Evaluating the effects of Liraglutide, Empagliflozin and Linagliptin on mild cognitive impairment reversion in patients with type 2 diabetes (LIGHT-MCI): a multicenter, randomized controlled trial with an extension phase

Statistical Analysis Plan

The Data Management and Statistical Analysis Plan is directed to support the aims of the LIGHT-MCI Study

Clinical trials.gov Number	NCT05313529
Protocol version and date	Version 5.0, Nov 17, 2025
SAP version and date	Version 1.0, Nov 4, 2025
SAP version and date	Version 2.0, Dec 12, 2025
SAP version and date	Version 3.0, May 16, 2026
SAP version and date	Version 4.0, July 10, 2026

SAP version history		
Version Date	**SAP Version**	**Details of Changes**
Nov 4, 2025	Version 1.0	
Dec 12, 2025	Version 2.0	Additions to the details of safety outcome, safety population, and analysis of primary outcome and secondary outcomes.
May 16, 2026	Version 3.0	1. Updated the sensitivity analyses to include a tipping-point analysis under alternative MNAR assumptions and to modify the definitions of the best- and worst-case scenarios. 2. Respecified the per-protocol analysis as a separate supportive analysis. 3. Added a post-hoc exploratory analysis of accelerated cognitive decline at week 48.
July 10, 2026	Version 4.0	As the 76-week follow-up has been completed, week 76 data will be included in the secondary outcome analyses.

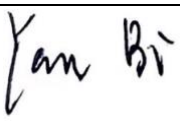
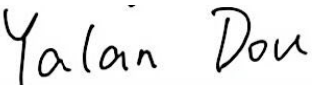
Signatures		
Name and role	**Signature**	**Date**
Prof Yan Bi (Chief investigator)		July 10, 2026
Prof Weili Yan (Senior Statistician)		July 10, 2026
Yalan Dou (Trial Statistician)		July 10, 2026

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation	Term
AD	Alzheimer's disease
MCI	mild cognitive impairment
GLP-1 RAs	glucagon-like peptide-1 receptor agonists
SGLT2	sodium–glucose cotransporter 2
DPP4	dipeptidyl peptidases 4
MRI	magnetic resonance imaging
TEAEs	treatment-emergent adverse events
SAEs	treatment-emergent serious adverse events
MD	mean diffusivity
MoCA	Montreal Cognitive Assessment
RBANS	Repeatable Battery for the Assessment of Neuropsychological Status
IADL	instrumental activity of daily living
IPTW	inverse probability of treatment weighting
FAQ	Functional Activities Questionnaire
MMSE	Mini-Mental State Examination
TMT	Trail Making Test
BADL	Basic activities of daily living
FA	fractional anisotropy
FSL	FMRIB Software Library
DTI	diffusion tensor imaging
FW-WM	fractional volume of free water in white matter
DTI-ALPS	diffusion tensor image analysis along the perivascular space
PVSVF-WM	perivascular space volume fraction in white matter
HbA_{1c}	glycated haemoglobin
AESI	TEAEs of Special Interest
Aβ	Aβamyloid beta
p-tau	phosphorylated tau
NfL	neurofilament light chain
GFAP	glial fibrillary acidic protein
NIA-AA	National Institute on Aging-Alzheimer's Association
ITT	Intent-to-Treat
PP	Per-protocol
SOC	System Organ Clas
PT	Preferred Term
GEE	generalized estimating equation
GLM	generalized linear model
CI	confidence intervals
SD	standard deviation
SPM	Statistical Parametric Mapping
MAR	Missing-at-random
MICE	multiple imputation by chained equations

1. INTRODUCTION

Cognitive dysfunction is increasingly recognized as an emerging complication of diabetes, contributing to profound debilitation and mortality[1,2]. According to disease severity, cognitive dysfunction is often divided into three stages: the preclinical, mild cognitive impairment (MCI), and dementia. Regrettably, the currently available pharmacological therapies for dementia have limited efficacy and undesirable side effects. MCI represents the critical prodromal stage with high dementia risk, making early identification and intervention essential for dementia prevention[3,4]. Despite recommendations for routine cognitive screening, current clinical guidelines lack effective management strategies[5].

Recent evidence highlights that certain antidiabetic agents, particularly glucagon-like peptide-1 receptor agonists (GLP-1 RAs) and sodium–glucose cotransporter 2 (SGLT2) inhibitors, exhibit cardiovascular and renal benefits beyond glycemic control[6-9]. Emerging data suggest that these agents may also improve cognitive function in diabetes, though prospective head-to-head comparisons remain lacking[10-12]. Linagliptin, a dipeptidyl peptidases 4 (DPP4) inhibitor, demonstrated neutral cognitive effects in the CARMELINA and CAROLINA-COGNITION trials, thus serving as an ideal active control[13,14.]

Therefore, we have designed a multicenter, randomized, superiority trial to head-to-head to compare the efficacy of liraglutide and empagliflozin, against linagliptin (as an active comparator) for MCI reversion in type 2 diabetes. The results are expected to address knowledge gaps and inform clinical decision-making by providing robust evidence.

The purpose of this Statistical Analysis Plan is to define the outcome variables, statistical methods, and analysis strategies to address the objectives of the LIGHT-MCI study.

2. STUDY OBJECTIVES AND OUTCOMES

2.1. Study Objectives

2.1.1. Primary objective

The primary objective is to evaluate the effects of liraglutide and empagliflozin, each compared separately to linagliptin, as add-on therapies to metformin monotherapy, on MCI reversion at 48 weeks in adults with type 2 diabetes.

2.1.2. Secondary objectives

The secondary objectives are to evaluate the effects of liraglutide and empagliflozin, each compared separately to linagliptin, on general cognition, cognitive subdomains, basic and instrumental activities of daily living, magnetic resonance imaging (MRI)-derived brain structure and function, blood-based biomarkers for neurodegeneration, olfactory function, glucose and lipid metabolism, anthropometrics, and liver fibrosis and fatty liver levels.

2.1.3. Prespecified safety objectives

The pre-specified safety objectives are to assess the effects of liraglutide, empagliflozin, and linagliptin, as add-on therapy with the metformin monotherapy, on the incidence of treatment-emergent adverse events (TEAEs), treatment-emergent serious adverse events (SAEs), treatment-emergent adverse events leading to study drug discontinuation, and treatment-emergent adverse events of special interest.

2.2. Outcomes

The trial includes a core phase (from baseline to Week 48) and an extension phase (Week 76 will be pre-specified follow-up). Outcomes measured at baseline, Weeks 24, 48 and 76 will be categorized in the SAP as either (a) core phase analysis through Week 48, or (b) extension phase analysis including data through Week 76.

2.2.1. Primary outcome

The primary outcome is MCI reversion at 48 weeks of treatment.

- Description:

  MCI reversion is defined by meeting three criteria: an education-adjusted score of the Montreal Cognitive Assessment (MoCA) ≥26, no cognitive deficits in any explored subdomain (with scores above the cut-off scores in all tests as detailed in Appendix 1), including processing speed, executive function, immediate memory, visuospatial construction ability, language, attention and delayed memory, evaluated by Trail-Making Test, Stroop Color-Word Test and Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), respectively, and preservation of ability to perform instrumental activity of daily living (IADL) with a Functional Activities Questionnaire (FAQ) score <5.
- Time Frame: at week 48 of treatment.
- Type: binary variable (Yes / No).

2.2.2. Secondary outcomes

(1) Change in score of Mini-Mental State Examination (MMSE)

- Description: The MMSE (Chinese version) contains a total of 30 items that assess orientation, registration, attention and calculation, recall, and language. The MMSE score ranges from 0 to 30, with higher scores indicating better cognitive function[15].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(2) Change in score of MoCA

- Description: The MoCA (Chinese version) is a screening instrument for MCI comprised of 30 items to assess multiple cognitive domains (memory recall,

visuospatial abilities, executive functions, attention, language, and orientation to time and place; scores range from 0 to 30, with higher scores indicating better cognitive function). Participants received one additional point if they had a MoCA score < 30 and 12 years or less of formal school education[16].

- Time Frame: from baseline to weeks 24, 48, and 76 of treatment (the week- 76 assessment will be analyzed and reported after the extension phase).
- Type: repeatedly measured continuous variable.

(3) Change in total score of RBANS

- Description: The RBANS (Chinese version) is a brief neuropsychological screening battery with established test-retest reliability and age-appropriate normative data, which consists of 12 task tests assessing 5 cognitive domains, namely immediate memory, visuospatial/ constructional, language, attention and delayed memory, with a score range from 40 to 160. A higher RBANS score reflects a better global cognitive function[17].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(4) Change in the RBANS index score of immediate memory

- Description: The RBANS (Chinese version) is a brief neuropsychological screening battery with established test-retest reliability and age-appropriate normative data. The RBANS includes 12 sub-tests that generate five age-adjusted index scores and a total score. The five indices include immediate memory (consisting of list learning and story memory tests), visuospatial/constructional domain (consisting of figure copying and line orientation tests), language (consisting of picture naming and semantic fluency tests), attention (consisting of digit span and coding tests) and delayed memory (consisting of list recall, story recall, figure recall and list recognition

tests). Generally, a higher index score reflects a better cognitive subdomain function[18].

- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(5) Change in the RBANS index score of visuospatial/constructional

- Description: Measurements and instruments are the same as the RBANS index score of immediate memory.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(6) Change in the RBANS index score of language

- Description: Measurements and instruments are the same as the RBANS index score of immediate memory.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(7) Change in the RBANS index score of attention

- Description: Measurements and instruments are the same as the RBANS index score of immediate memory.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(8) Change in the RBANS index score of delayed memory

- Description: Measurements and instruments are the same as the RBANS index score of immediate memory.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(9) Change in aggregate time to test completion for Trail Making Test (TMT)

- Description: The TMT is a validated timed measure of processing speed, which consists of part A and part B (TMT-A and TMT-B). The time limits for performing the TMT-A and TMT-B are 180 and 300 seconds, respectively. Processing speed is estimated by the aggregate time in seconds to complete TMT-A and TMT-B such that less time indicate faster processing speed[19].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(10) Change in aggregate time to test completion for Victoria Stroop Color-Word Test

- Description: The Victoria Stroop Color-Word Test is a well-known measure of executive function. There are three indices including Stroop-dot, Stroop-word and Stroop-color word in the test. The color task consists of colored dots; the word task comprises ordinary words that are unrelated to the meaning of color; the color–word task consists of words written in color that indicate the meaning of the color, but the color of these words differs from the meaning of the word itself. The participants will be asked to quickly read the color of the dots or Chinese words on the cards. The sum of consuming time taken to read the three cards is used as an index of executive function performance. Less time indicates better executive function[20].

- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(11) Change in Barthel index

- Description: Barthel index will be measured to assess the Basic activities of daily living (BADL), which includes ten items assessing basic self‐care abilities (transfer, bathing, toileting, dressing, feeding). The score ranges from 0 to 100, with higher scores indicating better function, and 100 points indicating complete independence[21].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(12) Change in FAQ score

- Description: The FAQ measures IADLs such as preparing balanced meals and managing personal finances. Sum scores range from 0 to 30, with higher scores indicating worse function[22]
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(13) Change in MRI-derived total brain volume

- Description: Philips 3.0T scanner will be used for magnetic resonance scanning (Ingenia 3.0T, Philips Medical Systems, Eindhoven, The Netherlands). T1 images will be processed using Freesurfer software (version 8.1.0, https://surfer.nmr.mgh.harvard.edu) to calculate intracranial volume, total gray matter volume, total white matter volume, cerebrospinal fluid volume and volumes of 68 cortical and 14 subcortical regions based on the Desikan-

Killiany atlas[23]. The total brain volume will be calculated as the sum of the gray matter and white matter volumes.

- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(14) Change in MRI-derived total gray matter volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(15) Change in MRI-derived total white matter volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(16) Change in MRI-derived cerebrospinal fluid volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(17) Change in MRI-derived frontal cortical gray matter volume

- Description: Measurements and instruments are the same as total brain volume. Frontal gray matter volume will be calculated as the sum of all frontal lobe regions parcellated by FreeSurfer (Desikan-Killiany atlas).

- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(18) Change in MRI-derived parietal cortical gray matter volume

- Description: Measurements and instruments are the same as total brain volume. Parietal gray matter volume will be calculated as the sum of all parietal lobe regions parcellated by FreeSurfer (Desikan-Killiany atlas).
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(19) Change in MRI-derived temporal cortical gray matter volume

- Description: Measurements and instruments are the same as total brain volume. Temporal gray matter volume will be calculated as the sum of all temporal lobe regions parcellated by FreeSurfer (Desikan-Killiany atlas).
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(20) Change in MRI-derived occipital cortical gray matter volume

- Description: Measurements and instruments are the same as total brain volume. Occipital gray matter volume will be calculated as the sum of all occipital lobe regions parcellated by FreeSurfer (Desikan-Killiany atlas).
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(21) Change in MRI-derived insula cortical grey matter volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(22) Change in MRI-derived hippocampal volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(23) Change in MRI-derived parahippocampal volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(24) Change in MRI-derived entorhinal cortex volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(25) Change in MRI-derived inferior parietal lobule volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(26) Change in MRI-derived precuneus volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(27) Change in MRI-derived cuneus volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(28) Change in MRI-derived thalamus volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(29) Change in MRI-derived caudate nucleus volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to week 48, and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(30) Change in MRI-derived putamen volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(31) Change in MRI-derived pallidum volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(32) Change in MRI-derived amygdala volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(33) Change in MRI-derived accumbens volume

- Description: Measurements and instruments are the same as total brain volume.
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(34) Change in MRI-derived white matter hyperintensity volume (WMH)

- Description: Total white matter hyperintensity volume will be measured from FLAIR images using the Lesion Segmentation Tool for Statistical Parametric Mapping (SPM) 12.0[24].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(35) Change in mean fractional anisotropy (FA) in white matter derived from diffusion tensor imaging (DTI)

- Description: Diffusion Toolbox in FMRIB Software Library (FSL, version 6.0.7, http://fsl.fmrib.ox.ac.uk/fsl/fslwiki) will be used to process all DTI-scans, and generate FA maps. FA from DTI assessment is a global microstructural integrity measure, with higher value indicating better integrity[25].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(36) Change in mean diffusivity (MD) in white matter derived from DTI

- Description: Measurement and instruments are the same as FA. MD describes the overall diffusion and motion of water molecules, with higher MD values reflect more water mobility[25].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(37) Change in fractional volume of free water in white matter (FW-WM)

- Description: FW-WM maps will be constructed from preprocessed diffusion-weighted images using a regularized bi-tensor model and the open-source software package Diffusion Imaging in Python (also known as Dipy) algorithm (https://dipy.org/). Elevated FW -WM suggests the stagnation of fluid drainage caused by glymphatic dysfunction[26].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(38) Change in diffusion tensor image analysis along the perivascular space (DTI-ALPS) index

- Description: The ALPS index will be analyzed on DTI images using a method developed and validated by Taoka et al[27]. ALPS index = (mean (Dxxproj, Dxxassoc)) ⁄ (mean (Dyyproj, Dzzassoc)). The ALPS indices in the bilateral hemispheres will be calculated, respectively, and the mean will be used in further analyses. The DTI-ALPS index will be calculated from the diffusivity along the deep medullary vein at the level of the lateral ventricle body, as a measure of perivascular clearance activity in the human brain. It is validated as a non‐invasive proxy for glymphatic functioning, with higher value indicating better function[28].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(39) Change in perivascular space volume fraction in white matter (PVSVF-WM)

- Description: PVSs will be mapped from T2W images using an automated and highly reliable quantification method, following the pipeline of previous studies[29]. The PVS volume fraction (PVSVF) = PVS volume/intracranial volume. The PVSVF-WM is a quantitative metric that reflects the distribution and dilatation of perivascular spaces within cerebral white matter. An elevated volume fraction may suggest impaired clearance function of brain interstitial fluid[30].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(40) Change in functional MRI-derived odor-induced brain activation

- Description: Task-evoked activation will be assessed to estimate scent effects on brain activation. Bilateral parahippocampus, amygdala, piriform cortex, insula, orbitofrontal cortex, and hippocampus in Automated Anatomical

Labeling templates and Brodmann areas 28 and 34 (entorhinal cortices) will be extracted and merged as olfactory regions of interest (ROIs) for further analyses.

- Time Frame: from baseline to the 48 weeks of treatment.
- Type: continuous variable.

(41) Change in resting-state functional MRI-derived brain network connectivity

- Description: The whole-brain network functional connectivity analysis will be carried out using the CONN Toolbox (v.22.v2407)[31]. The Schaefer–Yeo 7-network functional atlas will be used to parcellate the cortex into 400 volumetric functional parcels, and the Melbourne Subcortical Atlas will be used to parcellate the subcortex into 54 functional nuclei, yielding a total of 454 regions. The 400 cortical parcellation will be chosen per previous recommendations. Each cortical region will be assigned to one of seven canonical resting-state functional networks: default mode network (DMN), ventral attention network (VAN), sensorimotor network (SMN), limbic network, visual network, dorsal attention network (DAN) and central executive network (CEN) (also known as frontoparietal). For each participant, the pre-processed resting-state fMRI time series will be spatially averaged across all voxels composing each region, yielding an averaged time series for each region. The Pearson correlation coefficient will be computed between all pairs of regions from the combined cortical and subcortical atlases to provide a measure of functional connectivity, yielding a 454*454 symmetric connectivity matrix for each individual.[32,33]
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(42) Change in plasma amyloid beta (Aβ)42/Aβ40 concentration

- Description: Plasma Aβ40 and Aβ42 concentration will be measured using the Neurology 4-plex E assay kit from Quanterix.
- Time Frame: from baseline to the 48 weeks of treatment.
- Type: continuous variable.

(43) Change in plasma phosphorylated tau (p‐tau)181 concentration

- Description: Plasma p-tau 181 concentration will be measured using the p-Tau181 Advantage PLUS kit from Quanterix.
- Time Frame: from baseline to the 48 weeks of treatment.
- Type: continuous variable.

(44) Change in plasma p‐tau 217 concentration

- Description: Plasma p-tau 217 concentration will be measured using the ALZpath p-Tau217 Advantage PLUS kit from Quanterix.
- Time Frame: from baseline to the 48 weeks of treatment.
- Type: continuous variable.

(45) Change in plasma neurofilament light chain (NfL) concentration

- Description: Plasma NfL concentration will be measured using the Neurology 4-plex E assay kit from Quanterix.
- Time Frame: from baseline to the 48 weeks of treatment.
- Type: continuous variable.

(46) Change in plasma glial fibrillary acidic protein (GFAP) concentration

- Description: Plasma GFAP concentration will be measured using the Neurology 4-plex E assay kit from Quanterix.
- Time Frame: from baseline to the 48 weeks of treatment.
- Type: continuous variable.

(47) Change in olfactory threshold score

- Description: Olfactory Function Assessment by Computerized Testing (OLFACT®, Osmic Enterprises, Inc.) will be administered to assess olfactory threshold, identification, and memory. The olfactory threshold test (score range, 1-14) will be a forced-choice paradigm based on serial dilutions of N-butanol solutions. Higher scores indicate better olfactory function[34,35].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(48) Change in olfactory identification score

- Description: Olfactory Function Assessment by Computerized Testing (OLFACT®, Osmic Enterprises, Inc.) will be administered to assess olfactory threshold, identification, and memory. The olfactory identification (score 0–30) and memory (score 0–20) subtests consisted of two sequential phases will be separated by a 10-min interval. During task A, participants will be performed forced-choice identification of 10 distinct odors. In task B, they will be presented with 20 odors (10 from task A and 10 novel distractors) and required to complete two tasks per odor: identify the odor, then classify it as 'old' or 'new' (relative to task A). Higher scores indicate better olfactory function[34,35].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(49) Change in olfactory memory score

- Description: Olfactory Function Assessment by Computerized Testing (OLFACT®, Osmic Enterprises, Inc.) will be administered to assess olfactory threshold, identification, and memory. The olfactory identification (score 0–30) and memory (score 0-20) subtests will consist of two sequential phases

separated by a 10-min interval. During task A, participants will be performed forced-choice identification of 10 distinct odors. In task B, they will be presented with 20 odors (10 from task A and 10 novel distractors) and required to complete two tasks per odor: identify the odor, then classify it as 'old' or 'new' (relative to task A). Higher scores indicate better olfactory function[34,35].

- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(50) Change in olfactory assessment total score

- Description: The sum of the olfactory threshold, identification, and memory scores will be referred to as the olfactory assessment total scores, which ranged from 0 to 64 points. Higher scores indicate better olfactory function[34,35].
- Time Frame: from baseline to weeks 48 and 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: continuous variable.

(51) Change in glycated haemoglobin (HbA_{1c}) levels from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Repeatedly measured continuous variable)

(52) Change in fasting plasma glucose from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Repeatedly measured continuous variable)

(53) Change in 2-hour postprandial plasma glucose from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(54) Change in fasting insulin from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(55) Change in fasting C-peptide levels from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(56) Change in 2-hour postprandial insulin from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(57) Change in 2-hour postprandial C-peptide levels from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(58) Change in pancreatic islet function (HOMA2-β) estimated from fasting C-peptide via HOMA2 Calculator (v2.2.3, http://www.dtu.ox.ac.uk/homacalculator/) from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(59) Change in insulin sensitivity (HOMA2-S) estimated from fasting C-peptide via HOMA2 Calculator (v2.2.3, http://www.dtu.ox.ac.uk/homacalculator/) from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(60) Change in insulin resistance indexes (HOMA2-IR) estimated from fasting C-peptide via HOMA2 Calculator (v2.2.3, http://www.dtu.ox.ac.uk/homacalculator/) from baseline to weeks 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Continuous measurements)

(61) Change in fasting serum triglyceride levels from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Repeatedly measured continuous variable)

(62) Change in fasting serum total cholesterol levels from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Repeatedly measured continuous variable)

(63) Change in fasting serum high-density lipoprotein-cholesterol levels from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Repeatedly measured continuous variable)

(64) Change in fasting serum low-density lipoprotein-cholesterol levels from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).
(Repeatedly measured continuous variable)

(65) Change in visceral fat area measured by Inbody 720 analyser from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(66) Change in percent body fat measured by Inbody 720 analyser from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(67) Change in liver stiffness measurement quantified by FibroTouch® system from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(68) Change in liver fat controlled attenuation parameter quantified by FibroTouch® system from baseline to weeks 24, 48 and 76 (the week 76 assessment will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(69) Change in blood pressure from baseline to week 48. (Measurements will be performed every 8 weeks until week 76, with data collected after week 48 will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(70) Change in body weight every 8 weeks from baseline to week 48. (Measurements will be performed every 8 weeks until week 76, with data collected after week 48 will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(71) Change in body mass index every 8 weeks from baseline to week 48. (Measurements will be performed every 8 weeks until week 76, with data

collected after week 48 will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(72) Change in waist circumference every 8 weeks from baseline to week 48. (Measurements will be performed every 8 weeks until week 76, with data collected after week 48 will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(73) Change in hip circumference every 8 weeks from baseline to week 48. (Measurements will be performed every 8 weeks until week 76, with data collected after week 48 will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(74) Change in waist/hip ratio every 8 weeks from baseline to week 48. (Measurements will be performed every 8 weeks until week 76, with data collected after week 48 will be analyzed and reported after the extension phase).

(Repeatedly measured continuous variable)

(75) MCI reversion at 76 weeks of treatment.

- Description: Measurements and definition are the same as the primary outcome.
- Time Frame: at week 76 of treatment (the week 76 assessment will be analyzed and reported after the extension phase).
- Type: binary variable (Yes / No).

2.2.3. Safety outcomes

A TEAE is defined as any untoward medical occurrence that emerges during a defined

treatment period, having been absent pretreatment, or worsens relative to the pretreatment state, and does not necessarily have to have a causal relationship with this treatment. A SAE is any TEAE from this study that results in one of the following outcomes: death, initial or prolonged inpatient hospitalization, a life-threatening experience (that is, immediate risk of dying), persistent or significant disability/incapacity, congenital anomaly/birth defect, considered significant by the investigator for any other reason.

All TEAEs will be summarized, and the following categories will be further summarized:

- **SAEs**
- **TEAEs leading to study drug discontinuation**
- **TEAEs of special interest**
  - Allergic reaction
  - Injection site reaction
  - Gastrointestinal adverse events (abdominal pain, bloating, constipation, diarrhoea, nausea, vomiting, etc.)
  - Hypoglycemic events
  - Acute coronary syndrome
  - Stroke events
  - Urinary tract infection
  - Fractures

Hypoglycemic events

According to the classification criteria for hypoglycemia of the American Diabetes Association (ADA)/European Association of Diabetes (EASD) 2017 version, hypoglycemia is defined as follows:

- Warning level (Grade 1) of hypoglycemia with a plasma glucose concentration no higher than 70 mg/dL (3.9 mmol/L)

- Symptomatic hypoglycemia: plasma glucose concentration not greater than 70 mg/dL (3.9 mmol/L) but greater than or equal to 54 mg/dL (3.0 mmol/L) with symptoms associated with hypoglycemia.
- Asymptomatic hypoglycemia: a plasma glucose concentration no greater than 70 mg/dL (3.9 mmol/L) but greater than or equal to 54 mg/dL (3.0 mmol/L) without symptoms associated with hypoglycemia.
    - Unclassified hypoglycemia: plasma glucose concentration no higher than 70 mg/dL (3.9 mmol/L) but greater than or equal to 54 mg/dL (3.0 mmol/L) with no documented information on hypoglycemic symptoms.
    - Clinically significant hypoglycemia (Grade 2) with a plasma glucose concentration no greater than 54 mg/dL (3.0 mmol/L)
    - Symptomatic hypoglycemia: a plasma glucose concentration of less than 54 mg/dL (3.0 mmol/L) accompanied by symptoms associated with hypoglycemia.
    - Asymptomatic hypoglycemia: plasma glucose concentration not greater than 54 mg/dL (3.0 mmol/L) without symptoms associated with hypoglycemia.
    - Unclassified hypoglycemia: plasma glucose concentration no higher than 54 mg/dL (3.0 mmol/L) with no documented information on hypoglycemic symptoms.
    - Severe hypoglycemia (Grade 3).

2.2.4. Post-hoc exploratory outcome

Accelerated cognitive decline at week 48 will be assessed as a post-hoc exploratory outcome using a regression-based index[14,36]. For each cognitive measure, accelerated cognitive decline of MoCA or RBANS was defined as a regression-based index at or below the 16th percentile of the analysis population distribution, approximately corresponding to one standard deviation below the mean. Accelerated cognitive decline in this analysis was defined as accelerated cognitive decline on either MoCA or RBANS

score.

2.2.5. Case ascertainment and case definitions

(1) Type 2 diabetes

Type 2 diabetes will be defined per American Diabetes Association criteria: HbA1c ≥ 6.5%, fasting glucose ≥ 126 mg/dL (≥ 7.0mmol/L), two-hour oral glucose tolerance test ≥ 200 mg/dL (≥11.1mmol/L), random glucose≥ 200 mg/dL (≥11.1mmol/L), or use of anti-diabetes medications[37].

(2) Mild cognitive impairment

MCI diagnosis followed the Peterson criteria: cognitive concern from the patient or an informant; objective evidence of cognitive impairment: education-adjusted MoCA score ≤ 25 and ≥ 18; or ≥1.0 SD below the mean of age-specific and education-specific groups on any cognitive subdomain; preservation of independence in daily living abilities: Barthel Index score ≥ 60; and absence of dementia[38,39].

3. STUDY DESIGN

3.1. Design

The LIGHT-MCI study is an investigator-initiated, multicenter, open-label, parallel-group, randomized, superiority trial designed to evaluate the efficacy of three add-on therapies to metformin monotherapy for MCI reversion in T2DM patients. Using a 1:1:1 allocation ratio, participants will be randomly assigned to receive either liraglutide, empagliflozin, or linagliptin for the 48-week core study period. Those completing the core study may opt for an additional 28-week extension phase after providing informed consent, resulting in a total possible intervention duration of 76 weeks. The trial is registered at ClinicalTrials.gov (Identifier: NCT05313529).

3.2. Trial Sites

The trial will be conducted in five hospitals, including Nanjing Drum Tower Hospital, Affiliated Hospital of Medical School, Nanjing University, Nanjing First Hospital, Wuxi People's Hospital, Changzhou Second People's Hospital, and Nanjing Jiangning Hospital.

3.3. Interventions

After discontinuing other anti-diabetes medications besides metformin under guidance for one week, participants will be given another anti-diabetes medication of liraglutide, empagliflozin or linagliptin with an intervention duration of 48 weeks in the core study. The participants who entered the extension study will continue on the same study medication to which they are randomized in the core study up to 28 weeks.

- **Tested drug 1: Liraglutide**

  Liraglutide (Victoza) is supplied as a disposable pre-filled pen with the dosage of 18mg for 3ml per vial. Participants will initiate treatment with liraglutide at a starting dose of 0.6 mg administered via daily subcutaneous abdominal injection. The dose will be escalated in a stepwise manner over an 8-week titration period, with incremental increases to 1.2 mg and subsequently to the target maintenance dose of 1.8 mg, based on individual tolerance as assessed by the investigator. The protocol incorporates a flexible dose adjustment strategy: (a) individualized adjustment of dose-escalation intervals within the 8-week period, (b) dose reduction to 1.2 mg as the maximum tolerated dose when necessary, and (c) protocol-permitted treatment pauses of variable duration for managing persistent adverse events.

- **Tested drug 2: Empagliflozin**

  Empagliflozin tablets (Jardiance) are available in a specification of 10mg per tablet and are taken orally. The tablets should be stored in a cool, dark place and kept sealed, with a shelf life of 24 months. The medication is manufactured by Boehringer Ingelheim Pharma GmbH & Co.KG (Germany) and packaged

by Boehringer Ingelheim Pharmaceuticals (Shanghai) Co., Ltd. Patients are instructed to take one 10mg tablet daily, orally before breakfast.

- **Tested drug 3: Linagliptin**

  Linagliptin tablets (Tradjenta) are available in a specification of 5mg per tablet and are taken orally. The tablets should be stored in a cool, dark place and kept sealed, with a shelf life of 24 months. This medication is manufactured by Boehringer Ingelheim Pharmaceuticals Inc. Patients are instructed to take one 5mg tablet daily, orally before breakfast.

- **Additional treatment criteria**

  If glucose control does not meet standard during follow-up (fasting blood glucose ≥11.1mmol/L more than twice or HbA_{1c} ≥ 8.5%), additional intervention with gliclazide (starting from 30mg once daily and maximum dose not exceeding 120mg) or basal insulin (insulin glargine injection) will be given. The actual glycemic control target and additional treatment therapy will be determined by the investigators according to the individual conditions of participants.

3.4. Randomization and Blinding

This trial adopts a block randomization method with a block size of 6, with each center competing for enrollment. The randomization list will be centrally generated by the statistician of the LIGHT-MCI trial platform before the start of the trial. This randomization list and a document describing the randomization procedure will be kept confidential by the LIGHT-MCI statistician who are not involved in participant screening or treatment. Participants meeting the eligibility criteria will be assigned a random number by a web-based system (JianDaoYun, Fanruan Software Co., Ltd.) and will be allocated to the "liraglutide," "empagliflozin," or "linagliptin" group by randomization in a 1:1:1 ratio after all baseline assessments are completed. When it is confirmed that the subjects are successfully enrolled, the researcher applies for the randomization allocation process, fills in the date and subject label, and the

randomization process manager matches the random number and returns it to the applicant to complete the single randomization allocation process.

Given the inherent challenges of complete blinding due to differing drug administration routes (oral vs. injectable), this study has implemented a rigorous evaluator blinding approach to uphold research integrity. The study involves four independent teams with strictly segregated workflows, including intervention, evaluation, statistical and monitoring team. The intervention team is dedicated to participant enrollment and clinical management. The evaluation team conducts all participant assessments and collects anonymized data while remaining blinded to group allocation. All assessors underwent standardized blinding protocol training and work in controlled environments free of treatment identifiers. The statistical team, unaware of study cohorts, performs data cleaning and verification according to standard procedures. Finally, the statistical team will lock the database once declared complete and accurate, and will perform statistical analysis based on the pre - specified plan.

If accidental unblinding occurs, it's immediately documented (time, cause, personnel) and reported to the independent data and safety monitoring committee. The unblinded assessor is replaced, and affected data is flagged for sensitivity analysis. Preventive measures include enhanced staff training on blinding protocols, regular environmental audits, and ongoing process monitoring by the study coordinator. All unblinding events and corrective actions will be transparently reported in the final publication. Statistical analyses will also include pre - and post - unblinding comparisons to assess potential outcome impacts.

3.5. Sample Size

This superiority trial is powered for two independent primary comparisons (liraglutide vs. linagliptin and empagliflozin vs. linagliptin), each tested at two-sided $\alpha=0.05$ without multiplicity adjustment. Based on our pilot data, we assumed 48-week MCI

reversion rates of 50% (liraglutide), and 30% (empagliflozin). Previous studies of DPP4 inhibitors have shown neutral effect on cognitive function but the reversion rates from MCI to normal cognition is estimated to be approximately 10% to 15% in clinic-based studies, and therefore its MCI reversion rates is assumed to be 15% in our study.[40,41] These assumptions yielded absolute differences of 35% (liraglutide vs linagliptin) and 15% (empagliflozin vs linagliptin). Using PASS 15.0 (NCSS, Kaysville, UT; www.ncss.com/software/pass) with 80% power and 1:1:1 allocation, the required sample sizes were 25 per group for liraglutide and 118 per group for empagliflozin comparisons. To ensure adequate power for both pre-specified tests, we selected the larger sample size (118 per group), ultimately enrolling 132 per arm (total N=396) after accounting for 10% attrition.

4. ANALYSIS POPULATIONS

4.1. Study Population Data Sets

Three study populations will be considered in the analysis as follows:

- **Intent-to-Treat (ITT) population**

  ITT population will consist of all randomized patients and analyzed according to their allocated treatment group, regardless of adherence, treatment discontinuation, protocol deviations or use of rescue medication.

- **Per protocol (PP) population**

  PP population will include only those patients who completed the study according to the protocol and reached all primary endpoints without major deviations.

  Treatment compliance rate (%) will be calculated as:

  100 × (Number of doses actually received / Total number of doses prescribed to be received). Compliance will be categorized into four levels: <50%, 50%–75%, 75%–100%, and >100%. Patients with a treatment compliance rate of <50% will be defined as intervention deviation and excluded from the PP population.

- **Safety population**

Safety population will include all patients who received at least one dose of study medication and will be used for safety analysis. In case of violation of the randomization scheme, patients will be classified according to the treatment they actually received.

4.2. Analysis Close Date

The analysis close date for the core phase is defined as the date on which the last randomized participant completed the Week 48 visit and data required for the Week 48 analyses have been finalized. Data through Week 76 (extension phase) will be locked separately for further analyses.

4.3. Data Cleaning

All data cleaning will be performed and documented in the trial database (Jiandaoyun software developed by Fanruan company) to ensure that there are no erroneous entries and that all missing data is properly coded. A data lock will be performed prior to analysis (core phase lock for Week 48 analyses, extension phase lock for Week 76 analyses). After database lock, the final analysis datasets will be delivered to the statistics team for analysis.

5. STATISTICAL ANALYSES

In this trial, core phase analysis will use data through week 48 only, and extension phase analysis will use data at baseline, weeks 24, 48 and 76. The primary manuscript will present Week-48 core phase analysis. Week-76 analyses will be reported in a separate paper or supplementary material.

5.1. Demographic and Baseline Characteristics

Descriptive statistics of demographic data and baseline characteristics will be summarized by treatment group in the ITT population. Continuous outcomes will be

summarized by treatment arm using number of subjects, mean, standard deviation (SD), median, interquartile range (IQR), minimum and maximum, as appropriate. Categorical variables will be summarized as counts and percent.

- **Pre-specified baseline variables:** age, sex, education, smoking habits, alcohol intake, diabetes duration, body weight, body mass index, waist and hip circumference, waist/hip ratio, systolic and diastolic blood pressure, fasting and 2-hour postprandial plasma glucose, HbA_{1c}, fasting and 2-hour postprandial insulin, fasting and 2-hour postprandial C-peptide, HOMA2-β, HOMA2-S, HOMA2-IR, triglyceride, total cholesterol, high and low-density lipoprotein-cholesterol, visceral fat area, percent body fat, liver stiffness measurement, liver fat controlled attenuation parameter, olfactory assessment total scores, olfactory threshold, identification and memory scores, baseline scores of MMSE, MoCA, RBANS total and index, Barthel index, and FAQ, aggregate time to test completion for TMT, and Victoria Stroop Color-Word Test, MRI-derived brain structure volume, FA, MD, FW-WM, DTI-ALPS, PVSVF-WM, plasma biomarkers (Aβ42/40 ratio, p-tau181, p-tau 217, GFAP and NfL), APOE4 carrier status (carriers and noncarriers), APOE genotype (E2/E2, E2/E3, E2/E4, E3/E3, E3/E4, E4/E4), use of dementia symptomatic medication at baseline (yes or no), and clinical subgroup (MoCA-Impaired and MoCA-Intact but impaired on domain-specific cognitive tests).
- **Baseline comorbidities:** hypertension, dyslipidemia, cardiovascular diseases, and cerebrovascular disease.
- **Medical history**: Medical history will be recorded using standardized medical terminology. Conditions will be categorized and grouped for analysis to ensure consistency.
- **Prior/Concomitant medication:** Prior medications are defined as medications that started before receiving the first dose of the study treatment and that have been stopped before receiving the first dose of the study treatment. Concomitant medications are defined as medications taken from the first dose date to the last dose date + 28 days, including medications that start before the

first dose of the study drug and are ongoing after the first dose, and medications that started on or after the first dose date to the last dose date + 28 days.

5.2. Handling of missing data

5.2.1. Missing baseline covariates

Missing baseline covariates will be imputed using simple imputation methods in the covariate adjusted analysis, should the missing values for a particular covariate be less than 5%. Continuous variables will be imputed using the mean or median simple within the treatment group, while binary and categorical variables will be imputed using the mode within the group. If the missing values for a covariate are ≥5% then they will be imputed by the multiple imputation[42].

5.2.2. Missing primary and secondary outcomes

Missing data for the primary outcome will be handled using Multiple Imputation by Chained Equations (MICE), under a missing at random (MAR) assumption. The imputation will be performed in the ITT population before the primary outcome analyses. The imputation model will include treatment group and baseline covariates (age, sex, education, APOE4 carrier status, and clinical subgroup at baseline). The missing primary outcome will be imputed using logistic regression. We will generate 10 imputed datasets using a random seed of 1015. The analyses will be performed within each imputed dataset, and the results will be pooled using Rubin's rules to obtain overall estimates and 95% confidence intervals (CIs)[42,43]. Missing data for the secondary outcomes will not be imputed.

5.3. Primary Outcome Analysis

5.3.1. Primary analysis

The primary outcome is a binary outcome: MCI reversion at 48 weeks of treatment and

will be summarized using number (%) of patients achieving reversion by treatment group. Two independent primary comparisons (liraglutide vs linagliptin; empagliflozin vs linagliptin) are prespecified.

The primary analysis will be performed in the ITT population using data after multiple imputation (Section 5.2.2). The primary analysis will use a generalized estimating equation (GEE) model with Poisson distribution and log link function, treating group as a categorical variable (linagliptin as reference) and center as a cluster effect, with exchangeable covariance structure. If it does not converge, an independent covariance structure will be used. The margin treatment effects of liraglutide and empagliflozin as compared with linagliptin (reference) will be estimated as point risk ratios (RRs) with unadjusted 95% CI. Robust standard errors will be used for standard error and CI calculations to allow for under- or over-dispersion.

The treatment effects of the two drugs and P values and will be used to draw the main conclusion. The GEE model with Poisson distribution and identity link function will also be used to estimate risk difference (RD) together with unadjusted 95% CIs.
If the above GEE model dose not converge, the following models will be fitted until convergence is achieved:

- GEE model with binomial distribution and log link function;
- GEE model with binomial distribution and logit link function, from which odds ratio will be converted into RR[44].

5.3.2. PP analysis

The primary analysis (Section 5.3.1) will be repeated in the PP population[45].

5.3.3. Sensitivity analyses

5.3.3.1. Covariate adjusted analysis

Covariate adjusted analysis will be performed in the ITT population after imputation of missing baseline covariates and primary outcome (Section 5.2). The primary analysis

(Section 5.3.1) will be performed with adjustment for the following covariates: age (continuous variable), sex (binary variable: male or female), education (categorical variable), APOE4 carrier status (binary variable: carriers or noncarriers), and clinical subgroup (categorical variable: MoCA-Impaired and MoCA-Intact but impaired on domain-specific cognitive tests) at baseline to estimate adjusted RR and 95% CI.

If the covariates adjusted model does not converge, inverse probability of treatment weighting (IPTW) method will be used[46]. We will first calculate a propensity score with treatment as the dependent variable, and all covariates listed above as independent variables through a logistic regression model, and then perform an IPTW analysis. Adjusted RR and its two-sided 95% CI will be estimated.

5.3.3.2. Analysis based on best-/worst-case scenario imputation

(1) Best-case scenario: all missing primary outcome will be imputed as reversion in the liraglutide and empagliflozin groups, and as non-reversion in the linagliptin group.

(2) Worst-case scenario: all missing primary outcome will be imputed as non-reversion in the liraglutide and empagliflozin groups, and as reversion in the linagliptin group[47].

The primary analysis (Section 5.3.1) will be repeated in the ITT population.

5.3.3.3. Complete-case analysis

The primary analysis (Section 5.3.1) will be repeated in the ITT population without imputation (complete cases only).

5.3.3.4. Tipping point analysis

Sensitivity to departures from the MAR assumption for the missing primary outcome will be assessed using a tipping-point analysis in the ITT population[48]. For each

prespecified primary comparison with linagliptin, varying numbers of reversion and non-reversion will be assigned to the missing primary outcomes in each group, and all possible combinations of missing outcome values across the two groups will be examined. Observed primary outcomes will be kept unchanged[49]. For each combination, the method of primary analysis (Section 5.3.1) will be repeated to estimate the RRs, unadjusted 95% CIs, and two-sided P value.

Tipping points will be defined as combinations of missing outcome values for which superiority is no longer established, as indicated by a two-sided P value ≥ 0.05. Additional tipping points will include the RR point estimate ≤1.00, the unadjusted 95% CI includes 1.00, or the RR point estimate ≤1.20. The RR ≤1.20 threshold will be considered exploratory and used to describe attenuation of the treatment effect.

5.3.4. Subgroup analyses

Subgroup analyses will be performed using the same statistical methods for the primary analysis (Section 5.3.1).

Prespecified subgroups include:

- Age at baseline (dichotomized at the sample median value: < median and ≥ median)
- Sex (male, female)
- Education (≤12, and >12 years)
- APOE4 mutation genotype at baseline (carriers, noncarriers)
- Clinical subgroup (MoCA-Impaired, MoCA-Intact but impaired on domain-specific cognitive tests)

Subgroup analyses may also base on categorized variables of the above continuous variable based on median value to make sample sizes of the categories comparable to achieve greater power. Subgroup analysis will be performed in the ITT population after multiple imputation. P values are for treatment-by-subgroup interaction will be reported.

5.4. Secondary Outcome Analysis

All secondary outcomes will be analyzed for superiority test, differences in outcomes (liraglutide vs linagliptin, empagliflozin vs linagliptin) will be reported. Secondary outcome analyses will be based on the complete cases in ITT and PP populations. We will report point estimates, 95% CIs, and p values without multiplicity adjustment.

5.4.1. Analysis of binary secondary outcomes

Binary secondary outcomes will be summarized as number (%) of patients and analyzed with the same GEE model for the primary analysis (Section 5.3.1). The crude and adjusted RR, corresponding 95% CI and p values will be reported.

The secondary outcome of reversion from MCI to normal cognition at week 76 will be analyzed with week 48 reversion as repeated binary outcome. Risk ratios, 95% CIs, and P values at each visit will be estimated using a mixed-effect model with Poisson distribution, log link and robust standard errors, including treatment group, visit, and the treatment-by-visit interaction as fixed effects, with random intercepts for study center and participants.

5.4.2. Analysis of continuous secondary outcomes

Continuous outcomes will be summarized by treatment arm using number of subjects, mean (SD), median (IQR), minimum and maximum, as appropriate.

Single measurement outcomes will be analyzed using GEE model with Gaussian distribution and identity link function, treating group as a categorical variable (linagliptin as reference), center as a cluster effect, and baseline value of the outcome as covariates. The crude and adjusted mean differences of outcomes with their two-sided 95% CI between the groups (liraglutide vs linagliptin; empagliflozin vs linagliptin) will be derived. Normality assumption for the residuals will be assessed using Q-Q plot. A win ratio method will be used if normality assumption is seriously violated[50].

For repeatedly measured continuous outcomes, linear mixed model will be used with treatment, visit, interaction between treatment and visit, center and subject as random effects, and baseline value of the given outcome as a covariate. Mean differences with 95% CIs and p values between the groups will be derived at each visit.

5.4.3. Analysis of functional neuroimaging outcomes

A multi-stage GLM strategy will be employed for neuroimaging outcomes, where first-level individual analysis quantifies individual-level brain activation and second-level analysis measures group-wise differences. A flexible factorial design will be employed to compare brain activation between each active treatment group (liraglutide and empagliflozin) and the active comparator group (linagliptin). For task-based fMRI odor-induced activation, multiple comparisons correction will be performed using Gaussian Random Field (GRF) theory. A cluster-based thresholding approach will be applied: an initial voxel-level threshold of $p < 0.005$ and a cluster level of $p < 0.05$ will be used to define candidate clusters. For whole-brain network connectivity, a 454*454 functional connectivity matrix will be computed for each participant based on the Schaefer-Yeo (400 parcels) and Melbourne Subcortical (54 nuclei) atlases. Group-level comparison of connection-level differences will be conducted using the Network-Based Statistic (NBS) method within the GLM framework, with significant connections defined by an uncorrected P threshold of 0.005.

This neuroimaging-specific analytical strategy is selected to appropriately address the high dimensionality, spatial-temporal auto-correlation, and multiple comparison burden inherent to functional imaging data, which are not adequately handled by GEE-based approaches used for clinical endpoints. Group-level statistical analyses will be performed using SPM 12 and the CONN Toolbox. The GRF-based cluster correction will be implemented through DPABI v6.1_220101's standard statistical inference procedures. Results will be visualized as statistical parametric maps overlaid on standard brain templates for activation differences, and as graphs or matrices highlighting significant network components for connectivity differences.

These prespecified neuroimaging analyses are exploratory for hypothesis-generating; results will be presented as imaging-derived statistical maps or network components.

5.5. Safety analyses

Safety analyses will be performed in the safety population according to the actual treatment group. Core phase safety summaries (baseline-week 48) and overall safety including extension (baseline-week 76) will be presented separately.

The following adverse events will be summarized: any TEAEs, SAEs, TEAEs leading to study drug discontinuation, TEAEs of special interest, etc.

Adverse events will be summarized as counts and percentage by treatment group. Treatment group comparison will be assessed using a Chi-square or Fisher's exact tests and with P value report.

5.6. Exploratory mediation analyses

This analysis aims to test whether change in brain structure or function is a mediator of the treatment effect. Changes in neuroimaging measures (e.g., change in brain structure or function from baseline to Week 48) may be evaluated as potential mediators of the effect of treatment on the primary outcome (MCI reversion at Week 48). These analyses will be performed only if the following preconditions are met: (1) a superior effect of Liraglutide or Empagliflozin arm on the primary outcome versus Linagliptin (control) is observed at week 48, and (2) the same arm shows a superior effect on the candidate mediator (e.g. change in brain structure or function) at week 48. Mediation analyses are pre-specified in the Appendix and details the causal assumptions, directed acyclic graph (DAG), and modeling approach. Results from mediation analyses will be presented as exploratory and will not affect primary confirmatory conclusions.

5.7. Post-hoc exploratory analyses

Accelerated cognitive decline at week 48 will be assessed as a post-hoc exploratory outcome using a regression-based index[14,36]. The analysis will include participants with complete data of MoCA and RBANS scores at baseline and week 48, as well as non-missing age, sex, education, treatment group, and study center. Regression-based indices will be calculated separately for MoCA and RBANS scores. For each measure, the week 48 score will be predicted using a linear regression model that include baseline score, age, sex, and education as predictors. The regression-based index will be calculated as the observed week 48 score minus the predicted week 48 score, divided by the standard deviation of the residuals. For each cognitive measure, accelerated cognitive decline will be defined as a regression-based index at or below the 16th percentile of the analysis population distribution, approximately corresponding to one standard deviation below the mean. Accelerated cognitive decline in this analysis will be defined as accelerated cognitive decline on either MoCA or RBANS score.

The primary analysis (Section 5.3.1) will be repeated to estimate RRs and 95% CIs. No imputation will be performed for this analysis, and no adjustment will be applied.

6. GENERAL CONSIDERATIONS FOR DATA ANALYSES

SAS (version 9.4) will be used to perform all data analyses and generate the majority of data displays. Stata or R may also be used for some data analyses and generating statistical graphs.

6.1. Reporting Guidelines

We will follow the Consolidated Standards of Reporting Trials (CONSORT) 2025 statement: updated guidelines for reporting cluster randomized trials (http://www.consort-statement.org/).

6.2. Participant Disposition and Flow Chart

A CONSORT flow diagram will be prepared to summarize participant disposition throughout the study. The flow chart will illustrate the number of participants screened, enrolled, randomized, treated, followed up, and analyzed in each treatment arm, as well as those included in the ITT and PP populations.

The following information will be presented:

- The total number of participants screened, the number who failed screening, and the reasons for screen failure.
- The number and percentage of participants randomized and receiving at least one dose of the study medication by treatment group.
- The number and percentage of participants who discontinued study medication prematurely, together with reasons for discontinuation.
- The number and percentage of participants who withdrew prematurely from the study, with reasons for withdrawal (e.g., safety, adverse events, loss to follow-up, or other reasons).
- The number and percentage of participants who completed the Week 48 (core phase) and Week 76 (extension phase) visits.
- The number and percentage of participants included in each analysis population (ITT, PP, and Safety), summarized by treatment group.

6.3. Graphical Displays

Mean values for some continuous outcomes by treatment and visit will be plotted.

6.4. Multiplicity

Multiplicity adjustment will not apply to the primary or secondary outcome analyses. Two independent comparisons for the primary outcome are prespecified (Liraglutide vs Linagliptin, Empagliflozin vs Linagliptin), each will be tested at two-sided $\alpha = 0.05$.

7. TABLE CONTENTS

Table 1. Summary statistics of baseline information.

Characteristic	Liraglutide (N=)	Empagliflozin (N=)	Linagliptin (N=)
Age - year			
Male sex - no.(%)			
Education- no.(%)			
≤12 years			
>12 years			
Body Mass Index - kg/m^2			
Cerebrovascular disease- no.(%)			
Cardiovascular disease- no.(%)			
Hypertension- no.(%)			
APOE carrier- no.(%)			
E2/E2			
E2/E3			
E2/E4			
E3/E3			
E3/E4			
E4/E4			
Diabetes			
Diabetes duration - years			
HbA_{1c}- %			
Antidiabetic medications at screening- no.(%)			
Insulin secretagogues			
Glycosidase inhibitors			
Basal insulin			
Clinical outcomes			
MMSE score			
MoCA score			
RBANS total score			

Characteristic	Liraglutide (N=)	Empagliflozin (N=)	Linagliptin (N=)
RBANS index score			
Immediate memory			
Visuospatial/constructional			
Language			
Attention			
Delayed memory			
Trail Making Test - seconds			
Victoria Stroop Color-Word Test - seconds			
Barthel index score			
FAQ score			
MoCA-impaired clinical subgroup - no. (%)			
Brain structure volume – cm^3			
Total brain volume			
Gray matter volume			
White matter volume			
Cerebrospinal fluid volume			
White matter hyperintensity volume			

Table 2. Primary and secondary outcomes.

Variable	Liraglutide (N=)	Empagliflozin (N=)	Linagliptin (N=)
Primary outcome			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Risk difference vs. Linagliptin (95% CI)			-
Key secondary outcomes			
Change in score of MMSE (95% CI)			
Mean difference vs. Linagliptin (95% CI)			-
Change in total score of RBANS (95% CI)			
Mean difference vs. Linagliptin (95% CI)			-
Change in RBANS index score (95% CI)			
Mean difference vs. Linagliptin (95% CI)			-
Change in aggregate time to test completion for TMT (95% CI)			
Mean difference vs. Linagliptin (95% CI)			-
Change in aggregate time to test completion for Victoria Stroop Color-Word Test (95% CI)			
Mean difference vs. Linagliptin (95% CI)			-
Change in MRI-derived total brain volume (95% CI)			
Mean difference vs. Linagliptin (95% CI)			-
Change in MRI-derived gray matter volume (95% CI)			

Mean difference vs. Linagliptin (95% CI)			
Change in MRI-derived white matter volume (95% CI)			
Mean difference vs. Linagliptin (95% CI)			
Change in MRI-derived cerebrospinal fluid volume (95% CI)			
Mean difference vs. Linagliptin (95% CI)			
Change in MRI-derived white matter hyperintensity volume			
Mean difference vs. Linagliptin (95% CI)			

Change is defined as the change from baseline to the 48 weeks of treatment (i.e., value at Week 48 minus baseline value).

Table 3. Sensitivity Analysis of the Primary Outcome.

Analysis	Liraglutide (N=)	Empagliflozin (N=)	Linagliptin (N=)
PP population (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Risk difference vs. Linagliptin (95% CI)			-
Covariate adjusted analysis in ITT population (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Risk difference vs. Linagliptin (95% CI)			-
Best-case scenarios imputation in ITT population (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Risk difference vs. Linagliptin (95% CI)			-
Worst-case scenarios imputation in ITT population (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Risk difference vs. Linagliptin (95% CI)			-
Complete-case analysis in ITT population (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Risk difference vs. Linagliptin (95% CI)			-

Table 4. Subgroup Analysis of the Primary Outcome.

Subgroups	Liraglutide (N=)	Empagliflozin (N=)	Linagliptin (N=)
Age			
< median(n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
≥ median (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Sex			
Male (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Female (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Education			
≤ 12 years (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
>12 years (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
APOE4 carrier status at baseline			

Carriers (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Non-carriers (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			-
Clinical subgroup			
MoCA-Impaired (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			
MoCA-Intact but impaired on domain-specific cognitive tests (n=)			
MCI reversion of at 48 weeks, no. (%)			
Risk ratio vs. Linagliptin (95% CI)			

Table 5. Secondary repeated measured outcomes estimated by linear mixed model. (Core phase analysis)

Outcome	Liraglutide (N=)	Empagliflozin (N=)	Linagliptin (N=)	Liraglutide vs.	Empagliflozin vs. Linagliptin
				Mean difference (95%	Mean difference (95% CI)
Changes in score of MoCA from					
Week 24					
Week 48					
Changes in HbA_{1c} level from baseline					
Week 24					
Week 48					
Changes in BMI level from baseline					
Week 24					
Week 48					

Table 6. Adverse Events (safety population).

Events	Liraglutide (N=)	Empagliflozin (N=)	Linagliptin (N=)
Any treatment-emergent adverse events			
Treatment-emergent serious adverse events			
Treatment-emergent adverse events leading to study drug discontinuation			
Treatment-emergent adverse events of special interest			
Allergic reaction			
Injection site reaction			
Gastrointestinal adverse events			
Hypoglycemic events			
Acute Coronary Syndrome			
Stroke events			
Urinary tract infection			
Fractures			

8. REFERENCE

1. Biessels GJ, Strachan MWJ, Visseren FLJ, et al. Dementia and cognitive decline in type 2 diabetes and prediabetic stages: towards targeted interventions. Lancet Diabetes Endocrinol 2014;2:246–55.
2. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020;396:413–46.
3. Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: Mild cognitive impairment [RETIRED]: Report of the Guideline Development, Dissemination, and Implementation Subcommittee of the American Academy of Neurology. Neurology (ECronicon) 2018;90:126–35.
4. Kantarci K, Weigand SD, Przybelski SA, et al. Risk of dementia in MCI: combined effect of cerebrovascular disease, volumetric MRI, and 1H MRS. Neurology (ECronicon) 2009;72:1519–25.
5. American Diabetes Association Professional Practice Committee. 2. Classification and Diagnosis of Diabetes: Standards of Medical Car in Diabetes—2022. Diabetes Care 2022;45:S17–38.
6. Heerspink HJL, Stefánsson BV, Correa-Rotter R, et al. Dapagliflozin in Patients with Chronic Kidney Disease. N Engl J Med 2020;383:1436–46.
7. McMurray JJV, Solomon SD, Inzucchi SE, et al. Dapagliflozin in Patients with Heart Failure and Reduced Ejection Fraction. N Engl J Med 2019;381:1995–2008.
8. Kosiborod MN, Abildstrøm SZ, Borlaug BA, et al. Semaglutide in Patients with Heart Failure with Preserved Ejection Fraction and Obesity. N Engl J Med 2023;389:1069–84.
9. Rossing P, Baeres FMM, Bakris G, et al. The rationale, design and baseline data of FLOW, a kidney outcomes trial with once-weekly semaglutide in people with type 2 diabetes and chronic kidney disease. Nephrol Dial Transplant 2023;38:2041–51.
10. Cao B, Rosenblat JD, Brietzke E, et al. Comparative efficacy and acceptability of antidiabetic agents for Alzheimer's disease and mild cognitive impairment: A systematic review and network meta-analysis. Diabetes Obes Metab 2018;20:2467–71.
11. Tian S, Jiang J, Wang J, et al. Comparison on cognitive outcomes of antidiabetic agents for type 2 diabetes: A systematic review and network meta-analysis. Diabetes Metab Res Rev 2023;39:e3673.
12. Wu C-Y, Shapiro L, Ouk M, et al. Glucose-lowering drugs, cognition, and dementia: The clinical evidence. Neurosci Biobehav Rev 2022;137:104654.

13. Biessels GJ, Verhagen C, Janssen J, et al. Effect of Linagliptin on Cognitive Performance in Patients With Type 2 Diabetes and Cardiorenal Comorbidities: The CARMELINA Randomized Trial. Diabetes Care 2019;42:1930–8.
14. Biessels GJ, Verhagen C, Janssen J, et al. Effects of linagliptin vs glimepiride on cognitive performance in type 2 diabetes: results of the randomised double-blind, active-controlled CAROLINA-COGNITION study. Diabetologia 2021;64:1235–45.
15. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: a practical method for grading the cognitive state of patients for the physician. J Psychiatr Res. 1975;12:189–98.
16. Nasreddine Z. S., Phillips N. A., Bedirian V., et al. The montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J. Am. Geriatr. Soc 2025;53:695–699.
17. Randolph C, Tierney M, Mohr E, et al. The repeatable battery for the assessment of neuropsychological status (RBANS): preliminary clinical validity. J Clin Exp Neuropsyc.1998; 20:310–9.
18. Randolph C, Tierney M, Mohr E, et al. The repeatable battery for the assessment of neuropsychological status (RBANS): preliminary clinical validity. J Clin Exp Neuropsyc.1998;20:310–9.
19. Sánchez-Cubillo I, Periáñez JA, Adrover-Roig D, et al. Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. J Int Neuropsychol Soc. 2009;15:438–450.
20. Houx PJ, Jolles J, Vreeling FW. Stroop interference: aging effects assessed with the Stroop Color-Word Test. Exp Aging Res. 1993; 19:209–24.
21. *Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. Md State Med J. 1965; 14:61–5.*
22. Gold DA. An examination of instrumental activities of daily living assessment in older adults and mild cognitive impairment. Journal of Clinical and Experimental Neuropsychology. 2012; 34:11–34.
23. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 2006;31:968–980.
24. Schmidt P, Gaser C, Arsic M, et al. An automated tool for detection of FLAIR-hyperintense white-matter lesions in Multiple Sclerosis. Neuroimage 2012; 59: 3774–83.
25. Pasternak O, Sochen N, Gur Y, et al. *Free water elimination and mapping from diffusion MRI. Magn Reson Med. 2009;62:717-730.*
26. Taoka T., Masutani Y., Kawai H., et al. Evaluation of glymphatic system activity with the diffusion MR technique: diffusion tensor image analysis along the perivascular space (DTI-ALPS) in Alzheimer's disease cases. *Jpn. J. Radiol.*2017; 35 172–178.

27. Taoka T, Masutani Y, Kawai H, et al. Evaluation of glymphatic system activity with the diffusion MR technique: diffusion tensor image analysis along the perivascular space (DTI-ALPS) in Alzheimer's disease cases. Jpn J Radiol. 2017;35(4):172-178.
28. Ballerini L, Lovreglio R, Valdes Hernandez MDC, et al. Perivascular spaces segmentation in brain MRI using optimal 3D filtering. Scientific Rep. 2018;8:2132.
29. Sepehrband F, Barisano G, Sheikh-Bahaei N, et al. Image processing approaches to enhance perivascular space visibility and quantification using MRI. Sci Rep. 2019;9(1):12351.
30. Schaefer, A. et al. Local–global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb. Cortex. 2018;28:3095–3114.
31. Rubinov, M. , Sporns, O., et al. Complex network measures of brain connectivity: Uses and interpretations. NeuroImage. 2009;52(3):1059–1069.
32. Van Essen, D. C., Glasser, M. F., Dierker, D. L. et al, Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. Cereb. Cortex. 2012;22:2241–2262.
33. Yeo, B. T. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J. Neurophysiol. 106, 1125–1165 (2011).
34. Zhang Z, Zhang B, Wang X, et al. Altered odor-induced brain activity as an early manifestation of cognitive decline in patients with type 2 diabetes. Diabetes. 2018;67(5):994-1006.
35. Zhang Z, Zhang B, Wang X, et al. Olfactory dysfunction mediates adiposity in cognitive impairment of type 2 diabetes: insights from clinical and functional neuroimaging studies. Diabetes Care. 2019;42(7): 1274-1283.
36. Temkin NR, Heaton RK, Grant I, Dikmen SS. Detecting significant change in neuropsychological test performance: a comparison of four models. J Int Neuropsychol Soc.1999; 5: 357–69.
37. American Diabetes Association Professional Practice C. 2. Diagnosis and. classification of diabetes: standards of Care in Diabetes-2024. Diabetes Care. 2024;47(Suppl 1): S20-S42.
38. Petersen RC. Mild cognitive impairment as a diagnostic entity. J Intern Med. 2004;256(3):183–94.
39. Petersen RC, Lopez O, Armstrong MJ, et al. Practice guideline update summary: mild cognitive impairment: report of the guideline development, dissemination, and implementation subcommittee of the American academy of neurology. Neurology. 2018;90(3):126–35.
40. Malek-Ahmadi M. Reversion from Mild Cognitive Impairment to Normal Cognition: A Meta-Analysis. Alzheimer Dis Assoc Disord. 2016;30:324–30.
41. Canevelli M, Grande G, Lacorte E, et al. Spontaneous Reversion of Mild Cognitive Impairment to Normal Cognition: A Systematic Review of Literature and Meta-Analysis. J Am Med Dir Assoc. 2016;17:943–8.

42. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ.2009; 338: b2393.
43. Little R, Rubin D. Multiple imputation for nonresponse in surveys. Wiley. 1987; 10: 9780470316696.
44. Localio AR, Margolis DJ, Berlin JA. Relative risks and confidence intervals were easily computed indirectly from multivariable logistic regression. J Clin Epidemiol. 2007;60(9):874-82.
45. Cheng DM, Hogan JW. The Sense and Sensibility of Sensitivity Analyses. N Engl J Med. 2024;391(11):972-974.
46. Morris TP, Walker AS, Williamson EJ, White IR. Planning a method for covariate adjustment in individually randomised trials: a practical guide. Trials. 2022;23(1):328.
47. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.
48. Liu Y, Zhou K, Sims KD. Tipping Point Analysis: Assessing the Potential Impact of. Missing Data. JAMA.2025; 334: 265–6.
49. Liublinska V, Rubin DB. Sensitivity analysis for a partially missing binary outcome in a two-arm randomized clinical trial. Stat Med. 2014;33(24):4170-85.
50. Wang D, Pocock S. A win ratio approach to comparing continuous non-normal outcomes in clinical trials. Pharm Stat. 2016;15(3):238-45.
51. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods. 2013;18:137-150.
52. Vanderweele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. Epidemiology. 2014;25:300-6.
53. Lee H, Cashin AG, Lamb SE, et al. A guideline for reporting mediation analyses of randomized trials and observational studies: the AGReMA statement. JAMA. 2021;326:1045–56.
54. Imai K, Keele L, Tingley D. A general approach to causal mediation analysis. Psychol Methods. 2010;15:309–34.

9. APPENDIXES

9.1. Appendix 1. The cutoff scores for various cognitive subdomains

Scores/Education	Average (SD)						Cognitive impairment					
	Primary education		Secondary education		Higher education		Primary education		Secondary education		Higher education	
RBANS												
Immediate memory	70 (15)		76 (15)		85 (15)		55		61		70	
Visuospatial construction	99 (13)		105 (13)		109 (13)		87		91		96	
Language	96 (9)		98 (9)		100 (9)		85		89		91	
Attention	96 (12)		104 (12)		110 (12)		84		92		98	
Delayed memory	88 (14)		92 (14)		96 (14)		74		78		82	
Age (years)	40-60	60-80	40-60	60-80	40-60	60-80	40-60	60-80	40-60	60-80	40-60	60-80
Processing speed	148(66)	202(80)	128(60)	160 (66)	99(40)	148(64)	214	282	188	226	140	212
Executive function	86(27)	93(31)	72(21)	80(22)	66(19)	73(20)	113	93	85	124	102	93

Immediate memory, visuospatial construction ability, language, attention, delayed memory, processing speed and executive function are assessed by Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). Processing speed and executive function are evaluated with the following time tests: the Trail Making Test (parts A and B) and the Stroop Colour-Word Test (parts I, II and III). A longer time for completion (in seconds) represents worse processing speed and executive function. RBANS scores are standardized for age. SD, standard deviation.

9.2. Appendix 2. LIGHT-MCI Study variable list

NO	Variable	
Baseline characteristics		
1	Age, y	N
2	Sex	B
3	Education, y	N
4	Diabetes duration, y	N
5	Weight, kg	N
6	Body mass index, kg/m^2	N
7	Smoking habits	B
8	Alcohol intake	B
9	Waist circumference, cm	N
10	Hip circumference, cm	N
11	Waist/hip ratio	N
12	Systolic blood pressure, mmHg	N
13	Diastolic blood pressure, mmHg	N
14	Fasting plasma glucose, mmol/L	N
15	2-hour postprandial plasma glucose, mmol/L	N
16	HbA_{1c}, %	N
17	HbA_{1c}, mmol/L	N
18	Fasting insulin, mmol/L	N
19	2-hour postprandial insulin, mmol/L	N
20	Fasting C-peptide, mmol/L	N
21	2-hour postprandial C-peptide, mmol/L	N
22	HOMA2-β	N
23	HOMA2-S	N
24	HOMA2-IR	N
25	Triglyceride, mmol/L	N
26	Total cholesterol, mmol/L	N
27	High-density lipoprotein-cholesterol , mmol/L	N
28	Low-density lipoprotein-cholesterol, mmol/L	N
29	Visceral fat area, cm^2	N
30	Percent body fat, %	N
31	Liver stiffness measurement, KPa	N
32	Liver fat controlled attenuation parameter, dB/m	N

33	Interventions	C (0= linagliptin，1=liraglutide, 2=empagliflozin)
34	Olfactory assessment total scores	N
35	Olfactory threshold scores	N
36	Olfactory identification scores	N
37	Olfactory memory scores	N
38	MMSE score	N
39	MoCA score	N
40	RBANS total score	N
41	Immediate memory score	N
42	Visuospatial construction score	N
43	Language score	N
44	Attention score	N
45	Delayed memory score	N
46	Trial Making Test, seconds	N
47	Victoria Stroop Color-Word Test, seconds	N
48	FAQ score	N
49	Barthel index	N
50	Plasma Aβ42/40 ratio	N
51	Plasma p-tau 181, pg/ml	N
52	Plasma p-tau 217, pg/ml	N
53	Plasma GFAP, pg/ml	N
54	Plasma NfL, pg/ml	N
55	APOE 4 carrier status (carriers and noncarriers)	B
56	APOE genotype (E2/E2, E2/E3, E2/E4, E3/E3, E3/E4, E4/E4)	T
55	Dementia symptomatic mediation use (yes or no)	B
57	Clinical subgroup (MoCA-Impaired and MoCA-Intact but impaired on domain-specific cognitive tests)	B
58	Intracranial volume, cm^3	N
59	Total brain volume, cm^3	N
60	Gray matter volume, cm^3	N
61	White matter volume, cm^3	N
62	Cerebrospinal fluid volume, cm^3	N
63	White matter hyperintensity volume, cm^3	N
64	Frontal lobe gray matter volume, cm^3	N
65	Temporal lobe gray matter volume, cm^3	N
66	Parietal lobe gray matter volume, cm^3	N
67	Occipital lobe gray matter volume, cm^3	N
68	Insular lobe gray matter volume, cm^3	N
	Subcortical gray matter nuclei volumes, cm^3	
69	Thalamus	N
70	Putamen	N

71	Pallidum	N
72	Amygdala	N
73	Caudate	N
74	Accumbens	N
	AD signature region volume, cm^3	
75	Hippocampus	N
76	Parahippocampal	N
77	Entorhinal	N
78	Inferior parietal lobule	N
79	Precuneus	N
80	Cuneus	N
81	Mean FA	N
82	MD	N
83	FW-WM	N
84	DTI-ALPS	N
85	PVSVF-WM	N
86	Hypertension	B
87	Dyslipidemia	B
88	Cardiovascular diseases	B
89	Cerebrovascular disease	B
	Antidiabetic medications at screening	
90	Insulin secretagogues	B
91	Glycosidase inhibitors	B
92	Basal insulin	B
Week 24 variable list		
1	Weight, kg	N
2	Body mass index, kg/m^2	N
3	Waist circumference, cm	N
4	Hip circumference, cm	N
5	Waist/hip ratio	N
6	Systolic blood pressure, mmHg	N
7	Diastolic blood pressure, mmHg	N
8	Fasting plasma glucose, mmol/L	N
9	HbA_{1c}, %	N
10	HbA_{1c}, mmol/L	N
11	Triglyceride, mmol/L	N
12	Total cholesterol, mmol/L	N
13	High-density lipoprotein-cholesterol, mmol/L	N
14	Low-density lipoprotein-cholesterol, mmol/L	N
15	Visceral fat area, cm^2	N
16	Percent body fat, %	N
17	Liver stiffness measurement, KPa	N

18	Liver fat controlled attenuation parameter, dB/m	N
19	MoCA score	N
Week 48 variable list		
1	Weight, kg	N
2	Body mass index, kg/m^2	N
3	Waist circumference, cm	N
4	Hip circumference, cm	N
5	Waist/hip ratio	N
6	Systolic blood pressure, mmHg	N
7	Diastolic blood pressure, mmHg	N
8	Fasting plasma glucose, mmol/L	N
9	2-hour postprandial plasma glucose, mmol/L	N
10	HbA_{1c}, %	N
11	HbA_{1c}, mmol/L	N
12	Fasting insulin, mmol/L	N
13	2-hour postprandial insulin, mmol/L	N
14	Fasting C-peptide, mmol/L	N
15	2-hour postprandial C-peptide, mmol/L	N
16	HOMA2-β	N
17	HOMA2-S	N
18	HOMA2-IR	N
19	Triglyceride, mmol/L	N
20	Total cholesterol, mmol/L	N
21	High-density lipoprotein-cholesterol, mmol/L	N
22	Low-density lipoprotein-cholesterol, mmol/L	N
23	Visceral fat area, cm^2	N
24	Percent body fat, %	N
25	Liver stiffness measurement, KPa	N
26	Liver fat controlled attenuation parameter, dB/m	N
27	Olfactory assessment total scores	N
28	Olfactory threshold scores	N
29	Olfactory identification scores	N
30	Olfactory memory scores	N
31	MMSE score	N
32	MoCA score	N
33	RBANS total score	N
34	Immediate memory score	N
35	Visuospatial construction score	N
36	Language score	N
37	Attention score	N
38	Delayed memory score	N
39	Trial Making Test, seconds	N

40	Victoria Stroop Color-Word Test, seconds	N
41	FAQ score	N
42	Barthel index	N
43	Plasma Aβ42/40 ratio	N
44	Plasma p-tau 181, pg/ml	N
45	Plasma p-tau 217, pg/ml	N
46	Plasma GFAP, pg/ml	N
47	Plasma NfL, pg/ml	N
48	Dementia symptomatic mediation use (yes or no)	B
49	MCI reversion (yes or no)	B
50	Intracranial volume, cm^3	N
51	Total brain volume, cm^3	N
52	Gray matter volume, cm^3	N
53	White matter volume, cm^3	N
54	Cerebrospinal fluid volume, cm^3	N
55	White matter hyperintensity volume, cm^3	N
56	Frontal lobe gray matter volume, cm^3	N
57	Temporal lobe gray matter volume, cm^3	N
58	Parietal lobe gray matter volume, cm^3	N
59	Occipital lobe gray matter volume, cm^3	N
60	Insular lobe gray matter volume, cm^3	N
	Subcortical gray matter nuclei volumes, cm^3	
61	Thalamus	N
62	Putamen	N
63	Pallidum	N
64	Amygdala	N
65	Caudate	N
66	Accumbens	N
	AD signature region volume, cm^3	
67	Hippocampus	N
68	Parahippocampal	N
69	Entorhinal	N
70	Inferior parietal lobule	N
71	Precuneus	N
72	Cuneus	N
73	Mean FA	N
74	MD	N
75	FW-WM	N
76	DTI-ALPS	N
77	PVSVF-WM	N
78	Additional antidiabetic treatment therapy	B

Note: T, text variable; D, The date type; N, Continuous variable; B, binary variable.

9.3. Appendix 3. Exploratory mediation analysis

- **Exploratory objectives**

This section details a pre-specified, exploratory, hypothesis-generating analysis. Its purpose is to investigate the potential mechanistic role of brain structural or functional improvement, only if a specific hierarchy of outcomes is observed: namely, if a superior effect on the primary outcome (MCI reversion) is found for an intervention arm and that same arm also shows a superior effect on the proposed mediator (brain structural or functional improvement). Should the preconditions for analysis be met, the following analytic approach will be employed.

- **Conceptual framework and causal model**

A critical assumption for the validity of our mediation analysis is that of temporal or logical ordering—that alterations in brain structure or function causally precede and lead to MCI reversion. Although both the mediator and the outcome are measured at the same time point (week 48), they represent the cumulative change over the entire trial period. We explicitly assume that the treatment first alters brain structure or function, which in turn facilitates clinical reversion. This assumption is grounded in established neurobiological models of cognitive function[51] We acknowledge that the simultaneous measurement does not allow us to empirically test the direction of causality. Therefore, the results of the mediation analysis must be interpreted as providing evidence consistent with the hypothesized mechanistic pathway, rather than as definitive proof of causation. The extensive sensitivity analyses described below are primarily designed to assess the robustness of our findings to violations of this key assumption.

Figure 1 depicts the directed acyclic graph (DAG) of the hypothesized effect pathways, where A denoted the exposure (treatment assignment; A_i=1 for liraglutide or empagliflozin; A_i=0 for linagliptin) and Y will be the achievement of MCI reversion from baseline to week 48. M represented the mediator of interest (M= change in the

brain structure or function from baseline to week 48).

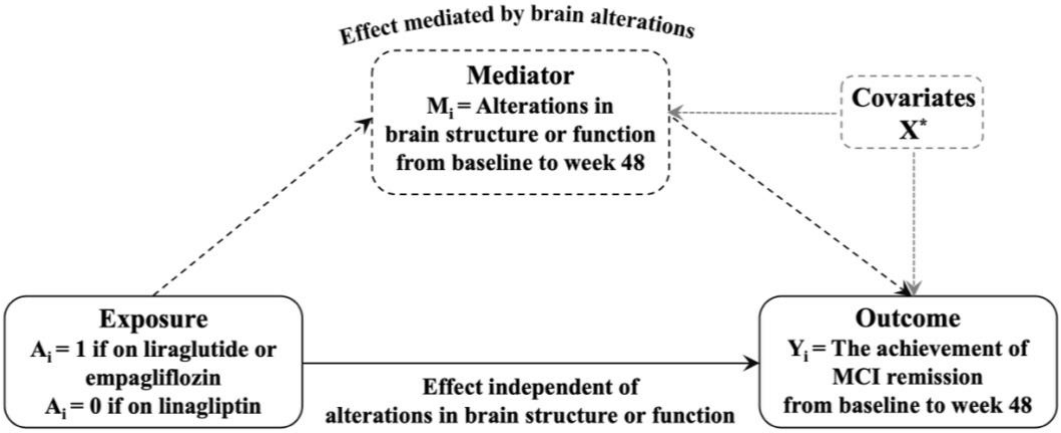


Figure 1. Directed acyclic graph for the exploratory mediation analysis

The effect of the treatment (liraglutide or empagliflozin) on the clinical outcome will be decomposed into 2 parts: a natural direct effect that will be independent of alterations in brain structure or function, and an indirect effect that will be mediated by alterations in brain structure or function. By comparing the relative magnitudes of the direct and indirect effects, the roles that brain structural or functional alterations played in achieving MCI reversion will be determined. On the basis of the assumed pathways shown in this figure, a GEE model with Gaussian distribution and identity link function, will be used to depict the effect of the treatment on the mediator while adjusting for the influence of relevant covariates. Outcome models will be constructed using GEE model with Poisson distribution and log link function, treating group as a categorical variable (linagliptin as reference) and center as a cluster effect, with exchangeable covariance structure. Within the mediation analysis framework, it is assumed that no unmeasured confounding variables are affecting the relationship between the mediator and the outcome. Thus, a separate sensitivity analysis assessed the impact of unobserved mediator–outcome confounding variables.
MCI = Mild Cognitive Impairment. *Covariates: baseline age, sex, brain structural or functional measure, education, APOE4 carrier status, and clinical subgroup (MoCA-Impaired and MoCA-Intact but impaired on domain-specific cognitive tests).

The causal mediation analysis (hereafter called "mediation analysis") will be conducted within the potential outcome framework (see key assumptions in Table 2). A GEE model with Gaussian distribution and identity link function, center as a cluster effect, will be used to describe the effect of the treatment assignment (liraglutide or empagliflozin vs. linagliptin) on the mediator (alterations in brain structure or function from baseline to week 48) while adjusting for the influence of relevant baseline covariates. To meet the requirements of the mediation analysis, the following baseline

variables will be included in all models: age (continuous), sex (binary), education (continuous), *APOE4* carrier status (binary), clinical subgroup (categorical: MoCA-Impaired vs. MoCA-Intact but impaired on domain-specific tests), and baseline brain structural or functional measure value (continuous).

The outcome models will be constructed using GEE models with Poisson distribution and log link function, treating group as a categorical variable (linagliptin as reference) and center as a cluster effect, with exchangeable covariance structure. The effect of treatment independent of alterations in brain structure or function (that is, the natural direct effect) and the effect of treatment mediated by alterations in brain structure or function (that is, the natural indirect effect), as defined by Valeri and VanderWeele[52], will be estimated. The corresponding standard errors (SEs) will be estimated via bootstrapping with 5,000 resamples, and the 95% CIs will be ascertained using bias-corrected and accelerated methods. The estimated proportion of the total effect explained by the mediator will be then calculated, along with the corresponding 95% CIs.

Among the standard assumptions for mediation analysis (Table 2), the no unmeasured mediator–outcome confounder assumption is not directly testable. Therefore, we conducted a sensitivity analysis based on the method of VanderWeele[52] to quantify the robustness of the main findings to a potential mediator–outcome confounder. This analysis estimates the degree of correlation required for an unmeasured confounder to nullify the indirect effect.

Furthermore, to confirm the validity of the selected statistical model, a model which included the exposure-by-mediator (A*M) interaction term as a covariate will be tested. Other sensitivity analyses will be done to assess the robustness of the models and conclusions: 1) inclusion of baseline BMI and HbA_{1c} (non-prespecified variables) as covariates in the mediation model; 2) conducting the analysis using a linear probability model for the outcome to assess the impact of the link function on the estimated

proportion mediated; and 3) performing a reverse analysis testing the implausible pathway to assess the concern of reverse causation.

All mediation analyses will be performed using the CAUSALMED procedure in SAS. We will report the results following A Guideline for Reporting Mediation Analyses (AGReMA) guidelines[53]. General assumptions for mediation analyses will be described by Imai et al.[54].

Assumptions for Mediation Analysis*

Assumptions	Definition	Explanation	Justification
A.1†	Y (a, m) ⊥ A\ X	No unmeasured confounding in the treatment–outcome relationship	Satisfied because of random treatment assignment.
A.2†	M (a) ⊥ A \ X	No unmeasured confounding in the treatment–mediator relationship	Satisfied because of random treatment assignment.
A.3	Y (a, m) ⊥ M(a) \ A = a, X	No unmeasured confounding in the mediator–outcome relationship for a given treatment (i.e., in the single-world model	This assumption is not testable. We conducted a simulation analysis to assess the robustness of the findings when there is an unmeasured confounding effect that influences both M and Y.
A.4 ‡	Y (a, m) ⊥ M(a^*) \X	No mediator–outcome confounder is affected by exposure. The assumption extends A3 to the independence of Y (a, m) and M (a^*) across the worlds of the potential outcomes (i.e., when the potential outcome Y (a, m) and mediator M (a^*) are assessed under different treatments). It essentially requires there is nothing on the pathway from exposure to the mediator that also affects the	The assumption is usually satisfied when the mediator occurs after the exposure, as is the case in our research setting.

Temporal (logical) ordering	The causal effect of M on Y is unidirectional	The assumption posits that the alterations in brain structure or function (M) causally precede and influence the achievement of MCI reversion (Y) in a logical sequence, even though both are measured at the same point.	This assumption is justified by strong biological plausibility derived from prior literature, which suggests that structural or functional neural changes are prerequisite for observable cognitive improvement. Although measured concurrently at week 48, the mediator (M) represents the cumulative change over the trial period, which is hypothesized to mechanistically lead to the outcome (Y). The primary concern to this assumption is reverse causation (Y causing M). To assess this concern, we will perform a reverse analysis testing the implausible pathway.

* Under the standard notation for mediation analysis: *A* represents the exposure (*A*=1 for liraglutide or empagliflozin; *A*=0 for linagliptin), *M* the continuous mediator (alterations in the brain structure or function from baseline to week 48), and *Y* the achievement of MCI reversion from baseline to week 48. *M (a)* represents the potential value of the mediator, had the exposure taken value a; *Y (a, m)* represents the potential outcome under exposure a and mediator value m; and *Y (a, M (a))* and *Y (a, M (a*))* are the potential outcomes in the single and cross-world models, respectively. Here, the single-world model refers to the setting where the potential mediator and outcome are assessed under the same treatment; the cross-world model refers to the setting where the potential mediator and outcome are assessed under different treatments.

† A.1 and A.2 are collectively known as the sequential ignorability assumptions.

‡ Also known as the cross-world ignorability assumption.