) alongside an all-cause mortality HR of 0.53 (95% CI 0.49–0.57) for semaglutide. Mortality differences of this size, between two incretin agents and, still more, a 71% relative reduction against another cardioprotective class, are not supported by randomized evidence and are consistent with residual differences between the compared populations, which would also affect dementia estimates.

4.3 E-value sensitivity analysis

The E-value analysis (Table 1) shows substantial variation in robustness. For Tang et al., the point-estimate E-value is 2.35, but the upper confidence limit yields only 1.25, so the interval remains compatible with an association that relatively modest residual confounding could displace. For Wang et al., the comparison with insulin gives E-values of 5.51 and 3.33, whereas the comparison with other GLP-1RAs, which reduces although it does not eliminate confounding by class and indication, gives 2.78 and 1.29. Across all seven comparators, the numerical extremes are sulfonylureas (HR 0.31; E-values 5.91 and 3.59) and SGLT2 inhibitors (HR 0.60; E-values 2.72 and 1.16); that the weakest association is against another cardioprotective class mirrors the pattern reported by Tang et al. The high values for the insulin comparison should be read alongside the strong channelling of insulin towards patients with more severe diabetes, comorbidity and frailty, since several such factors acting together could approach the required strength. The Nørgaard registry estimate yields 1.50 and 1.36, but because it is expressed per additional year of exposure, these values describe the sensitivity of that duration parameter rather than the confounding strength needed to explain the cumulative effect of long-term treatment. For Younis et al., the dementia association is highly sensitive at its confidence bound (E-value 1.11), and the mortality association yields 2.12 and 1.46. Against SGLT2 inhibitors, the dementia association yields 2.40 and 1.36, whereas the mortality association yields 6.35 and 4.85; for semaglutide versus SGLT2 inhibitors, the corresponding values are 2.17 and 1.88 for dementia and 3.18 and 2.90

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

for mortality. That the largest E-values in this study belong to an outcome the drug is not expected to influence to that degree illustrates that high E-values can coexist with substantial residual population differences. In the active-comparator cohort of Pai et al., in which SGLT2 inhibitors were associated with slightly lower dementia and mortality risks than GLP-1RAs, the E-values are 1.39 and 1.16 for dementia and 1.53 and 1.43 for mortality.

The da Silva et al. associations for tirzepatide versus semaglutide are the most extreme, with E-values of 16.15 and 8.56 for MCI and 12.81 and 7.15 for dementia. Conventionally, such values indicate considerable resistance to a single unmeasured confounder. Here, however, the magnitude of the association is itself difficult to reconcile with the pharmacological and randomized evidence, and it could reflect several simultaneous sources of systematic bias, including calendar-time differences, treatment selection, duration of observable history, diagnostic opportunity, healthy-initiator effects and outcome misclassification. E-values do not address such combinations.

Taken together, the sensitivity analysis neither establishes that the observational findings are spurious nor establishes causality. It shows that most of the key signals examined are sensitive to modest confounding at their confidence bounds, and it illustrates why statistical resistance to a single hypothetical confounder should not be equated with biological credibility. In addition, four of the six observational studies in Table 1 (Wang et al., Younis et al., da Silva et al. and Pai et al.) drew on the same TriNetX network, so their source populations may overlap and their concordance should not be read as independent replication.

Table 1: Exploratory E-value sensitivity analysis of key observational associations

Study and comparison	Effect estimate (95% CI)	E-value, point estimate	E-value, CI limit closest to null	Interpretation
A. Incretin-based therapy compared with a non-incretin or another incretin comparator				
Nørgaard et al.: GLP-1RA exposure duration, dementia	HR 0.89 per additional year (0.86–0.93)	**1.50**	**1.36**	Modest; interpretation limited because exposure is modelled per year

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Study and comparison	Effect estimate (95% CI)	E-value, point estimate	E-value, CI limit closest to null	Interpretation
Wang et al.: semaglutide vs insulin, first AD diagnosis	HR 0.33 (0.21–0.51)	**5.51**	**3.33**	Numerically resistant to a single confounder; comparator vulnerable to severity channelling
Wang et al.: semaglutide vs other GLP-1RAs, first AD diagnosis	HR 0.59 (0.37–0.95)	**2.78**	**1.29**	Confidence bound susceptible to relatively modest confounding
Tang et al.: GLP-1RA vs other second-line agents, AD and related dementias	HR 0.67 (0.47–0.96)	**2.35**	**1.25**	Confidence bound susceptible to modest confounding
da Silva et al.: tirzepatide vs semaglutide, MCI	RR 0.12 (0.06–0.22)	**16.15**	**8.56**	Very large E-value, but extreme effect size raises concern about multiple structural biases not addressed by the metric
da Silva et al.: tirzepatide vs semaglutide, dementia	RR 0.15 (0.09–0.26)	**12.81**	**7.15**	As above
Younis et al.: tirzepatide vs semaglutide, dementia	HR 0.69 (0.48–0.99)	**2.26**	**1.11**	Highly sensitive at the confidence bound
Younis et al.: tirzepatide vs semaglutide, all-cause mortality	HR 0.72 (0.58–0.90)	**2.12**	**1.46**	Large mortality association warrants scrutiny of residual population differences; it does not by itself demonstrate confounding
Younis et al.: tirzepatide vs SGLT2i, dementia	HR 0.66 (0.47–0.93)	**2.40**	**1.36**	Confidence bound susceptible to modest confounding
Younis et al.: tirzepatide vs SGLT2i, all-cause mortality	HR 0.29 (0.23–0.37)	**6.35**	**4.85**	Large mortality association raises concern about comparability and residual bias; it does not by itself demonstrate confounding
Younis et al.: semaglutide vs SGLT2i, dementia	HR 0.71 (0.63–0.78)	**2.17**	**1.88**	Moderate robustness; observational association in which residual confounding remains possible
Younis et al.: semaglutide vs SGLT2i, all-cause mortality	HR 0.53 (0.49–0.57)	**3.18**	**2.90**	Large mortality association warrants scrutiny of residual between-group differences
B. Active-comparator cohort with GLP-1RA as the reference arm (added during revision)				
Pai et al.: SGLT2i vs GLP-1RA (GLP-1RA as reference), dementia	HR 0.92 (0.87–0.98)	**1.39**	**1.16**	Association favours SGLT2i; highly sensitive at the confidence bound
Pai et al.: SGLT2i vs GLP-1RA (GLP-1RA as reference), all-cause mortality	HR 0.88 (0.85–0.91)	**1.53**	**1.43**	Concurrent mortality difference; competing risk and channelling require scrutiny

Source: prepared by the author(s) from the published estimates of the cited studies.

Note: inclusion rules established for this exploratory analysis in Section 2.4; the decision for every estimate considered is listed in Supplementary Table S2. E-values calculated according to VanderWeele and Ding, using the reciprocal for estimates below 1.0. Hazard ratios were treated as

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

approximations to risk ratios for this exploratory sensitivity analysis; the approximation is most defensible when cumulative outcome incidence is low and hazards are not strongly non-proportional, and where outcomes are common or hazards vary substantially over time the hazard ratio may diverge from a cumulative risk ratio, so the resulting E-value should be interpreted cautiously. E-values quantify sensitivity to unmeasured confounding only and do not address selection bias, differential surveillance, outcome misclassification, calendar-time effects, exposure misclassification or combinations of biases. AD, Alzheimer's disease; CI, confidence interval; GLP-1RA, glucagon-like peptide-1 receptor agonist; HR, hazard ratio; MCI, mild cognitive impairment; RR, relative risk; SGLT2i, sodium-glucose cotransporter-2 inhibitor.

Table 1 shows that robustness to unmeasured confounding varies widely across the key associations examined: most lose robustness at their confidence bounds, whereas the largest E-values belong to tirzepatide–semaglutide estimates whose very magnitude is difficult to reconcile with pharmacological evidence.

5. Randomized Prevention Signals from Secondary and Safety Outcomes

Seminer et al. meta-analysed 26 randomized trials comprising 164,531 participants and found fewer dementia or cognitive-impairment events with GLP-1RAs (odds ratio 0.55; 95% CI 0.35–0.86). This signal and the pooled analysis of Nørgaard et al. are partly overlapping rather than independent, because the later meta-analysis incorporates dementia-event data from cardiovascular trials included in the earlier pooled analysis. Randomization removes confounding by indication, but inference remains limited: dementia was captured through adverse-event reporting rather than systematic assessment, follow-up was short relative to a process that evolves over decades, and dementia subtype was not ascertained, leaving open the possibility that any reduction concerns vascular rather than Alzheimer-type dementia. These data justify dedicated prevention trials; they do not demonstrate prevention of AD.

6. Randomized Evidence in Established Neurodegenerative Disease

6.1 Early trials and ELAD

A systematic review found that randomized studies of GLP-1RAs in AD measuring amyloid, tau or cognition did not show consistent differences between treatment and control (Liang et al.). In the phase 2b ELAD trial, 204 people with mild-to-moderate AD without diabetes received liraglutide or placebo; the primary outcome, cerebral glucose metabolic rate, did not differ, and a favourable secondary executive-function

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

measure was not accompanied by consistent benefit on clinical or functional scales (Edison et al.).

6.2 EVOKE and EVOKE+

EVOKE and EVOKE+ were double-blind, placebo-controlled phase 3 trials conducted in 40 countries, enrolling participants aged 55–85 years with MCI or mild dementia due to amyloid-confirmed AD; EVOKE+ additionally permitted significant small-vessel pathology (Cummings et al.). In total, 3,808 participants received oral semaglutide, titrated to 14 mg daily, or placebo. The difference in Clinical Dementia Rating-Sum of Boxes (CDR-SB) change at week 104 was −0.08 points (95% CI −0.35 to 0.20; p=0.57) in EVOKE and 0.10 points (95% CI −0.17 to 0.38; p=0.46) in EVOKE+. Secondary functional and cognitive endpoints were also neutral, and pooled analyses did not show delayed progression from MCI to dementia. The confidence intervals are narrow enough to exclude benefits of the size generally regarded as clinically relevant.

Two further observations bear on the hypotheses discussed below. First, in the cerebrospinal fluid substudy semaglutide reduced core AD-related biomarkers, including p-tau181 and p-tau217, by up to approximately 10% (Cummings et al.). Second, EVOKE and EVOKE+ were not prospectively enriched for metabolic vulnerability: in the pooled population of 3,808 participants, 518 (13.6%) had type 2 diabetes and 591 (15.5%) had a body-mass index of 30 kg/m² or greater (Cummings et al.). Diagnosed type 2 diabetes and obesity are, however, imperfect proxies for insulin resistance, visceral adiposity, metabolic syndrome or hyperinsulinaemia, none of which was used for enrolment, so these figures show the absence of metabolic enrichment rather than low metabolic vulnerability. The trials therefore provide strong evidence regarding oral semaglutide in an early symptomatic AD population not selected for metabolic status, but substantially less information about whether a highly insulin-resistant or metabolically enriched subgroup might respond differently. This distinction should not be interpreted as evidence that such a responder subgroup exists; rather, it identifies a question that the trials were not specifically designed or powered to resolve.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

6.3 Target engagement without clinical benefit

Statistically significant biomarker change with a precisely neutral clinical outcome illustrates that target engagement is not disease modification. Three interpretations remain open: the biological effect may be too small, since a reduction of about 10% in fluid biomarkers is modest; the modified pathways may contribute to AD biology without being rate-limiting at the symptomatic stage; or the intervention may come too late, after synaptic and neuronal loss has made the clinical trajectory largely independent of those pathways.

6.4 Indirect evidence from Parkinson's disease

In a phase 3 trial, weekly exenatide for 96 weeks did not slow motor progression in Parkinson's disease compared with placebo, overturning encouraging phase 2 data (Vijiaratnam et al.). A phase 2 trial of lixisenatide in early Parkinson's disease reported a small difference in motor-score progression at 12 months, with frequent gastrointestinal adverse events (Meissner et al.). These trials test the translational plausibility of the drug class in another neurodegenerative disease; they do not establish efficacy or inefficacy in AD and do not confirm any particular explanation for the EVOKE result. They show, however, the same pattern of encouraging early or small trials not reproduced by a larger phase 3 trial in symptomatic patients.

6.5 Clinical meaningfulness

A meta-analysis of 14 randomized trials (1,260 participants) of GLP-1RAs in non-diabetic individuals with AD, MCI or Parkinson's disease found a small effect on global cognition (standardized mean difference 0.14; 95% CI 0.01–0.27), with an estimated probability of approximately 1% that it represented a clinically important benefit, and more frequent gastrointestinal adverse events and discontinuations (Elghanam and Kim). An umbrella review by Ballum et al. identified nine systematic reviews and meta-analyses addressing cognition in adults with type 2 diabetes; the authors reported convergence toward an association between GLP-1RA exposure and less cognitive decline, with favourable findings in some cognitive measures and non-significant results in others, and concluded that larger controlled trials are required.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Because umbrella reviews synthesize partially overlapping meta-analyses, these findings summarize the existing literature rather than adding an independent efficacy signal.

7. Tirzepatide: Does Dual GIP/GLP-1 Agonism Change the Equation?

No published randomized trial of tirzepatide with a cognitive or dementia outcome was identified through the search date, and the supplementary ClinicalTrials.gov search identified no registered randomized trial of tirzepatide with a cognitive outcome (Supplementary Table S1). Tirzepatide is the only dual GIP/GLP-1 receptor agonist with human cognitive data, and that evidence is exclusively observational. In the TriNetX comparison by da Silva et al. (44,470 matched patients per group), tirzepatide was associated with lower recorded incidence of MCI (RR 0.12; 95% CI 0.06–0.22) and dementia (RR 0.15; 95% CI 0.09–0.26), whereas the estimate for AD was imprecise and compatible with no difference (RR 0.48; 95% CI 0.22–1.01). Younis et al., using the same network, reported a more modest dementia association alongside the mortality difference discussed in Section 4.2.

The magnitude reported by da Silva et al. warrants particular caution. An 88% relative reduction in recorded MCI between two incretin-based therapies over short real-world follow-up is difficult to reconcile with the magnitude of effects observed in randomized metabolic and neurological studies. Potential explanations include differential calendar time, treatment selection, follow-up duration, diagnostic opportunity, outcome misclassification and healthy-initiator effects. That the most specific outcome (AD) shows the weakest and least precise association is also compatible with non-specific bias. Consequently, these findings are insufficient to establish either intrinsic neuroprotection from dual GIP/GLP-1 agonism or neurological superiority of tirzepatide over semaglutide. If a real difference existed, it could arise from GIP receptor signalling, greater weight loss, greater improvement in insulin resistance or differential cardiovascular effects, which only randomized comparison with target-engagement measures could separate.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Table 2: Key human evidence on incretin-based therapies and cognitive outcomes (selection criteria in Section 2.2)

Study	Design	Population	Exposure	Main finding	Principal threat to validity	Domain
Nørgaard et al., 2022	Pooled RCTs + national registry	T2D	GLP-1RAs	Pooled RCT HR 0.47; registry HR 0.89 per year	Dementia not a trial outcome; few events	**C, B**
Wang et al., 2024	Target-trial emulation	T2D without prior AD	Semaglutide	HR 0.33 vs insulin; HR 0.59 vs other GLP-1RAs	Channelling; EHR outcome specificity	**B**
Tang et al., 2025	Target-trial emulation	T2D	GLP-1RA, SGLT2i	HR 0.67 vs other agents; HR 0.97 vs SGLT2i	Residual confounding; shared mechanisms	**B**
Pai et al., 2024	Active-comparator matched cohort	T2D, age ≥50 years	SGLT2i vs GLP-1RA (both with metformin)	Dementia HR 0.92; mortality HR 0.88 (SGLT2i relative to GLP-1RA)	Channelling; exposure required for ≥12 months; competing risk	**B**
Younis et al., 2026	Target-trial emulations	T2D	Tirzepatide vs semaglutide	Tirzepatide: dementia HR 0.69 vs semaglutide and 0.66 vs SGLT2i; semaglutide vs SGLT2i: dementia HR 0.71; mortality HR 0.72, 0.29 and 0.53	Mortality imbalance	**B**
da Silva et al., 2026	Matched retrospective cohort	T2D	Tirzepatide vs semaglutide	MCI RR 0.12; dementia RR 0.15; AD RR 0.48 (CI includes 1)	Implausible magnitude; calendar-time bias	**B**
Seminer et al., 2025	Meta-analysis of 26 RCTs	164,531 participants	GLP-1RAs	OR 0.55 for dementia/cognitive impairment	Adverse-event ascertainment; short follow-up; partly overlaps with Nørgaard et al.	**C**
Edison et al., 2026	Phase 2b RCT	Mild–moderate AD, no diabetes	Liraglutide	Primary outcome negative	Sample size	**D**
Cummings et al., 2026	Two phase 3 RCTs	3,808 with amyloid-confirmed early AD	Oral semaglutide	No CDR-SB effect; CSF biomarkers reduced ≤10%	Central exposure; stage; no metabolic enrichment	**D**
Vijiaratnam et al., 2025	Phase 3 RCT	Parkinson's disease	Exenatide	No effect on motor progression	Disease stage	**D (indirect)**
Elghanam & Kim, 2026	Meta-analysis of 14 RCTs	AD, MCI, PD	GLP-1RAs	SMD 0.14; ~1% probability of clinical importance	Heterogeneity; small trials	**D**

Source: compiled by the author(s) from the cited studies.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

AD, Alzheimer's disease; CDR-SB, Clinical Dementia Rating-Sum of Boxes; CSF, cerebrospinal fluid; EHR, electronic health record; HR, hazard ratio; MCI, mild cognitive impairment; OR, odds ratio; PD, Parkinson's disease; RCT, randomized controlled trial; RR, relative risk; SGLT2i, sodium-glucose cotransporter-2 inhibitor; SMD, standardized mean difference; T2D, type 2 diabetes. Domains as defined in Section 2.3.

Table 2 organizes the principal studies by evidence domain and shows that favourable findings derive mainly from observational analyses and secondary outcomes of randomized metabolic trials, whereas the trials designed to test disease modification were negative.

8. Four Competing Explanations

Favourable and unfavourable findings come from different populations, designs and outcomes, so the literature need not contain a contradiction. Four non-exclusive explanations can account for the divergence, and they should be weighed explicitly rather than in favour of the one most encouraging for further investment (Table 3). Stage dependence holds that metabolic intervention may alter risk while metabolic and vascular contributors remain dominant, but not once neurodegeneration is established. Insufficient central exposure holds that the agents tested may not reach AD-relevant brain regions in pharmacologically meaningful concentrations. Confounded epidemiology holds that much or all of the observational signal may reflect channelling, healthy-initiator effects and detection bias. Indirect vascular mediation holds that incretin drugs may genuinely reduce dementia through cardiovascular, renal and cerebrovascular improvement, with benefit concentrated in vascular and mixed dementia rather than in AD pathology.

Table 3: Competing explanations and the evidence that would discriminate between them

Explanation	Supporting evidence	Evidence against or unresolved	Discriminating test
Stage dependence	Prevention signals arise pre-symptomatically; treatment trials fail in symptomatic AD and PD	Not tested prospectively; EVOKE included MCI, an early symptomatic stage	Pre-symptomatic RCT with biomarker-defined risk, stratified by stage
Insufficient central exposure	Very low and variable CSF semaglutide concentration relative to plasma in EVOKE; limited parenchymal access in animals	Peripheral effects may act indirectly; CSF biomarkers changed in EVOKE	Central target-engagement studies (CSF exposure plus receptor-occupancy or pharmacodynamic measures); brain-penetrant analogues; injectable high-dose arms

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Explanation	Supporting evidence	Evidence against or unresolved	Discriminating test
Confounded epidemiology	Low confidence-limit E-values; implausible effect sizes; mortality imbalances; similar SGLT2i associations	Randomized secondary/safety-event signal in metabolic and cardiovascular trials	RCT with dementia as prespecified outcome; negative-control outcomes in cohorts
Indirect vascular mediation	GLP-1RA ≈ SGLT2i; both improve cardiometabolic health	AD-specific data sparse and imprecise	Subtype-specific ascertainment (amyloid/tau biomarkers, MRI vascular burden) in prevention trials

Source: prepared by the author(s).

Table 3 indicates that the four explanations are not mutually exclusive and that each leaves a distinct observable signature, so that a single well-designed trial could generate evidence on several of them simultaneously.

The explanations imply different priorities. Because the explanations can coexist, they shift priorities by degree rather than in an all-or-none manner. If residual and structural confounding explains most of the observational signal, the rationale for large prevention trials would be substantially weakened unless independent mechanistic or target-engagement evidence emerges. If central exposure is limiting, pharmacology becomes a higher priority than trial population. If vascular mediation accounts for much of the benefit, incretin therapy would sit primarily within the vascular-risk prevention agenda, without thereby excluding an additional AD-specific component. A pre-symptomatic AD prevention trial is most strongly justified if stage dependence contributes materially and adequate central exposure can be demonstrated.

9. The Metabolic Window Hypothesis of Incretin Neuroprotection

The following model is a hypothesis-generating interpretation developed after consideration of the currently available evidence and should not be interpreted as an explanation established by EVOKE; its falsifiability is a property of its formulation, not evidence that it is true. We formalize the stage-dependent explanation as follows (Figure 1):

> *The neurological benefit of incretin-based therapies, if any, is stage dependent. It is greatest while metabolic and vascular abnormalities remain major, modifiable contributors to cognitive risk and neuronal injury is limited, and it declines as irreversible neurodegenerative burden accumulates.*

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

The hypothesis extends a stage-dependence principle previously articulated in anti-amyloid therapeutic development (Sperling et al.) to metabolic repurposing, linking the window to measurable metabolic and biomarker states rather than to chronological age or clinical stage alone. It rests on three independently testable assumptions: (a) the agent achieves adequate central or effective peripheral target engagement; (b) metabolic dysfunction contributes causally, not merely as a marker, to AD-related neurodegeneration in at least a subset of individuals; and (c) this contribution is relatively larger before substantial synaptic and neuronal loss.

If the hypothesis is correct: (P1) incretin therapy slows decline on a preclinical cognitive composite relative to placebo in a pre-symptomatic trial; (P2) the effect is larger with greater baseline insulin resistance or visceral adiposity and lower baseline neurodegeneration markers; (P3) the effect is attenuated with increasing baseline plasma p-tau217 or tau-PET burden, producing a detectable interaction; and (P4) if a direct neural component is claimed, the cognitive effect is at least partly independent of weight loss and blood-pressure change in mediation analysis. The hypothesis should be considered refuted, or reduced to the vascular-mediation explanation, if an adequately powered pre-symptomatic trial with confirmed target engagement shows no effect on cognitive trajectory, if no gradient of effect across metabolic or neurodegeneration strata is found, or if benefit is confined to vascular outcomes and small-vessel imaging markers without change in AD-specific biomarkers or cognition. The hypothesis also makes predictions that separate it from the three alternative explanations. Unlike insufficient central exposure, it predicts that benefit depends on baseline metabolic and neurodegenerative state even when target engagement is confirmed; unlike confounded epidemiology, it requires a treatment-by-phenotype interaction within randomized comparisons; and unlike indirect vascular mediation, it requires change in AD-specific biomarkers or cognition that is not fully explained by vascular outcomes.

Importantly, the hypothesis does not depend exclusively on completion of a large, dedicated prevention trial. Several of its predictions, in particular the interactions between treatment effect, metabolic vulnerability and baseline neurodegenerative burden (P2 and P3), could be examined through prespecified secondary analyses of

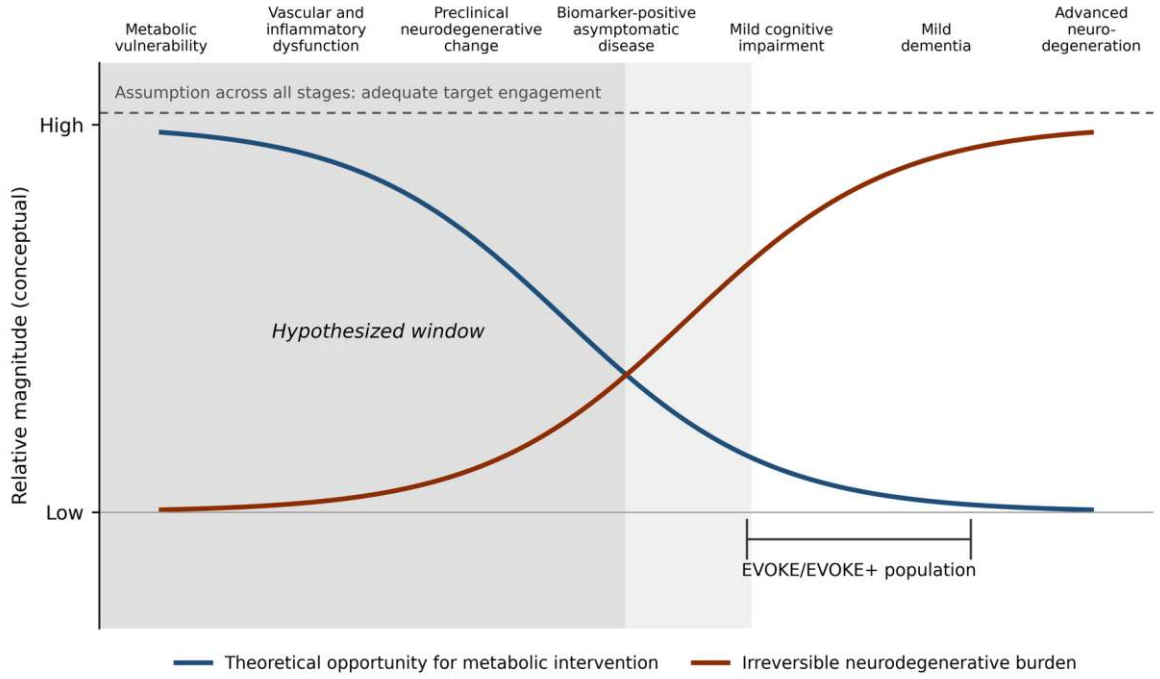
CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

randomized trials with stored biospecimens, biomarker-rich prospective cohorts, or harmonized individual-participant datasets. Such analyses would not replace a definitive prevention trial, but they could provide earlier evidence supporting or weakening the stage-dependent model and help determine whether a large prospective trial is scientifically justified.

Figure 1: The Metabolic Window Hypothesis of Incretin Neuroprotection



Source: prepared by the author(s).

Conceptual representation of two opposing gradients across the metabolic-to-neurodegenerative continuum: the theoretical opportunity for metabolic intervention declines as irreversible neurodegenerative burden increases. The shaded area marks the hypothesized window; the bracket indicates the disease stages enrolled in EVOKE/EVOKE+; the dashed line denotes the assumption of adequate target engagement across all stages. Curves are illustrative and not derived from data. The model is hypothesis-generating and is not a treatment algorithm.

10. Precision Incretin-Neurodegeneration Framework

Because AD is heterogeneous, benefit confined to a subset would be diluted by broad enrolment. EVOKE and EVOKE+ were not prospectively enriched for metabolic vulnerability and therefore neither confirm nor exclude such dilution. The framework is proposed for pre-symptomatic and prodromal research settings, where this question can be tested prospectively; it is a research tool, not a clinical

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

recommendation. To move towards reproducibility without assigning clinical weights that the evidence cannot support, each domain is expressed as a categorical research variable (Table 4, Figure 2). The framework is a proposed operational scaffold, not yet a fully operationalized classification: it has not been validated, and the rules for combining several variables within a domain and for handling discordant results (for example, positive plasma p-tau217 with an indeterminate Aβ42/40 ratio, or elevated NfL without MRI atrophy) must be prespecified in each protocol before enrolment. Laboratory cut-offs are deliberately not fixed here, because they depend on assay, platform and population; they should be prespecified in each protocol. No validated pharmacogenomic test identifies cognitive responders to any incretin agent, and GLP1R and GIPR variants are listed only as candidate modifiers.

Table 4: Proposed operational scaffold of the Precision Incretin-Neurodegeneration Framework

Domain	Proposed operational variables	Research categories
Metabolic	Type 2 diabetes and/or HOMA-IR above a prespecified threshold; visceral adiposity	Low / intermediate / high vulnerability
AD biology	Plasma p-tau217; Aβ42/40 ratio	Negative / indeterminate / positive
Neurodegeneration	NfL; MRI atrophy; tau-PET where available	Low / intermediate / high burden
Vascular	White-matter hyperintensities; prior stroke; hypertension; chronic kidney disease	Low / high burden
Functional reserve	Frailty index; grip strength; activities of daily living	Preserved / reduced
Genetic	APOE ε4 (candidate: GLP1R, GIPR variants)	Non-carrier / heterozygous / homozygous

Source: prepared by the author(s).

Table 4 expresses each framework domain as categorical research variables; reproducible classification will additionally require assay-dependent cut-offs and combination rules, which are deliberately left to be prespecified in each protocol.

Three research phenotypes are defined by provisional rules. The early metabolic-predominant phenotype requires high metabolic vulnerability, low neurodegeneration burden and preserved functional reserve. The mixed phenotype combines high metabolic vulnerability with intermediate AD-biology or neurodegeneration

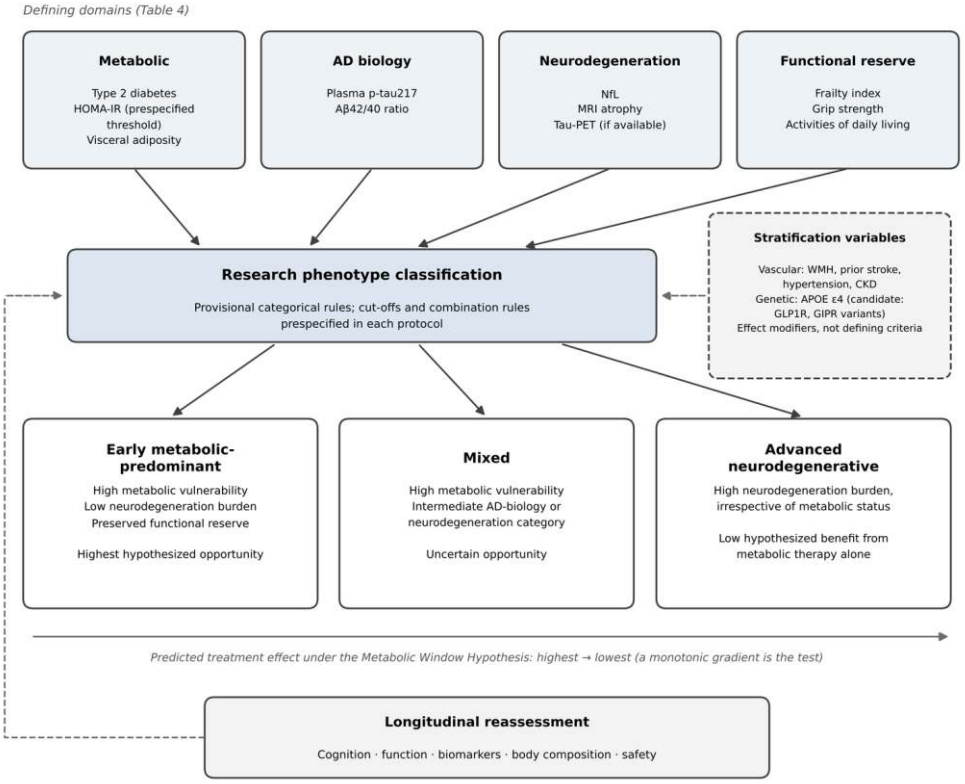
CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

categories. The advanced neurodegenerative phenotype is defined by high neurodegeneration burden irrespective of metabolic status. EVOKE/EVOKE+ enrolled people who had already passed the pre-symptomatic phase, but participants were not classified by neurodegeneration burden, so the trial population cannot be mapped onto this phenotype. Vascular and genetic categories serve as stratification and effect-modification variables rather than defining criteria, so that the vascular-mediation explanation is tested rather than assumed. Because the rules are stated before any such trial, the Metabolic Window Hypothesis predicts a monotonic gradient of treatment effect across the three phenotypes, and the absence of such a gradient would count against it. Future validation should evaluate the framework's reproducibility, prespecified biomarker classification rules and ability to identify treatment-effect heterogeneity in prospective randomized studies. Predictive utility should be assessed independently of the data used to develop the classification.

Figure 2: Precision Incretin-Neurodegeneration Framework



Source: prepared by the author(s).

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Four defining domains (metabolic, AD biology, neurodegeneration and functional reserve) converge on a provisional categorical phenotype classification (Table 4), with vascular and genetic variables entering only as stratification and effect-modification variables. The classification assigns participants to one of three research phenotypes along a gradient of predicted treatment effect; longitudinal reassessment of cognition, function, biomarkers, body composition and safety feeds back into classification. The framework guides research enrolment and stratification and is not a validated treatment-selection algorithm.

11. A Trial to Discriminate Between Explanations

Repeating EVOKE in another symptomatic population would not be informative. A more informative trial would enrol adults aged approximately 55–75 years with type 2 diabetes or marked insulin resistance, visceral adiposity, preserved independence, normal cognition or very early MCI, and elevated biological risk of AD, screened by plasma p-tau217 above a validated threshold and confirmed by amyloid PET or CSF in all participants or, at minimum, in a prespecified stratum, so that any benefit can be attributed to AD-specific prevention rather than to vascular or metabolic effects in people without AD pathology. Three arms are proposed: subcutaneous semaglutide and tirzepatide at the highest tolerated doses, and matching placebo injection, all on protocolized optimized metabolic care. Injectable formulations maximize systemic and potential central exposure. Randomization would be stratified by APOE ε4, diabetes status, baseline p-tau217 tertile, insulin-resistance index and MRI cerebrovascular burden, and a mechanistic substudy would measure cerebrospinal fluid drug concentrations and pharmacodynamic markers to provide a direct measure of CSF exposure and an indirect assessment of central pharmacological engagement; because CSF concentration does not establish parenchymal concentration or receptor occupancy, receptor-occupancy imaging or other target-engagement methods should be added where available.

The primary endpoint would be change in a validated preclinical cognitive composite; secondary endpoints would include conversion to MCI and dementia with biomarker-based subtype adjudication, CDR-SB, fluid biomarkers, MRI atrophy and white-matter hyperintensities, body composition and sarcopenia, cardiovascular events and functional independence. The primary estimand should follow ICH E9(R1), addressing intercurrent events such as discontinuation for gastrointestinal intolerance. Because gastrointestinal effects and weight loss make participant

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

unblinding likely, outcome assessors should be independent of treating clinicians and blinding integrity should be assessed formally.

Critically, the trial should be sized for the treatment-by-phenotype interaction rather than only for the average effect, because the interaction is the direct test of the hypothesis. Formal calculation requires specifying the composite's variance, the minimal clinically important difference, alpha, power, attrition and the interaction magnitude, which should be derived from natural-history cohorts. Interaction tests typically require several times the sample size of main-effect tests, and follow-up of probably four to five years would be needed, which argues for multinational collaboration or an adaptive platform design.

12. Clinical and Translational Implications

Current evidence does not justify prescribing semaglutide or tirzepatide for the purpose of preventing or treating dementia outside clinical research. For patients with approved metabolic or cardiovascular indications, possible neurological effects remain a legitimate research question but should not be presented as an established benefit or used to justify treatment.

Three translational considerations follow. First, a biomarker acquires clinical value only when it identifies a relevant phenotype, changes a decision and improves a patient-relevant outcome; work on implementing pharmacogenomics in a universal health system has shown that this translation depends on workflows, governance and equitable access, not on test availability (Floriano and Menezes), and EVOKE shows from the opposite direction that biomarker modification cannot substitute for clinical benefit. Multimodal risk models, including machine-learning approaches, may assist in integrating heterogeneous clinical and biomarker data; previous work has highlighted both the diagnostic potential of artificial intelligence and the need to address validation and responsible clinical implementation (Floriano et al.). Second, weight loss, a therapeutic goal in obesity, is a marker of poor prognosis in older adults with cognitive impairment, and incretin therapies suppress appetite and were less well tolerated in neurodegenerative-disease trials (Elghanam and Kim); prevention

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

studies should therefore measure lean mass, grip strength, gait speed, nutritional status and frailty, not body weight alone. Third, should efficacy be demonstrated in a defined subgroup, implementation in older adults with multimorbidity would require structured medication management, including reconciliation, monitoring of gastrointestinal and nutritional adverse effects and coordination across specialties, a role for which clinical pharmacists are well placed.

13. Limitations

This is a structured critical review with study selection by a single reviewer, a bibliographic search restricted to one primary database, supplemented by citation tracking and a trial-registry search, so that statements about the absence of evidence are limited to the sources searched, and a domain-based qualitative bias appraisal (Supplementary Table S3) rather than a formal risk-of-bias instrument or certainty grading, so selection and interpretation bias cannot be excluded. The E-value analysis is exploratory and restricted to key studies, relies on published estimates, approximates hazard ratios as risk ratios and addresses only unmeasured confounding. Several observational sources share a single electronic-health-record network, so their populations may overlap. The two randomized prevention signals are partly overlapping and derive from secondary or safety outcomes. Much of the prevention evidence uses electronic-health-record diagnoses with limited subtype specificity, tirzepatide data have short follow-up, and several key sources were published in 2026 with limited independent scrutiny. The Metabolic Window Hypothesis and the Precision Incretin-Neurodegeneration Framework are unvalidated conceptual proposals intended to structure research, not to guide care.

14. Conclusions

Biological plausibility, observational association, biomarker change and clinical efficacy are not interchangeable forms of evidence, and the incretin–dementia literature illustrates the cost of treating them as such. GLP-1RAs are repeatedly associated with lower dementia incidence, but exploratory E-value analysis shows that most of the key associations examined are sensitive to modest confounding at

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

their confidence bounds, and the largest ones, for tirzepatide, are of a magnitude that warrants particular caution. The largest direct test, in 3,808 people with early symptomatic AD, showed no clinical slowing despite modest biomarker change, and the phase 3 exenatide trial in Parkinson's disease, an indirect comparator, followed the same pattern.

Disease stage, inadequate central exposure, confounded epidemiology and indirect vascular mediation can each explain the gap between prevention signals and treatment failure, and current evidence does not allow them to be ranked with confidence. The Metabolic Window Hypothesis, a hypothesis-generating post-EVOKE interpretation, formalizes the stage-dependent explanation with explicit predictions and refutation conditions, and the Precision Incretin-Neurodegeneration Framework, still a provisional scaffold, and the proposed trial are designed to help discriminate among all four. The informative question is no longer whether semaglutide or tirzepatide treat AD, but which biologically defined individual, at which point on the metabolic-to-neurodegenerative continuum and with what degree of target engagement, has a clinically meaningful probability of benefit.

The principal scientific contribution of this review is a falsifiable research framework for distinguishing stage-dependent metabolic benefit from inadequate target engagement, residual epidemiological confounding and indirect vascular mediation. Its value lies in guiding future hypothesis testing rather than establishing clinical efficacy or identifying validated treatment responders.

Declarations

Ethics approval. Not applicable. This is a review of previously published literature and did not involve human participants or identifiable data.

Consent for publication. Not applicable.

Funding. No specific funding was received for this work.

Author contributions. João Fernandes Floriano: conceptualization, methodology, literature search, formal analysis (E-value calculations), writing of the original draft and review and editing of the manuscript. Maria Elizabeth Menezes: critical scientific

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

review of the methodology and interpretation, and review and editing of the manuscript. Both authors approved the final version and agree to be accountable for its content.

Conflicts of interest. The authors declare no conflicts of interest related to the subject of this manuscript.

Data availability. No original data were generated. Search strings are provided in Supplementary Table S1; the E-value formula and worked calculations are given in Supplementary Material S4 and can be reproduced from the published estimates in Table 1; Supplementary Tables S2 and S3 document the E-value inclusion decisions and the bias-domain appraisal.

Use of generative artificial intelligence. During preparation of this manuscript, João Fernandes Floriano used ChatGPT (OpenAI) and Claude (Anthropic) as assistive tools for language editing, structural organization and revision, refinement of literature-search terminology, cross-checking of bibliographic information and preparation of conceptual figures. The tools were not used to generate or analyse original patient-level data. All scientific statements, numerical results, interpretations, references and final wording were reviewed and verified by the authors against the cited sources. The authors take full responsibility for the accuracy, integrity and content of the manuscript.

Works Cited

Andrade, L. J. O., et al. "Brain Insulin Resistance and Alzheimer's Disease: A Systematic Review." *Dementia & Neuropsychologia*, vol. 18, 2024, e20230032, https://doi.org/10.1590/1980-5764-DN-2023-0032.

Baethge, C., et al. "SANRA: A Scale for the Quality Assessment of Narrative Review Articles." *Research Integrity and Peer Review*, vol. 4, 2019, 5, https://doi.org/10.1186/s41073-019-0064-8.

Ballum, H., et al. "The Effect of Glucagon-Like Peptide 1 (GLP-1) Receptor Agonists on Cognition: A Systematic Review of Systematic Reviews and Meta-Analyses." *Journal of Affective Disorders*, vol. 402, 2026, 121310, https://doi.org/10.1016/j.jad.2026.121310.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Corcoran, E., et al. "The Effects of GLP-1 Receptor Agonists on Alzheimer's Pathophysiology: A Systematic Review." *Molecular and Cellular Neuroscience*, vol. 137, 2026, 104091, https://doi.org/10.1016/j.mcn.2026.104091.

Cummings, J. L., et al. "Efficacy and Safety of Oral Semaglutide 14 mg (Flexible Dose) in Early-Stage Symptomatic Alzheimer's Disease (Evoke and Evoke+): Two Phase 3, Randomised, Placebo-Controlled Trials." *The Lancet*, vol. 407, no. 10544, 2026, pp. 2167–79, https://doi.org/10.1016/S0140-6736(26)00459-9.

da Silva, A. M. P., et al. "Tirzepatide versus Semaglutide for the Prevention of Mild Cognitive Impairment, Dementia, and Alzheimer's Disease in Type 2 Diabetes: A Real-World, Retrospective Cohort Study." *Journal of Diabetes and Its Complications*, vol. 40, no. 5, 2026, 109306, https://doi.org/10.1016/j.jdiacomp.2026.109306.

Edison, P., et al. "Liraglutide in Mild to Moderate Alzheimer's Disease: A Phase 2b Clinical Trial." *Nature Medicine*, vol. 32, no. 1, 2026, pp. 353–61, https://doi.org/10.1038/s41591-025-04106-7.

Elghanam, Y., and E. Kim. "Evaluating the Clinical Effects of GLP-1 Receptor Agonists for Alzheimer's and Parkinson's Diseases Using Minimal Clinically Important Difference: Systematic Review and Meta-Analysis." *Archives of Pharmacal Research*, vol. 49, no. 5, 2026, pp. 691–712, https://doi.org/10.1007/s12272-026-01615-y.

Floriano, João Fernandes, and Maria Elizabeth Menezes. “Da Evidência à Política Pública: A Integração da Farmacogenômica ao Sistema Único de Saúde como Modelo Normativo para Sistemas de Cobertura Universal em Países de Renda Média.” *Veredas do Direito*, vol. 23, no. 12, 2026, e237600, https://doi.org/10.18623/rvd.v23.7600.

Floriano, João Fernandes, et al. “Aplicações da Inteligência Artificial no Diagnóstico em Saúde: Avanços Tecnológicos, Desafios e Perspectivas Futuras.” *Veredas do Direito*, vol. 23, no. 10, 2026, e237118, https://doi.org/10.18623/rvd.v23.7118.

Gabery, S., et al. "Semaglutide Lowers Body Weight in Rodents via Distributed Neural Pathways." *JCI Insight*, vol. 5, no. 6, 2020, e133429, https://doi.org/10.1172/jci.insight.133429.

Gupta, P., et al. "GLP-1 Receptor Agonists, Metabolic Syndrome, and Alzheimer's Disease: Lessons and Opportunities from the EVOKE Trials." *Journal of Neuroendocrinology*, vol. 38, no. 8, 2026, e70246, https://doi.org/10.1111/jne.70246.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Lam, Y. W., et al. "GLP-1 Receptor Agonists and Reduced Dementia Risk: Real-World Evidence Stacks Up." *British Journal of Clinical Pharmacology*, 2026, pp. 1–4, https://doi.org/10.1002/bcp.70451.

Liang, Y., et al. "Clinical Evidence for GLP-1 Receptor Agonists in Alzheimer's Disease: A Systematic Review." *Journal of Alzheimer's Disease Reports*, vol. 8, no. 1, 2024, pp. 777–89, https://doi.org/10.3233/ADR-230181.

Meissner, W. G., et al. "Trial of Lixisenatide in Early Parkinson's Disease." *The New England Journal of Medicine*, vol. 390, no. 13, 2024, pp. 1176–85, https://doi.org/10.1056/NEJMoa2312323.

Nørgaard, C. H., et al. "Treatment with Glucagon-Like Peptide-1 Receptor Agonists and Incidence of Dementia: Data from Pooled Double-Blind Randomized Controlled Trials and Nationwide Disease and Prescription Registers." *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, vol. 8, no. 1, 2022, e12268, https://doi.org/10.1002/trc2.12268.

Overgaard, R. V., et al. "Clinical Pharmacokinetics of Oral Semaglutide: Analyses of Data from Clinical Pharmacology Trials." *Clinical Pharmacokinetics*, vol. 60, no. 10, 2021, pp. 1335–48, https://doi.org/10.1007/s40262-021-01025-x.

Pai, Y.-W., et al. “Association of Sodium-Glucose Cotransporter 2 Inhibitors with Risk of Incident Dementia and All-Cause Mortality in Older Patients with Type 2 Diabetes: A Retrospective Cohort Study Using the TriNetX US Collaborative Networks.” *Diabetes, Obesity and Metabolism*, vol. 26, no. 11, 2024, pp. 5420–30, https://doi.org/10.1111/dom.15918.

Rhea, E. M., et al. "State of the Science on Brain Insulin Resistance and Cognitive Decline Due to Alzheimer's Disease." *Aging and Disease*, vol. 15, no. 4, 2024, pp. 1688–725, https://doi.org/10.14336/AD.2023.0814.

Seminer, A., et al. "Cardioprotective Glucose-Lowering Agents and Dementia Risk: A Systematic Review and Meta-Analysis." *JAMA Neurology*, vol. 82, no. 5, 2025, pp. 450–60, https://doi.org/10.1001/jamaneurol.2025.0360.

Sperling, R. A., et al. "Testing the Right Target and Right Drug at the Right Stage." *Science Translational Medicine*, vol. 3, no. 111, 2011, 111cm33, https://doi.org/10.1126/scitranslmed.3002609.

Stringhi, R., et al. "GLP-1 Receptor Agonists for Alzheimer's Disease: Lessons from Trials and Translational Challenges." *Cell Reports Medicine*, vol. 7, no. 8, 2026, 102918, https://doi.org/10.1016/j.xcrm.2026.102918.

Tang, H., et al. "GLP-1RA and SGLT2i Medications for Type 2 Diabetes and Alzheimer Disease and Related Dementias." *JAMA Neurology*, vol. 82, no. 5, 2025, pp. 439–49, https://doi.org/10.1001/jamaneurol.2025.0353.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

VanderWeele, T. J., and P. Ding. "Sensitivity Analysis in Observational Research: Introducing the E-Value." *Annals of Internal Medicine*, vol. 167, no. 4, 2017, pp. 268–74, https://doi.org/10.7326/M16-2607.

Vijiaratnam, N., et al. "Exenatide Once a Week versus Placebo as a Potential Disease-Modifying Treatment for People with Parkinson's Disease in the UK: A Phase 3, Multicentre, Double-Blind, Parallel-Group, Randomised, Placebo-Controlled Trial." *The Lancet*, vol. 405, no. 10479, 2025, pp. 627–36, https://doi.org/10.1016/S0140-6736(24)02808-3.

Wang, W., et al. "Associations of Semaglutide with First-Time Diagnosis of Alzheimer's Disease in Patients with Type 2 Diabetes: Target Trial Emulation Using Nationwide Real-World Data in the US." *Alzheimer's & Dementia*, vol. 20, no. 12, 2024, pp. 8661–72, https://doi.org/10.1002/alz.14313.

Younis, A., et al. "Target Trial Emulations for Tirzepatide, Semaglutide and SGLT2-inhibitors for Dementia in Patients with Type 2 Diabetes: Real World Evidence from a Retrospective Cohort Study." *Diabetes Research and Clinical Practice*, vol. 235, 2026, 113083, https://doi.org/10.1016/j.diabres.2026.113083.

Supplementary Material

Supplementary Table S1: PubMed/MEDLINE search strings (last run 23 September 2026)

Core search:

("Glucagon-Like Peptide-1 Receptor Agonists"[Mesh] OR "GLP-1 receptor agonist*"[tiab] OR GLP-1RA*[tiab] OR semaglutide[tiab] OR liraglutide[tiab] OR dulaglutide[tiab] OR exenatide[tiab] OR lixisenatide[tiab] OR tirzepatide[tiab] OR "GIP/GLP-1"[tiab]) AND ("Alzheimer Disease"[Mesh] OR "Dementia"[Mesh] OR "Cognitive Dysfunction"[Mesh] OR alzheimer*[tiab] OR dementia[tiab] OR "mild cognitive impairment"[tiab] OR "cognitive decline"[tiab] OR neurodegenerat*[tiab] OR "Parkinson Disease"[Mesh])

Records retrieved (core search): n = 843

Mechanistic search:

(semaglutide[tiab] OR liraglutide[tiab] OR tirzepatide[tiab] OR "GLP-1"[tiab]) AND ("insulin resistance"[tiab] OR "brain insulin"[tiab] OR amyloid[tiab] OR tau[tiab] OR "p-

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

tau"[tiab] OR neuroinflammation[tiab] OR microglia[tiab] OR "blood-brain barrier"[tiab] OR cerebrovascular[tiab])

Records retrieved (mechanistic search): n = 2,975

Searches were last updated on 23 September 2026. Record counts were taken from the PubMed search history on the date of the final search, and the search history file was retained to allow audit. Together, the two strategies retrieved 3,818 records before deduplication. The archived search history documents the strategies, dates, times and counts, but not the individual records or screening decisions, so September duplicates and full-text counts cannot be derived from it.

Study selection: records were screened by title and abstract by the first author alone against the criteria in Section 2.2; as this is a structured critical review, no PRISMA-type flow diagram is provided and full-text counts are not reported. The synthesis cites 27 works; 11 studies are tabulated in Table 2, of which 6 observational studies contribute estimates to the E-value analysis (Table 1). Post-EVOKE articles were identified through the PubMed search and reference lists.

Search update (PubMed, 8 October 2026), a separate run of the same strategies: core search, 854 records; mechanistic search, 3,002 records; 354 PMIDs retrieved by both, giving 3,502 unique records. These are retrieval counts for the rerun, not screening counts for the original review.

Trial-registry search (ClinicalTrials.gov, 8 October 2026): intervention “tirzepatide”; condition “dementia OR Alzheimer OR cognitive impairment OR cognition OR cognitive decline”. Records retrieved: n = 1 (NCT07677865, a completed observational study). No registered randomized trial of tirzepatide with a cognitive outcome was identified.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Supplementary Table S2: Observational estimates considered for the E-value analysis: inclusion decisions and reasons

Study and comparison	Estimate (95% CI)	Included in E-value analysis	Reason
Nørgaard et al., 2022: registry, GLP-1RA exposure per additional year, dementia	HR 0.89 (0.86–0.93)	**Yes**	Observational; incretin exposure; CI excludes null
Nørgaard et al., 2022: pooled RCTs, GLP-1RA vs placebo, dementia	HR 0.47 (0.25–0.86)	**No**	Randomized estimate; the E-value addresses unmeasured confounding in observational data
Wang et al., 2024: semaglutide vs insulin, first AD diagnosis	HR 0.33 (0.21–0.51)	**Yes**	Illustrative contrast reported in the source abstract (strongest association named by the authors)
Wang et al., 2024: semaglutide vs other GLP-1RAs, first AD diagnosis	HR 0.59 (0.37–0.95)	**Yes**	Illustrative contrast reported in the source abstract (weakest association named by the authors); comparator shares class and indication
Wang et al., 2024: semaglutide vs the five other comparators, first AD diagnosis	Metformin 0.38 (0.24–0.59); DPP-4i 0.40 (0.26–0.63); SGLT2i 0.60 (0.37–0.98); sulfonylureas 0.31 (0.20–0.48); thiazolidinediones 0.43 (0.26–0.70) (Figure 2A of the source)	**No**	Not analysed; the two contrasts reported in the source abstract were retained as illustrative (Section 2.4). Sulfonylureas and SGLT2i are the numerical extremes
Tang et al., 2025: GLP-1RA vs other second-line agents, ADRD	HR 0.67 (0.47–0.96)	**Yes**	Observational; incretin exposure; CI excludes null
Tang et al., 2025: SGLT2i vs other second-line agents, ADRD	HR 0.57 (0.43–0.75)	**No**	Exposure is not an incretin-based therapy
Tang et al., 2025: GLP-1RA vs SGLT2i, ADRD	HR 0.97 (0.72–1.32)	**No**	CI includes the null
Younis et al., 2026: tirzepatide vs semaglutide, dementia	HR 0.69 (0.48–0.99)	**Yes**	Observational; CI excludes null
Younis et al., 2026: tirzepatide vs semaglutide, all-cause mortality	HR 0.72 (0.58–0.90)	**Yes**	Mortality reported alongside a cognitive outcome (indicator of residual differences)
Younis et al., 2026: tirzepatide vs SGLT2i, dementia	HR 0.66 (0.47–0.93)	**Yes**	Observational; CI excludes null (added in revision)
Younis et al., 2026: tirzepatide vs SGLT2i, all-cause mortality	HR 0.29 (0.23–0.37)	**Yes**	Mortality reported alongside a cognitive outcome (added in revision)
Younis et al., 2026: semaglutide vs SGLT2i, dementia	HR 0.71 (0.63–0.78)	**Yes**	Observational; CI excludes null (added in revision)
Younis et al., 2026: semaglutide vs SGLT2i, all-cause mortality	HR 0.53 (0.49–0.57)	**Yes**	Mortality reported alongside a cognitive outcome (added in revision)

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Study and comparison	Estimate (95% CI)	Included in E-value analysis	Reason
Pai et al., 2024: SGLT2i vs GLP-1RA, dementia	HR 0.92 (0.87–0.98)	**Yes**	Active-comparator cohort with GLP-1RA as reference; CI excludes null; added during revision, direction preserved
Pai et al., 2024: SGLT2i vs GLP-1RA, all-cause mortality	HR 0.88 (0.85–0.91)	**Yes**	Mortality reported alongside a cognitive outcome; added during revision
Pai et al., 2024: SGLT2i vs DPP-4i	HR 0.62 (0.59–0.65)	**No**	Neither arm is an incretin-based therapy
da Silva et al., 2026: tirzepatide vs semaglutide, MCI	RR 0.12 (0.06–0.22)	**Yes**	Observational; CI excludes null
da Silva et al., 2026: tirzepatide vs semaglutide, dementia	RR 0.15 (0.09–0.26)	**Yes**	Observational; CI excludes null
da Silva et al., 2026: tirzepatide vs semaglutide, AD	RR 0.48 (0.22–1.01)	**No**	CI includes the null
Seminer et al., 2025; Edison et al., 2026; Cummings et al., 2026; Vijiaratnam et al., 2025; Meissner et al., 2024; Elghanam and Kim, 2026	Randomized estimates	**No**	Randomized evidence; outside the scope of an analysis of unmeasured confounding
Ballum et al., 2026	Umbrella review	**No**	Synthesizes overlapping reviews; no single study-level relative-risk estimate
Lam et al., 2026; Gupta et al., 2026; Stringhi et al., 2026	Commentary or narrative review	**No**	No original estimate

Inclusion rules established for this exploratory analysis in Section 2.4. ADRD, Alzheimer's disease and related dementias; CI, confidence interval; EHR, electronic health record; HR, hazard ratio; MCI, mild cognitive impairment; RCT, randomized controlled trial; RR, relative risk; SGLT2i, sodium-glucose cotransporter-2 inhibitor.

Supplementary Table S3: Qualitative appraisal of observational studies across the six bias domains

Study	Confounding by indication / channelling	Healthy initiator / adherer	Immortal time / prevalent user	Detection / surveillance	Outcome misclassification	Competing risk of death
Nørgaard et al., 2022 (registry)	Serious concern	Serious concern	Some concern (cumulative exposure modelling)	Some concern	Serious concern (registry diagnoses, all-cause dementia)	Some concern
Wang et al., 2024	Serious concern (insulin comparator); moderate (GLP-1RA comparator)	Serious concern	Low–some concern (new-user emulation)	Serious concern	Serious concern (EHR-coded AD)	Some concern (3-year follow-up)

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/09/2026 - **Accepted:** 13/10/2026
Vol: 13.21
DOI: 10.66104/pp7zdq81
Pages: 1-36

Study	Confounding by indication / channelling	Healthy initiator / adherer	Immortal time / prevalent user	Detection / surveillance	Outcome misclassification	Competing risk of death
Tang et al., 2025	Some concern (active comparators)	Serious concern	Low–some concern (new-user emulation)	Some concern	Serious concern (ADRD codes)	Some concern
Younis et al., 2026	Serious concern	Serious concern (mortality imbalance)	Low–some concern (new-user emulation)	Some concern	Serious concern (EHR-coded dementia)	Serious concern (2-year follow-up, large mortality differences)
Pai et al., 2024	Serious concern (treatment selection despite matching)	Some–serious concern (≥12-month persistence requirement)	Serious concern (exposure required for ≥12 months before follow-up)	Some–serious concern	Serious concern (EHR-coded dementia)	Serious concern (concurrent mortality difference)
da Silva et al., 2026	Serious concern (calendar time, prescribing)	Serious concern	Some concern	Serious concern (diagnostic opportunity)	Serious concern (MCI coding)	Some concern (short follow-up)

Qualitative judgements by a single reviewer (low, some or serious concern) across the six domains described in Section 2.4; these are not ROBINS-I assessments and no numerical score was applied. EHR, electronic health record; ADRD, Alzheimer's disease and related dementias; MCI, mild cognitive impairment.

Supplementary Material S4: E-value calculation

For a protective estimate (RR < 1), the reciprocal RR* = 1/RR was used. The E-value for the point estimate is E = RR* + √[RR* × (RR* − 1)] (VanderWeele and Ding). For the confidence interval, the same formula was applied to the reciprocal of the limit closest to the null (the upper limit for protective estimates); if that limit had crossed 1, the E-value would have been 1.

Worked example (Wang et al., semaglutide vs insulin): RR* = 1/0.33 = 3.03; E = 3.03 + √(3.03 × 2.03) = 3.03 + 2.48 = 5.51. Upper limit 0.51: RR* = 1.96; E = 1.96 + √(1.96 × 0.96) = 1.96 + 1.37 = 3.33.

All values in Table 1 were recomputed with this formula during revision and agree with the published table to two decimal places. Hazard ratios were treated as approximate risk ratios (see Section 2.4).

CC BY