Abstract (Words): 300
Key points (Words): 96
Manuscript (Words): 3500
Tables and Figures: 5
Supplementary Materials, Tables: 6
Supplementary Materials, Figures: 3
References: 46

Type 2 diabetes mellitus accelerates brain aging and cognitive decline: complementary findings from UK Biobank and meta-analyses.

Botond Antal, M.S.[1,2*], Liam P. McMahon, B.A.[1,2*], Syed Fahad Sultan, M.S.[3], Andrew Lithen, B.S.[1,2], Deborah J. Wexler, M.D.[4], Bradford Dickerson, M.D.[2,5], Eva-Maria Ratai, Ph.D.[2], Lilianne R. Mujica-Parodi, Ph.D.[1,2,6]

* *Shared first authors, contributed equally to the manuscript.*

1. Department of Biomedical Engineering, Stony Brook University, Stony Brook, New York, USA
2. Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, Charlestown, Massachusetts, USA
3. Department of Computer Science, Stony Brook University, Stony Brook, New York, USA
4. Diabetes Center, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA
5. Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, USA
6. Department of Neurology, Stony Brook University School of Medicine, Stony Brook, New York, USA

For correspondence, please contact:

Lilianne R. Mujica-Parodi, Ph.D.
Director, Laboratory for Computational Neurodiagnostics
Professor, Department of Biomedical Engineering
Stony Brook University School of Medicine
Stony Brook, NY 11794-5281
Email: Lilianne.Strey@stonybrook.edu

Declaration of Interests and Sources of Funding

The research described in this paper was funded by the W. M. Keck Foundation (to LRMP) and the White House Brain Research Through Advancing Innovative Technologies (BRAIN) Initiative (NSFNCS-FR 1926781 to LRMP). None of the funding sources played any role in the design of the experiments, data collection, analysis, interpretation of the results, the decision to publish, or any aspect relevant to the study. DJW reports serving on data monitoring committees for Novo Nordisk. None of the authors received funding or in-kind support from pharmaceutical and/or other companies to write this manuscript.

Abstract

Background: Type 2 diabetes mellitus is known to be associated with neurobiological and cognitive deficits; however, their extent, overlap with aging effects, and the effectiveness of existing treatments in the context of the brain are currently unknown.

Methods: We characterized neurocognitive effects independently associated with T2DM and age in a large cohort of human subjects from the UK Biobank with cross-sectional neuroimaging and cognitive data. We then proceeded to evaluate the extent of overlap between the effects related to T2DM and age by applying correlation measures to the independently characterized neurocognitive changes. Our findings were complemented by meta-analyses of published reports with cognitive or neuroimaging measures for T2DM and healthy controls (HC). We also evaluated in a cohort of T2DM diagnosed individuals using UK Biobank how disease chronicity and metformin treatment interact with respect to the identified neurocognitive effects.

Findings: The UK Biobank dataset included cognitive and neuroimaging data (N=26,125) including 1,270 T2DM and 24,855 HC. Duration of T2DM ranged from 0–45 years (mean 9.7±7.9 years); 559 were treated with metformin alone, while 473 were unmedicated. Our meta-analysis evaluated 34 cognitive studies (N=22,231) and 60 neuroimaging studies: 30 of T2DM (N=866) and 30 of aging (N=1088). As compared to age, sex, and education-matched HC, T2DM was associated with marked cognitive deficits, particularly in *executive functioning* and *processing speed*. Likewise, we found that the diagnosis of T2DM was significantly associated with gray matter atrophy, primarily within the *ventral striatum*, *cerebellum*, and *putamen*, with reorganization of brain activity (decreased in the *caudate*, *frontal eye fields*, and *premotor cortex* and increased in the *subgenual area*, *thalamus*, *brainstem* and *posterior cingulate cortex*). The structural and functional changes associated with T2DM show marked overlap with the effects correlating with age but appear earlier, with disease duration linked to more severe neurodegeneration. Metformin treatment status was not associated with improved neurocognitive outcomes.

Interpretation: The neurocognitive impact of T2DM suggests marked acceleration of normal brain aging, by approximately 24% ± 10%; T2DM chronicity was associated with increased atrophy. As such, neuroimaging-based biomarkers may provide a valuable adjunctive measure of T2DM progression and treatment efficacy based on neurological effects.

Introduction

In 2018, ~6.6% of the population carried a diagnosis of Type 2 diabetes mellitus (T2DM) [1]. Patients with T2DM are at greater risk for developing dementia and Alzheimer's disease (AD) and have been reported to exhibit inferior cognitive performance when compared to age-matched healthy controls (HC) [2]. Several human neuroimaging studies have linked T2DM with brain atrophy and cognition [2-6]; recent research suggested that T2DM resulted in a more rapid rate of cognitive decline than typically associated with natural aging [7-9].

Despite strong preliminary evidence linking T2DM to neurological and cognitive decline, few patients with T2DM undergo a comprehensive neurocognitive evaluation as part of their clinical care [8,10,11]. This may reflect the fact that T2DM diagnosis often occurs in middle age, hindering dissociation of patients' cognitive changes from normal aging. Several studies published to date focused on the neurocognitive effects of T2DM include age-matched participants. However, because none has compared lifespan neurological changes to those experienced by equivalently aged patients with T2DM, it is currently unknown whether neurocognitive effects represent a T2DM-specific neurodegenerative pathway or the exacerbation of typical brain aging. Moreover, there remain limited data[12] evaluating the impact of chronicity or role of effective treatment in the progression of cognitive and neurological decline.

Routine clinical protocols typically focus on peripheral biomarkers (e.g., blood glucose and insulin levels, body fat percentage) as diagnostic modalities for T2DM. However, the neurological effects of T2DM may be apparent for many years before they can be detected by peripheral markers [3,8]. As such, by the time T2DM is diagnosed and treated by standard measures, patients may have already sustained irreversible brain damage. Thus, there are direct clinical implications with respect to defining the neurocognitive impact of T2DM and to determine how these negative sequelae might be prevented or treated [1].

Given these unknowns and their clinical importance, here we focus on addressing three questions. First, we establish T2DM neurocognitive effects, as compared to age, sex, and education-matched healthy controls (HC). To do so, we leverage the robust statistical power made possible by UK Biobank [13], the largest (N=26,125) neurocognitive lifespan dataset to date, with UK Biobank results then compared to a meta-analysis of the published literature (34 cognitive studies, 60 neuroimaging studies) to assess convergence. Second, we ask whether changes in the brain observed in T2DM represent normal aging, accelerated aging, or a non-aging-related degenerative pathway specific to T2DM. Third, we test whether T2DM chronicity exacerbates, and medication status ameliorates, the progression of neurocognitive effects.

Methods

Analysis of UK Biobank Dataset (N=26,113)

General Overview: UK Biobank data were analyzed for both cognitive and neuroimaging data (**SI Table 1**). The primary factor of interest was T2DM, which we dissociated from age-related effects by age matching T2DM and HC. To permit comparison of T2DM-specific effects to age-specific effects, we also assessed the same neurocognitive variables with age as a factor of interest from samples that excluded patients diagnosed with T2DM. To control for potential neurocognitive confounds, T2DM and HC were matched for not only age, but also sex and education (i.e., exact pairwise matching between groups for age, sex, and education). To ensure that all individuals in our T2DM sample had type 2, rather than type 1 diabetes, we only included individuals with a self-reported age of onset >20 years.

We fitted linear regression models to neurocognitive variables and quantified associated effects as the maximum likelihood point estimates and confidence intervals (95% CIs) of the corresponding regression coefficients. T2DM was accounted for as a binary factor with two states corresponding to healthy controls and individuals diagnosed with T2DM, whereas age was considered as a continuous factor with increments in years. Regression models were fit using the Statsmodels Python library[14]. To account for multiple comparisons, we applied Bonferroni correction to cognitive and structural results, and adjusted for false discovery rate in our neurofunctional results; *p* values were adjusted accordingly in all results.

To determine whether T2DM neurocognitive effects suggested non-aging-related degenerative pathways specific to T2DM, versus an exacerbation or acceleration of typical aging trajectories, we compared the progression of neurodegeneration seen in T2DM to that seen in relation to age across brain regions and cognitive domains using bivariate Pearson correlations.

Cognition: Data on five cognitive domains for 24,300 participants (T2DM: N=1,152, HC: N=23,148) were extracted from the UK Biobank dataset, including *abstract reasoning*, *executive function*, *processing speed*, *reaction time*, and *short-term memory*. Exact sample sizes varied across cognitive domains based on data availability, and therefore are noted separately for each result. We employed linear regression and considered the maximum likelihood estimates of coefficients belonging to age and T2DM to estimate their associations with performance in each of the five domains, as well as with a combined cognitive performance metric. The combined cognitive performance metric was derived from z-transformed task performance scores averaged across the individual domains for each subject. Effect sizes in cognition were quantified as percentages by dividing the estimated beta coefficient and 95% CIs of the factor of interest with the average performance of HC.

Brain Structure: Using structural MRI data from the UK Biobank dataset, we assessed the effects associated with T2DM (T2DM: N=982, HC: N=982) as compared to non-T2DM-specific age-related effects (N=14,836) on atrophy of gray matter volume; these findings were available as voxel counts for the whole brain as well as for 139 anatomical regions. For region-specific analyses, we coarse-grained the default unilateral parcellation provided by UK Biobank into 45 bilateral regions and corrected gray matter volumes for head size. We applied linear regression and quantified atrophy in each anatomical region as a relative percentage change in average gray matter volume by dividing the estimated regression coefficients and 95% CIs corresponding to the factor of interest with the average gray matter volume of HC.

Brain Function: Using functional MRI data from the UK Biobank dataset, we assessed the effects occurring with T2DM (T2DM: N=712; HC: N=712) as compared to non-T2DM-specific age-related effects (N=3,660) on resting-state brain activity. Data were accessed already preprocessed by UK Biobank according to their standard pipelines[15]. After transforming functional images to Montreal Neurological Institute (MNI) space, we performed spatial smoothing with a full width at half maximum (FWHM) of 5 mm, then quantified brain activity using the amplitude of low-frequency fluctuation[16] (ALFF). We used the program *3dRSFC*, which is a component of Analysis of Functional NeuroImages[17,18] (AFNI), to compute ALFF in voxel space. ALFF was computed from the 0.01–0.08 Hz frequency band, within a gray matter only brain mask. Computed voxel space ALFF values were normalized to the global mean of each individual subject. Statistical analyses were performed in voxel space using the Nistats Python library. We used a significance threshold of $p < 0.05$ and a minimum cluster size of 12 voxels (~100 mm^3) and controlled for multiple comparisons using false discovery rate (FDR).

Implications of T2DM Duration: To evaluate the implications of T2DM chronicity, we analyzed whole brain gray matter volume with time since T2DM diagnosis as a regressor. Time since diagnosis was derived from self-reported age at T2DM diagnosis. To improve the accuracy of these self-reported values, we considered their average from three visits, separated by multiple years. To estimate the degree to which T2DM progression was associated with gray matter loss relative to age, we calculated the ratio of regression

coefficients corresponding to T2DM duration and age, and expressed it as a percentage using Fieller's theorem[19] to quantify the confidence interval (95%) for this ratio.

Implications of Metformin Treatment: For patients with T2DM, we evaluated whether metformin, a first-line medication for the treatment of T2DM, was associated with improved outcomes in terms of cognition, atrophy, and/or brain activity. To isolate medication effects specific to metformin, we compared subjects who reported not taking any medications to treat T2DM, to subjects who reported taking metformin but no other medications. For these comparisons, we exact matched for sex and education, coarse matched for age (bin size of five years) and disease duration (bin size of three years). Since UK Biobank did not measure HbA1c levels, we also statistically controlled for BMI as the only available proxy measure for disease severity[20,21].

Meta-Analysis of Published Literature (N=24,185)

Search strategy and selection criteria (cognition): We conducted a literature search for peer-reviewed articles published up to August 28, 2020 from PubMed/Medline using the following search terms: "type-2-diabetes," "diabetes mellitus, type 2," "insulin-resistance," <AND> "cognition," "cognitive-function," "cognitive-dysfunction," "cognitive-performance," and "neuropsychological tests." Search results were filtered to include manuscripts that had undergone peer-review, were published in English with full-text availability, and reported relevant results. Our cognitive meta-analysis adhered to PRISMA guidelines [22].

We included studies that compared cognitive performance between people diagnosed with T2DM and healthy controls. We excluded studies that (a) included participants with neurological or psychiatric diagnoses, (b) utilized treatment interventions without first obtaining baseline cognitive measurements, (c) included only diagnostic threshold instruments for dementia (e.g., the Mini-Mental State Examination, or MMSE), (d) included a novel cognitive test without adequate explanation of the scoring procedures, (e) did not perform age and education-matching of the participants diagnosed with T2DM to their HC, or (f) failed to provide summary statistics needed to calculate effect sizes. In the latter case, the authors were contacted to obtain relevant data.

Our literature search yielded 219 articles; relevant reviews were also screened for eligible studies. Seventy-five articles were identified for full-text evaluation; 34 studies were eligible for inclusion. Among the studies that were excluded, eight featured inadequate testing or scoring procedures, 14 included secondary analyses of the same patient sample that was used in previous publications, and five failed to perform appropriate education-matching of the study groups. Furthermore, one longitudinal study did not report baseline scores and another reported inconsistent sample sizes. Fifteen authors were contacted to obtain data not provided in the text; three authors provided the data requested, and the remaining 12 studies were excluded. Eligible studies included a total of 4,735 subjects diagnosed with T2DM and 17,496 HC.

Data analysis (cognition): We extracted data including publication year, authors, sample demographics, and cognition from all included studies. We extracted baseline data only from longitudinal studies to avoid practice effects. We sorted individual cognitive tests into several domains, including *abstract reasoning*, *verbal memory*, *visual memory*, *working memory*, *information processing speed*, *executive function*, *short-term memory*, *verbal fluency*, *visuospatial construction*, and *motor speed* (**SI Table 2**).

Statistical analyses were performed using R version 3.6.1 [23] and the Metafor package version 2.4-0 [24]. Cognitive differences between participants diagnosed with T2DM and HC were determined by calculating standardized mean difference (SMD) effect sizes and 95% CIs for all cognitive domains. Effect size analyses were chosen to account for within-domain variability in the type and sensitivity of cognitive tests across different reports. We calculated effect sizes as Cohen's d by dividing the mean difference in group scores by the pooled standard deviation of individual domains [25]; an SMD (Cohen's d) of –1.0 was

interpreted as a difference of one standard deviation in the negative direction. We used random-effects models to account for variability between samples not due to sampling error with significance at $p < 0.05$ and effect-size heterogeneity was evaluated using values for Cochran's Q and I^2 [26]. Publication bias was evaluated with funnel plots. We applied Bonferroni correction to account for multiple comparisons across cognitive domains.

Search strategy and selection criteria (brain): We used NeuroQuery[27] to conduct a meta-analysis of published neurobiological results associated with T2DM and age. NeuroQuery is an automated Coordinate-Based Meta-Analysis (CBMA)[28-30] tool based on a database of z-scores collected by crawling through texts and tables of published research articles by an automated algorithm[31]. We utilized NeuroQuery to address limitations of standard approaches to meta-analyses of neuroimaging results, which rely on summary statistics and thus risk overfitting to what typically comprise a relatively small number of in-sample studies (i.e., they fail to generalize to out-of-sample studies). NeuroQuery optimizes for rigor and reproducibility by utilizing predictive modeling, a higher threshold for results than statistical significance. In a quantitative evaluation of its generalization performance with 16-fold cross validation and 10:90 test-train splits, NeuroQuery was found to accurately produce brain maps for out-of-sample neuroimaging studies[31]. We note that because of NeuroQuery's criteria for neuroimaging data quality and completeness in reporting, the algorithm draws only from journals that focus on functional neuroimaging results, and thus can exclude some general interest and non-neurological medical journals. This exclusion criterion is important to reduce false positives and ensure the quality and relevance of the compiled results. The database in total contains 149,000 neuroimaging papers and represents the single largest database of neuroimaging foci to date. By the law of large numbers NeuroQuery therefore provides the most unbiased approach to choosing representative papers, even at the risk of excluding relevant and well-cited articles specific to any one field.

Using the collected database of articles, NeuroQuery applies a multivariate model to predict the spatial distribution of voxel activations corresponding to a search term. The search terms we used to obtain the meta-analytic maps were: "diabetic" and "age". These terms identified the 30 most relevant neuroimaging studies for T2DM and 30 most relevant neuroimaging studies for age (**SI Table 3**). To account for any errors in the automated search results, the identified set of studies were cross-validated by an independent manual search using the same search terms for Google Scholar and PubMed to verify their relevance, as well as to confirm that they included T2DM age and sex-matched HC and T2DM (not Type 1 diabetes). In the T2DM datasets, 23 were fMRI (ALFF), two were structural (T1), three were FDG positron emission tomography (PET), and two were tractography (diffusion tensor imaging, DTI). In the age datasets, 22 were fMRI (ALFF), three were structural (T1), and five were tractography (diffusion tensor imaging, diffusion weighted imaging).

Data analysis (brain): For region and voxel level comparisons of the meta-analytic T2DM and Age maps from Neuroquery with their structural and functional counterparts from UK Biobank, the meta-analytic statistical maps were transformed onto comparable coordinate space and spatial resolution. At the voxel level, the meta-analytic maps were resampled to the standard MNI affine (the transformation matrix that maps from voxel indices of the data array to actual real-world locations of the brain; no registration was required as images were already aligned). For region level comparisons, the transformed voxel maps were coarse-grained to the 45 regions of interest from UK Biobank by masking with each individual region and computing the mean activation of the masked voxels as the representative region value.

Results

Cognitive correlates with age and T2DM

Individuals without T2DM showed age-based cognitive effects across all domains in the UK Biobank (**Fig. 1A)**. The strongest effects were observed in *executive function*, which showed 2.0% ± 0.1% decrease in performance per year (N=7,296, T= –30.6, p<1e–10) and *processing speed*, which showed 1.6% ± 0.1% decrease in performance per year (N=7,508, T= –41.0, p<1e–10;). Our analyses identified further cognitive deficits associated with T2DM, beyond typical age-related effects (**Fig. 1B)**. The strongest effects were observed in *executive function*, which showed an 8.8% ± 5.6% decrease in performance (T2DM: N=594; HC: N=594; T= –3.0, p=0.01), and *processing speed*, which showed a 6.3% ± 2.9% decrease in performance. (T2DM: N=612; HC: N=612; T= –4.3, p=0.0001). All other domains were also significantly affected: –3.0% ± 1.5% in *reaction time* (T2DM: N=1,149; HC: N=1,149; T= –4.0, p=0.0003), –4.3% ± 2.0% in *short-term memory* (T2DM: N=647; HC: N=647; T= –4.2, p=0.0001) and –4.0% ± 2.5% in *abstract reasoning* (T2DM: N=1,119; HC: N=1,119; T= –3.1, p=0.01). Our meta-analysis confirmed that individuals with T2DM exhibited markedly lower performance when compared to age, sex, and education-matched controls, over an even broader set of domains (**Fig. 1C**). These included *executive function* (K=18, d= –0.40, p=0.009), *short-term (~30 seconds) verbal memory* (K=23, d= –0.39, p=0.001), *verbal fluency* (K=25, d= –0.37, p=2e–8), *working memory* (K=12, d= –0.36, p=0.002), *abstract reasoning* (K=8, d= –0.36, p=1e–7), *information processing speed* (K=31, d= –0.34, p=5e–8), *visuospatial reasoning* (K=13, d= –0.32, p=4e–7), *delayed* (~20 minute) *verbal memory* (K=21, d= –0.21, p=0.005), and *short-term (~2-3 seconds) memory ("attention")* (k=16, d= –0.21, p=0.05) (**SI Table 4)**.

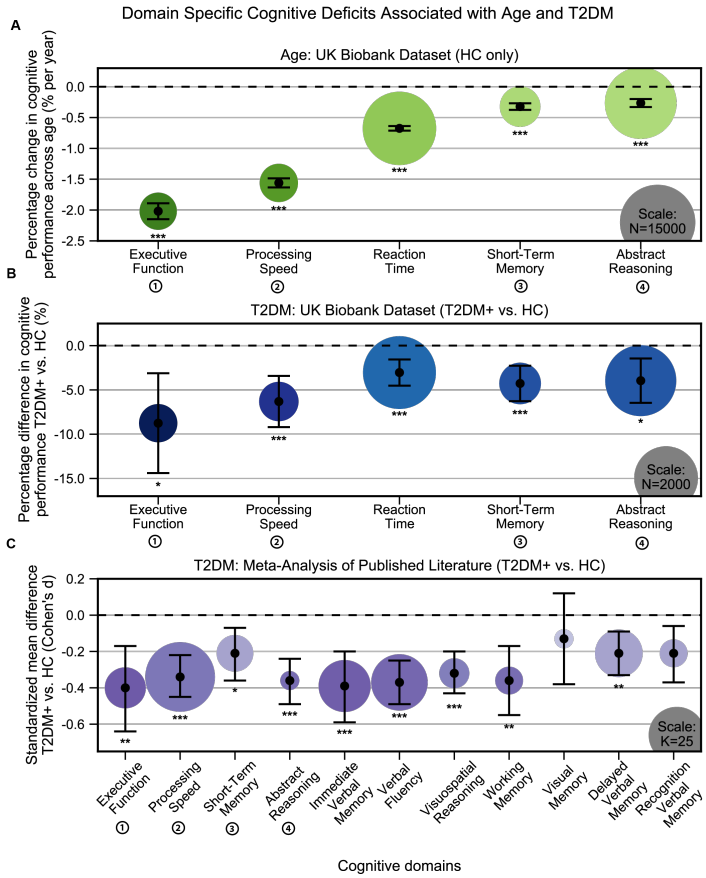


Figure 1: Cognitive deficits are apparent with respect to both age and T2DM diagnosis. A: Using the UK Biobank dataset, we performed a quantitative analysis of the effects related to age on cognitive performance across five cognitive domains. Associated changes were derived from estimated regression coefficients as percentages and are shown on the y axis. Age was associated with significant deficits in all five domains, with the strongest effects observed in *executive function* and *processing speed*. **B:** Using the same dataset, we also analyzed cognitive performance in T2DM, with negative values on the y-axis represent performance below that of age, sex, and education-matched HC. As per age effects, *executive function* and *processing speed* showed the highest magnitude changes. **C:** Cognitive deficits identified in UK Biobank data were confirmed by our meta-analysis, which included 11 domains from 34 studies. Average effect sizes (Cohen's d) corresponding to T2DM are shown on the y axis. Values below the cut-off line (y=0) indicate cases in which subjects with T2DM performed less well than age, sex, and education-matched HC. Numbers next to labels identify domains common across panels. Marker sizes represent sample sizes scaled as indicated in the bottom left corner of each panel. On Panel C, sample size indicates the number of individual studies. Error bars are 95% CI. *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$, Bonferroni corrected.

Neurobiological correlates with age and T2DM

Brain Atrophy: HC (N =14,836) showed a linear decrease in brain gray matter with age. This was most pronounced in the *ventral striatum*, which showed a 0.9% ± 0.04% decrease per year (T= –55.4, p<1e–10) and *Heschl's gyrus*, which also showed a 0.9% ± 0.03% decrease per year (T= –55.1, p<1e–10) **(Fig. 2A)**. As compared to their age-matched HC, T2DM patients showed further decreases in gray matter beyond typical age-related effects (T2DM: N=982; HC: N=982). These included both cortical and subcortical regions, with the most severe atrophy observed in the *ventral striatum*, which showed on average a 5.5% ± 1.4% further decrease in volume (T= –7.5, p<1e-10), in the *putamen* with an additional 4.7% ± 2.0% decrease in volume (T= –4.6, p=0.0002), and in the *cerebellum*, which showed a 5.0% ± 1.0% further decrease (T= –9.8, p<1e–10) **(Fig. 2B).**

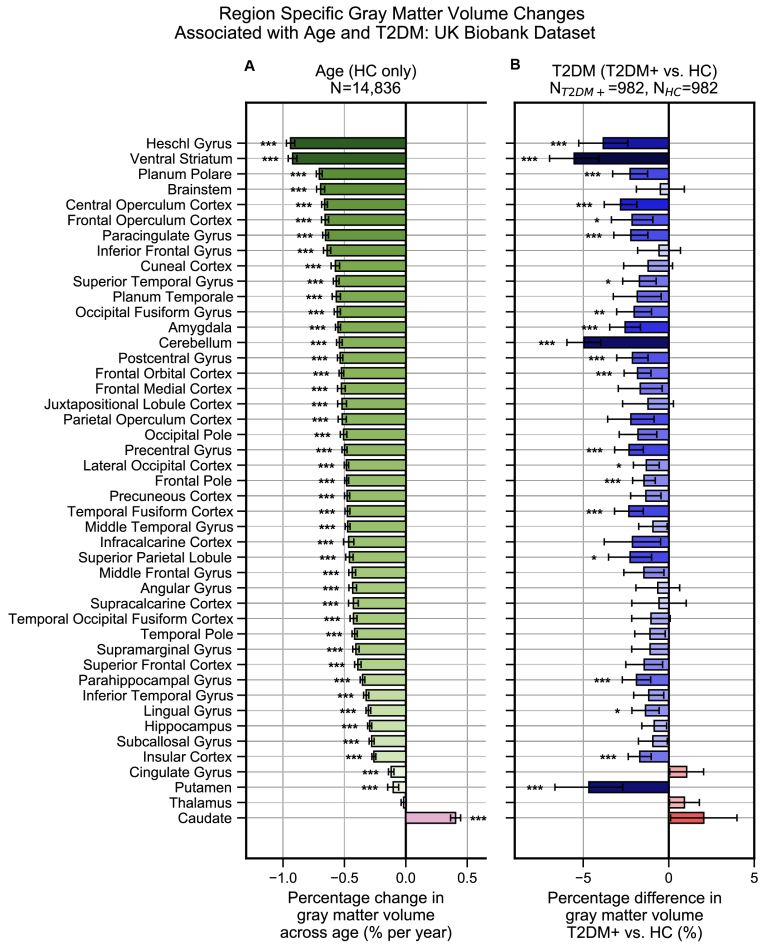


Figure 2: Widespread gray matter atrophy can be observed with respect to both age and T2DM diagnosis status. Using the UK Biobank dataset, we measured gray matter atrophy across 45 anatomical regions. Associated changes were derived from estimated regression coefficients as percentages and are shown on the x axes. **A:** We observed significantly decreased gray matter volume in both cortical and subcortical brain regions with respect to age in HC. Age was associated with an average of ~0.5% brain-wide decrease in gray matter volume per year, most prominently for the *ventral striatum* and *Heschl's gyrus* **B:** Gray matter atrophy was also seen in patients diagnosed with T2DM compared to age matched HC, most prominently for the *ventral striatum*, *cerebellum*, and *putamen*. The distribution of T2DM-related effects overlapped with those associated with age, with degeneration of the *ventral striatum* and preservation of the *thalamus* and *caudate*. Error bars are 95% CI.
*$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$, Bonferroni corrected.

Brain Activity: Age was associated with functional reorganization of brain activity (ALFF), rather than global decrease or increase. Brain activation in T2DM showed similar reorganization. Normalized to whole brain activity, both age (HC: N=3,660) and T2DM (T2DM: N=712, HC: N=712) were associated with decreased activation in the *caudate, premotor cortex* and *frontal eye fields*, with increased brain activity in the *subgenual area, thalamus* and *brainstem* (**Fig. 3A**).

Neuroquery: Our meta-analysis of 60 multimodal neuroimaging studies (30 age-specific, 30 T2DM-specific) independently identified the same regions as UK Biobank (*caudate*, *frontal eye fields*, *premotor*

cortex, thalamus), but additionally identified clusters of decreased activity in *Broca area* and the *superior temporal gyrus* and increased activity in the *posterior cingulate cortex*, and *angular gyrus* (**Fig. 3B**).

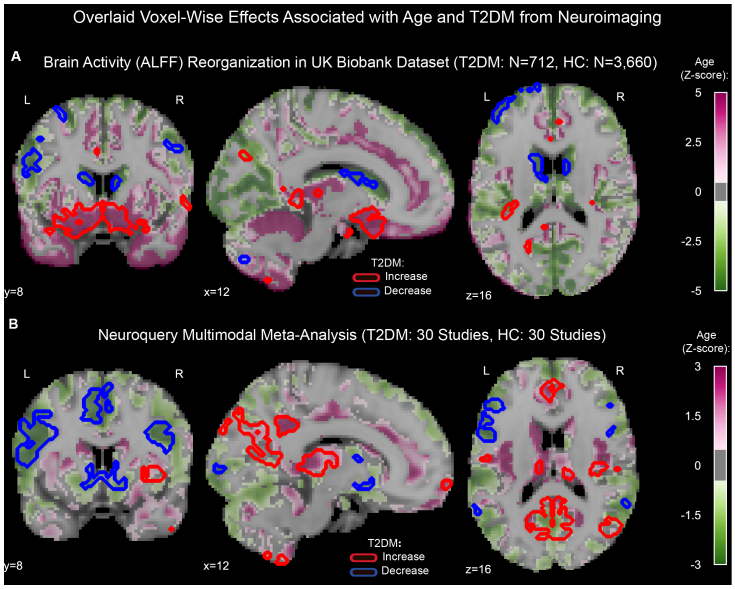


Figure 3: Overlap between age and T2DM with respect to reorganization of brain activity. A: For functional MRI data obtained from the UK Biobank dataset, we used the amplitude of low-frequency fluctuation (ALFF) to quantify brain activity. Effects linked to age are shown in the form of an unthresholded z-map represented by the pink-green color gradient, with pink indicating increased activation and green showing decreased. T2DM related effects were thresholded (minimum cluster size ~100mm^3, FDR p<0.05) to result in significant clusters. The outlines of these significant clusters are overlaid on the age-related z-map to demonstrate overlapping effects. The largest significant clusters with respect to T2DM were in the *subgenual area* (increased), the *caudate* (decreased), and *frontal eye fields* (decreased). All highlighted regions were similarly impacted across age, indicating substantial overlap between the two contrasts. **B:** Using multimodal neuroimaging data, we performed a meta-analysis for the same contrasts using NeuroQuery. We extracted contrast maps for age and T2DM with NeuroQuery and overlaid the outlines of thresholded (minimum cluster size ~100mm^3, FDR p<0.05) z-maps from T2DM on unthresholded z-maps belonging to age. The overlapping effects were evident in several regions, most importantly in the *cingulate gyrus*, *thalamus* and *premotor cortex*. These results support the hypothesis that neurodegeneration in both T2DM and aging may be associated with common mechanistic pathways.

Neurocognitive Changes associated with T2DM and Age Overlap, Consistent with Common Pathways.

Together, these analyses confirm that T2DM patients show evidence of neurocognitive deficits, with the most consistent and profound effects observed in structural atrophy (across all regions: T= –9.0, p<1e–10) (**SI Fig. 1B**). Even after controlling for education, cognitive deficits remained statistically significant (T= –3.8, p=0.0001) (**SI Fig. 1A)**. Both age and T2DM implicated the same areas of greatest vulnerability: for brain atrophy, this was the *ventral striatum*; for cognition, these were *executive function* and *processing speed*. When assessed across all brain regions, T2DM-related patterns in brain atrophy exhibited strong overlap with those associated with age (r=0.61, p=0.0002). Similarly, T2DM-related changes in brain activity (ALFF) also exhibited significant overlap with those associated with age (r=0.44, p=0.04). The meta-analysis, which included multimodal neuroimaging measures (not only atrophy and brain activity, but also glucose uptake via FDG-PET) also yielded equivalent results in terms of the overlap between neurobiological effects of T2DM and age (r=0.58, p=0.0005) (**SI Fig. 2**).

T2DM Chronicity Exacerbates Neurocognitive Symptoms.

Neurocognitive effects were more severe with increased disease duration, particularly for structural changes (T= –4.8, p=0.000003) (**Fig. 4**). T2DM progression was associated with 24% ± 10% acceleration of typical neurogenerative age-related effects, as per the linear shift along the horizontal time axis shown in **Fig 4**.

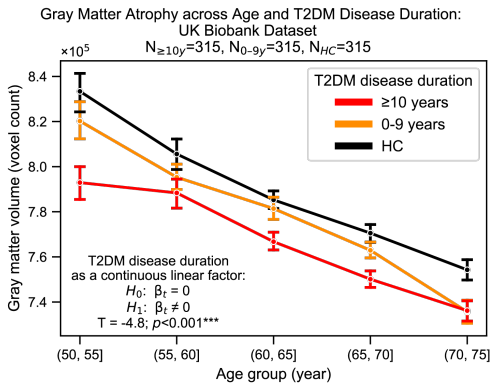


Figure 4: Progression of T2DM disease is significantly associated with gray matter atrophy, accelerating neurodegenerative effects seen in brain aging. For a quantitative evaluation of the impact of T2DM progression on gray matter volume, we considered time since T2DM diagnosis as the main factor of interest from the UK Biobank dataset. The T2DM+ cohort was divided into two groups based on disease duration (separated at 10 years) with a HC cohort also included for visualization purposes. We matched age, sex, education across these three groups and performed linear regression within T2DM+ subjects focusing on disease duration. Evaluation of our sample suggested that time since diagnosis was a significant factor, with each year after diagnosis of T2DM associated with an additional ~0.24 ± 0.10 years of brain aging beyond that of age-matched T2DM–. Error bars are standard error of the mean. $*P \leq 0.05$, $**P \leq 0.01$, $***P \leq 0.001$.

T2DM Patients Treated with Metformin Do Not Demonstrate Improved Neurocognitive Symptoms.

Even after matching groups for disease duration and controlling for BMI, T2DM patients who were treated with metformin alone (N=559) did not differ with respect to cognition or brain atrophy compared to T2DM patients who were unmedicated (N=473) (**SI Fig. 3**). Likewise, treatment status showed no significant impact on resting-state brain activity.

Discussion

The UK Biobank dataset confirms that T2DM patients show deficits in cognitive performance compared to HC, even after controlling for age, sex, and education, findings that were supported by meta-analysis of the published literature. These deficits in cognitive performance were accompanied by marked brain atrophy in the in T2DM sample as compared to age-matched HC. The atrophy was most severe (5% grey matter loss compared to HC) in the *ventral striatum*, a region critical to learning, decision making, goal-directed behavior, and cognitive control. These cognitive functions, collectively known as *executive functioning*, were (with processing speed) also those most affected by T2DM. Neurodegeneration for all regions was worsened with chronicity.

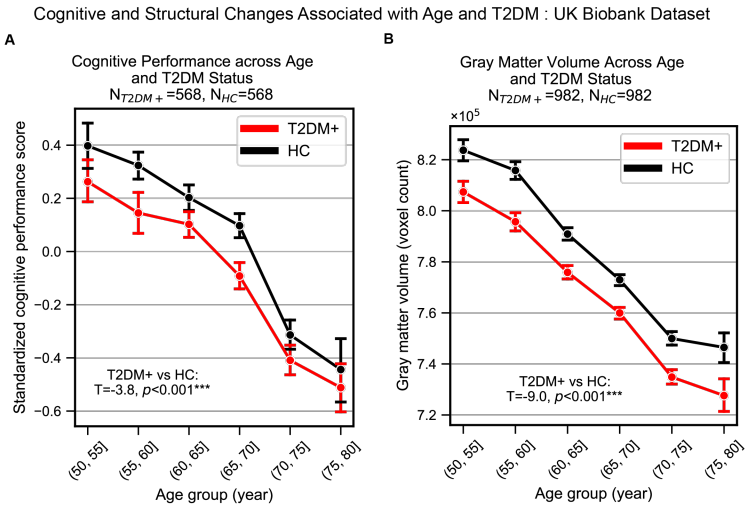
Our findings indicate that structural brain imaging, in particular, can provide a clinically valuable metric for identifying and monitoring neurocognitive effects associated with T2DM. Normalizing across sample sizes to compare the measures of neurocognitive effects: structural MRI, functional MRI, and cognitive testing, structural atrophy showed global effects that were more statistically robust (p<2e–10) than either global cognitive measures (p=0.0001) or global brain activity (p=0.002). One important advantage of structural MRI over cognitive testing is that the former avoids confounding associated with education and practice effects. Moreover, cognitive testing may be less interpretable in real-world clinical settings in which such matching is not feasible. Structural MRI also showed advantages as a biomarker over a functional MRI-derived measure of brain activation (ALFF). The reorganization of brain activity seen with T2DM may reflect the brain's switch to less metabolically expensive networks to conserve energy in the face of diminishing access to glucose, a pattern previously documented in aging[32-35]. Yet activation patterns that are spatially reorganized, rather than globally increased or decreased, are less straightforward to quantify. Moreover, functional MRI is an inherently more complex measure than structural MRI, reflecting both neuronal and hemodynamic influences. Each of these of these influences may be differentially affected by T2DM, further complicating its interpretation in a clinical setting.

The localization of brain atrophy in T2DM to the *ventral striatum*, followed by the *cerebellum*, may reflect the fact that these two brain regions contain the densest concentrations of insulin-dependent GLUT-4[36-39], as compared to non-insulin-dependent isoforms GLUT-1 and GLUT-3. The *ventral striatum* functions as a critical hub within the reward circuit, integrating inputs (including external stimuli) from both cortical and subcortical regions, and therefore is a key structure required for all learning. Rat studies have shown modulation of nitric oxide within the *ventral striatum* to control release of acetylcholine[40], a neurotransmitter severely reduced in dementia[41] and a target for its pharmaceutical treatment[42,43]. Release of nitric oxide is insulin dependent and reduced in T2DM[44]. Together, these suggest a potential mechanistic pathway between insulin resistance, atrophy of the *ventral striatum*, and widespread deficits with respect to learning. In this context, memory deficits may be primarily driven by failure to encode rather than failure to retrieve, which would be consistent with our results which did not identify the hippocampus as be one of the regions most affected. Importantly, the structural and functional changes associated with T2DM show marked overlap with age-related effects but appear earlier. This suggests that neurocognitive changes seen in T2DM may progress via a common mechanistic trajectory as normal brain aging, but which is accelerated.

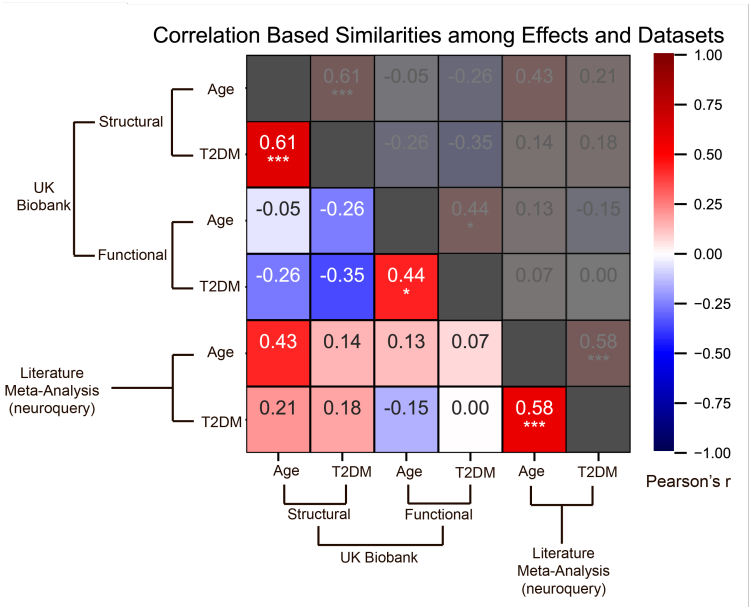
Our analyses had two limitations, inherent in the datasets analyzed, which represent important directions for future research. First, our use of a lifespan dataset permitted tracking how variables change with age, but not for the same subjects. A more rigorous assessment of phase shift between trajectories of neurodegeneration for patients with T2DM and HC would be made possible only with a longitudinal study. Second, while we had access to disease duration and BMI, we did not have HbA1c measures, which would have provided a more direct measure of disease severity. While metformin was not found to be associated with better neurocognitive measures, even when matched to unmedicated patients with equivalent disease duration and after controlling for BMI (a proxy measure for disease severity[20,21]), it was not possible to determine other diabetes-related characteristics. As such, our medication findings should be considered suggestive but not conclusive.

Consistent with findings from earlier studies that focused on the brain and energy metabolism [45,46], we suggest that the T2DM and its progression may be associated with accelerated brain aging. As T2DM results in compromised energy availability, brain structure and function undergo accelerated deterioration. We consider the possibility that, by the time T2DM is formally diagnosed, neuronal insulin resistance may have already resulted in significant brain damage. As such, our findings underscore the need for additional research into brain-based biomarkers for T2DM and treatment strategies that specifically target its neuro-cognitive effects [1].

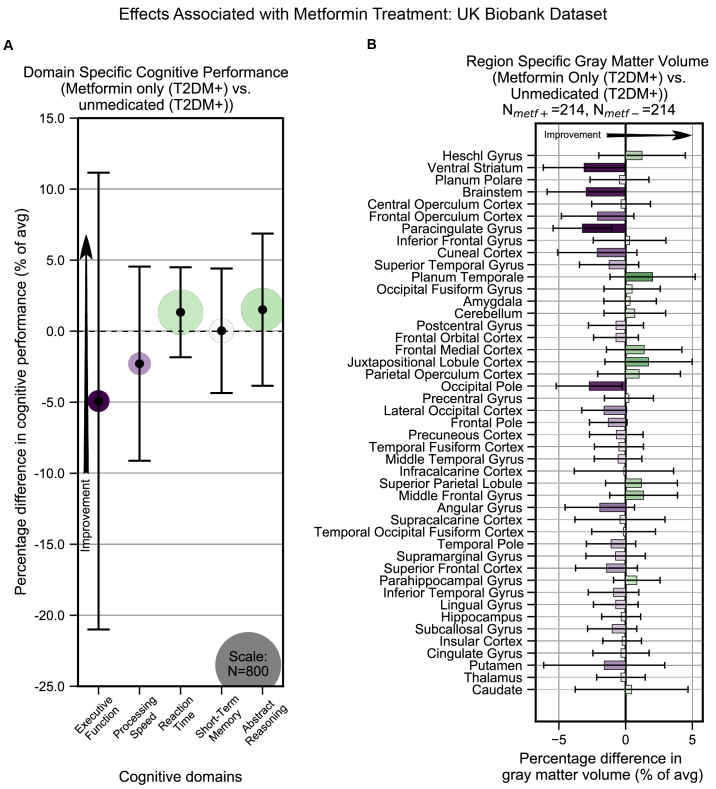
Supplementary Figures



Supplementary Figure 1: Aggregated whole brain measures represent the extent of accelerated brain aging with T2DM diagnosis. We used the UK Biobank dataset to address the extent by which aging is accelerated in individuals with T2DM. Subjects with T2DM were age, sex, and education matched with HC **A:** We quantified a gross cognitive metric from the combination of multiple z-scored performance scores from five cognitive domains. This metric yielded an effective representation of the general decline across age, the gap between HC versus subjects diagnosed T2DM, and the relative extent of these two phenomena. We observed significantly decreased cognitive performance in subjects with T2DM: an increase of 3.8 ± 2.0 years in age-related cognitive decline. **B:** An equivalent analysis was performed using whole brain gray matter volume. This metric yielded even stronger results compared to cognition. T2DM diagnosis was associated with significant atrophy: an increase of ~4.2 ± 0.9 years in age-related neurodegeneration. Error bars are standard error of the mean. The seemingly constant gap between the two cohorts is explained by uniformly distributed disease duration across the lifespan. *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.



Supplementary Figure 2: Effects of age and T2DM exhibited strong correlations across datasets and modalities. We considered six sets of previously characterized changes in association with the following: 1. Age contrast, gray matter volume in UK Biobank; 2. T2DM contrast, gray matter volume in UK Biobank; 3. Age contrast, brain activation in UK Biobank; 4. T2DM contrast, brain activation in UK Biobank; 5. Age contrast, brain structure/activation (aggregate) from NeuroQuery; 6. T2DM contrast, brain structure/activation (aggregate) from NeuroQuery. Corresponding effects from region/domain specific analyses were considered as inputs for correlation measures, which were then determined for all combinations of the six sets of effects. Age and T2DM were significantly correlated (Pearson's *r*) within all modalities, suggesting common trajectories between age and T2DM related effects. No other significant correlations were observed, however, between datasets or modalities. Given that structural and functional effects appeared to be unrelated within UK Biobank, and that the NeuroQuery results were a combination of both structural and functional results, we did not expect significant associations between modality-specific UK Biobank results and multimodal (structural and functional) NeuroQuery results. However, it is important to note that very different ways of acquiring and analyzing brain data independently replicate the correlations between age and T2DM related effects, suggesting that the association is highly robust. *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$, Bonferroni corrected.



Supplementary Figure 3: Treatment of T2DM patients with metformin had no impact on cognitive deficits or gray matter atrophy. We evaluated the UK Biobank dataset to determine whether treatment with metformin would prevent gray matter atrophy or the development of cognitive deficits associated with T2DM. Among T2DM diagnosed subjects only, we compared those subjects who reported using metformin but no other medications to those who reported not taking any medications to treat T2DM. We matched subjects for age, sex, education and T2DM disease duration, and controlled for BMI. The direction of theoretical improvement by metformin is indicated on both panels by an arrow. **A:** No statistically significant (α=0.05) differences in cognitive performance were detected when comparing subjects on metformin to unmedicated subjects **B:** Neither our analysis of gray matter atrophy detected any significant (α=0.05) improvements associated with metformin treatment. Error bars are 95% CI.

Supplementary Tables

Supplementary Table 1. Summary of All Relevant UK Biobank data-fields.

Variable	Designation	Instance Number
Diagnosis (T2DM)	2443	0, 2
Age	21003	0, 2
Sex	31	2
Education	6138	2
Age-of-onset (T2DM)	2976	0-2
Body Mass Index (BMI)	21001	2
Medication Status (Metformin)	20003	2
Gray-Matter Volume	25005-25006, 25782-25920	2
Resting-State MRI Images	20227	2
Matrix-Pattern Completion	20016	2
Alphanumeric Trail-Making Test	6350	2
Symbol-Digit Substitution	23324	2
Snap Game	20023	0, 2
Numeric Memory Test	4282	2

Supplementary Table 2. Summary of Cognitive Functions Assessed, with Corresponding Instruments.

Domains	Common Tests	Description
Verbal Memory	Rey & California Auditory Verbal Learning Tests, Hopkins Verbal Learning Test, Delayed Word Recall, Weschler Text Recall Sub, Word List Recall, Weschler Story Recall	Short and long-term recall of verbal information. Includes both auditory and visual encoding.
Executive Function	Trail Making (B), Stroop (III), Brixton Spatial Anticipation, Wisconsin Card Sort, Color Trails (2), Weschler Letter Number Sequencing	Top-down coordination of other cognitive domains (e.g., memory, motor function) to solve problems and manage cognitive resources. Often exhibited in tasks that require a degree of planning.
Verbal Fluency	Word & Semantic Fluency Tests, Controlled Oral word Association Test, Letter & Category Fluency Tests, Boston Naming Test	Language skills. Commonly measured by enumeration (e.g., name as many words as you can that begin with the letter "B").
Information Processing Speed	Trail Making (A), Digit Symbol Substitution, Stroop (I-II), Choice Reaction Time, Color Trails (1)	Speedy encoding and use of information. Often measured by time-to-completion in tasks that require the manipulation of presented information.
Working Memory	(Backwards) Digit Span, Corsi Block Tapping, N-back	Holding information for a short time for use on a current task. Characterized by both maintaining and manipulating stored information. Commonly measured by having subjects re-order learned information.
Visuospatial Reasoning	Rey-Osterreith Figure Copy, Taylor Complex Figure, Weschler Object Assembly	Manipulation or reconstruction of spatial information.
Abstract Reasoning	Raven's Progressive Matrices, Matrix Pattern Completion, Weschler Similarities, Standard Progressive Matrices	Manipulation of presented information to solve a problem without prior knowledge. Interrelated with fluid intelligence. Often presented as shape or logic puzzles.
Visual Memory	Location Learning, Weschler Visual Memory Subtest, Rey-Osterreith Delayed Recall, Face Recognition Test	Short and long-term recall of visually encoded information.

Harvey, (2019). Domains of Cognition and their Assessment. Dialogues of Clinical Neuroscience, 21(3), 227-237. doi:10.31887/DCNS.2019.21.3

Supplementary Table 3. Studies Identified as Most Relevant for Each Key Word by NeuroQuery Algorithm.

	T2DM studies	Age studies
1.	Chun-Xia Wang et al. 2014	György A Homola et al. 2012
2.	Xiangzhe Qiu et al. 2016	Peiying Liu et al. 2013
3.	Natalia García-Casares et al. 2016	G Juckel et al. 2012
4.	Z-L Wang et al. 2017	Natalie C Ebner et al. 2013
5.	Thomas J Marder et al. 2014	Michelle Hampson et al. 2012
6.	Franco Cauda et al. 2009	Sien Hu et al. 2012
7.	Ying Cui et al. 2015	Yu-Chien Wu et al. 2011
8.	Dae-Jin Kim et al. 2016	Rafat S Mohtasib et al. 2012
9.	Jung-Lung Hsu et al. 2012	Vonetta M Dotson et al. 2016
10.	Zhiye Chen et al. 2012	Estela Càmara et al. 2007
11.	Christopher M Marano et al. 2014	Harri Littow et al. 2010
12.	Olivia M Farr et al. 2016	Andrew P Merluzzi et al. 2016
13.	Dan-Miao Sun et al. 2017	Emily S Nichols et al. 2016
14.	Dewang Mao et al. 2015	Maria Morozova et al. 2016
15.	Rongfeng Qi et al. 2012	Kristen M Kennedy et al. 2009
16.	Dewang Mao et al. 2015	Chiara Chiapponi et al. 2013
17.	Xin Huang et al. 2016	Kathrin Cohen Kadosh et al. 2013
18.	Wenqing Xia et al. 2013	Quinton Deeley et al. 2008
19.	Po Lai Yau et al. 2009	Kristen M Kennedy et al. 2015
20.	Reza Tadayonnejad et al. 2019	Tatia M C Lee et al. 2006
21.	Chen Liu et al. 2014	Joshua Carp et al. 2011
22.	Yue Cheng et al. 2017	Esther H H Keulers et al. 2010
23.	Chuanming Li et al. 2014	Kristin Nordin et al. 2017
24.	Zhilian Zhao et al. 2014	Joshua Carp et al. 2010
25.	Xiaofen Ma et al. 2015	Mark B Schapiro et al. 2004
26.	Jessica A Turner et al. 2013	Nick S Ward et al. 2008
27.	Jiaxing Zhang et al. 2016	Nancy E Adleman et al. 2016
28.	Yingying Yue et al. 2015	Kaitlin L Bergfield et al. 2010
29.	Nicola Pannacciulli et al. 2006	Jenny R Rieck et al. 2017
30.	Xin Di et al. 2013	Marco Hirnstein et al. 2011

Supplementary Table 4. Study Estimates of Cognitive Meta-Analysis

Executive Function	**d = -0.40, K = 18, p < 0.001, Q = 186.9, I^2=88.4%**	
Studies	**Effect size (d)**	**Confidence Interval (95%)**
Bangen et al., 2015	-0.21	-0.33, -0.09
Biessels at al., 2001	-0.66	-1.41, 0.09
Brands et al., 2007	-0.54	-0.86, -0.21
Cui et al., 2014	-0.54	-1.07, -0.01
Garcia-Casares et al., 2014	-0.87	-1.45, -0.29
Kanaya et al., 2004	1.03	0.83, 1.24
Lindeman et al., 2001	-0.09	-0.26, 0.08
Liu et al., 2018	-0.54	-1.04, -0.04
Mehrebian et al., 2012	-1.35	-1.93, -0.76
Mogi et al., 2004	-0.36	-0.81, 0.08
Reijmer et al., 2016	-0.33	-0.80, 0.15
Ryan & Geckle. 2008	-0.31	-0.71, 0.08
Takeuchi et al., 2012	-0.62	-1.09, -0.15
Van den Berg et al., 2010	-0.65	-1.05, -0.24
Xia et al., 2010	-0.61	-1.07, -0.16
Yau et al., 2010	-0.14	-0.80, 0.51
Yeung et al., 2009	-0.50	-0.83, -0.18
Zhou et al., 2010	-0.48	-1.10, 0.15
Short-Term Verbal Memory	**d= –0.39, K=23, p<0.001, Q=143.6, I^2= 91.1%**	
Studies	**Effect size (d)**	**Confidence Interval (95%)**
Aberle et al., 2008	0.02	-0.31, 0.36
Arvanitakis et al., 2006	-0.05	-0.25, 0.14
Bangen et al., 2015	-0.13	-0.25, -0.02
Brands et al., 2007	-0.32	-0.64, 0.00
Cholerton et al., 2019	0.06	-0.13, 0.25
Cosway et al., 2001	-0.32	-0.81, 0.17
Cui et al., 2014	-0.07	-0.59, 0.46
Dai et al., 2017	-1.7	-2.24, -1.16
Garcia-Casares et al., 2014	-1.62	-2.26, -0.98
Liu et al., 2018	-0.49	-0.99, 0.01
Lowe et al., 1994	-0.06	-0.36, 0.25
Mattei et al., 2019	-0.3	-0.42, -0.19
Mehrebian et al., 2012	-1.71	-2.33, -1.10
Mogi et al., 2004	-0.29	-0.74, 0.16
Moran et al., 2013	0.31	0.16, 0.46
Reijmer et al., 2016	-0.23	-0.70, 0.24
Ryan & Geckle, 2008	-0.41	-0.81, -0.01
Takeuchi et al., 2012	-0.56	-1.03, -0.09
van den Berg et al., 2010	-0.39	-0.79, 0.01
van Harten et al., 2007	-0.35	-0.71, 0.01
Xia et al, 2015	-0.32	-0.77, 0.13
Yau et al., 2010	-0.70	-1.37, -0.03
Yeung et al., 2009	-0.37	-0.69, -0.04
Verbal Fluency	**d= –0.37, K=25, p<0.001, Q=101.3, I^2=83.3%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Aberle et al., 2008	0.06	-0.27, 0.39
Arvanitakis et al., 2006	-0.12	-0.31, 0.08
Atiea et al., 1995	-0.32	-0.86, 0.22
Bangen et al., 2015	-0.33	-0.44, -0.21
Brands et al., 2007	-0.50	-0.82, -0.17

Cholerton et al., 2019	-0.37	-0.56, -0.18
Cosway et al., 2001	-0.30	-0.79, 0.19
Dai et al., 2017	-0.83	-1.31, -0.35
Garcia-Casares et al., 2014	-0.92	-1.5, -0.34
Kanaya et al., 2004	-1.10	-1.31, -0.9
Liu et al., 2018	-0.44	-0.94, 0.05
Lowe et al., 1994	-0.46	-0.77, -0.15
Mankovsky et al., 2018	-0.32	-0.83, 0.19
Mattei et al., 2019	-0.33	-0.44, -0.21
Mehrebian et al., 2012	-0.18	-0.71, 0.34
Moran et al., 2013	-0.03	-0.18, 0.12
Rawlings et al., 2015	-0.36	-0.41, -0.31
Reijmer et al., 2016	-0.15	-0.62, 0.32
Solanki et al., 2009	-0.47	-0.93, -0.01
Takeuchi et al., 2012	-0.59	-1.06, -0.12
van Harten et al., 2007	-0.64	-1.01, -0.27
Yau et al., 2010	-0.07	-0.72, 0.59
Yeung et al., 2009	-0.15	-0.47, 0.17
Zhou et al., 2010	-0.58	-1.21, 0.06
Zihl et al., 2010	0.26	-0.46, 0.99
Working Memory	**d= –0.36, K=12, *p*<0.001, Q=30.2, I^2=72.2%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Arvanitakis et al., 2006	-0.04	-0.24, 0.15
Atiea et al., 1995	-0.48	-1.02, 0.07
Biessels et al., 2001	-0.63	-1.38, 0.12
Brands et al., 2007	-0.36	-0.69, -0.04
Lowe et al., 1994	-0.04	-0.35, 0.26
Mankovsky et al., 2018	-0.07	-0.57, 0.44
Mattei et al., 2019	-0.27	-0.38, -0.15
Ryan & Geckle, 2008	-0.34	-0.73, 0.06
Solanki et al., 2009	-1.33	-1.83, -0.84
Takeuchi et al, 2012	-0.56	-1.03, -0.10
van den Berg et al., 2010	-0.50	-0.90, -0.10
Yau et al., 2010	-0.17	-0.82, 0.49
Abstract Reasoning	**d= –0.36, K=8, *p*<0.001, Q=10.57, I^2=29.2%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Arvanitakis et al., 2006	-0.24	-0.43, -0.04
Bangen et al., 2015	-0.48	-0.59, -0.36
Brands et al., 2007	-0.19	-0.51, 0.13
Cosway et al., 2001	-0.10	-0.59, 0.38
Lowe et al., 1994	-0.41	-0.73, -0.10
Ryan & Geckle, 2008	-0.27	-0.67, 0.12
van den Berg et al., 2010	-0.38	-0.78, 0.02
Zihl et al., 2010	-1.06	-1.82, -0.29
Processing Speed	**d= –0.34, K=31, *p*<0.001, Q=227.3, I^2=82.3 %**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**

Aberle et al., 2008	0.00	-0.33, 0.33
Arvanitakis et al., 2006	-0.14	-0.34, 0.05
Atiea et al., 1995	0.05	-0.49, 0.59
Bangen et al., 2015	-0.22	-0.34, -0.11
Biessels et al., 2001	0.19	-0.54, 0.92
Brands et al., 2007	-0.26	-0.58, 0.06
Cholerton et al., 2019	-0.43	-0.62, -0.24
Cosway et al., 2001	-0.30	-0.79, 0.19
Cui et al., 2014	-0.68	-1.22, -0.14
Dai et al., 2017	0.25	-0.21, 0.72
Garcia-Casares et al., 2014	-0.66	-1.23, -0.09
Lindeman et al., 2001	-0.02	-0.19, 0.15
Liu et al., 2018	-0.36	-0.86, 0.13
Mattei et al., 2019	-0.37	-0.49, -0.25
Mehrebian et al., 2012	-1.23	-1.8, -0.65
Mogi et al., 2004	-0.59	-1.04, -0.14
Moran et al., 2013	0.16	0.02, 0.31
Naseer et al., 2014	-0.61	-1.25, 0.02
Rawlings et al., 2015	-0.65	-0.70, -0.60
Redondo et al., 2016	-0.54	-1.15, 0.07
Reijmer et al., 2016	-0.06	-0.53, 0.41
Ryan & Geckle, 2008	-0.43	-0.83, -0.03
Solanki et al., 2009	-0.93	-1.41, -0.46
Takeuchi et al., 2012	-0.61	-1.08, -0.14
van den Berg et al., 2010	-0.11	-0.51, 0.29
van Harten et al., 2007	-0.48	-0.84, -0.11
Xia et al, 2015	-0.27	-0.71, 0.18
Yau et al., 2010	-0.69	-1.37, -0.02
Yeung et al., 2009	-0.38	-0.70, -0.06
Zhou et al., 2010	-0.27	-0.89, 0.35
Zihl et al., 2010	-1.28	-2.07, -0.49

Visuospatial Reasoning	**d= –0.32, K=13, *p*<0.001, Q=27.3, I^2=56.6%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Arvanitakis et al., 2006	-0.11	-0.31, 0.09
Bangen et al., 2015	-0.18	-0.30, -0.07
Biessels et al., 2001	-0.25	-0.98, 0.49
Brands et al., 2007	-0.21	-0.53, 0.11
Garcia-Casares et al., 2014	-0.79	-1.36, -0.21
Lowe et al., 1994	-0.16	-0.47, 0.15
Mattei et al., 2019	-0.33	-0.45, -0.21
Moran et al., 2013	-0.56	-0.71, -0.41
Ryan & Geckle, 2008	-0.38	-0.78, 0.01
Takeuchi et al., 2012	-0.41	-0.87, 0.06
van den Berg et al., 2010	-0.16	-0.56, 0.24
Xia et al, 2015	-0.71	-1.17, -0.25
Zhou et al., 2010	-0.47	-1.10, 0.16

Delayed Verbal Memory	**d= –0.21, *p*<0.001, K=21, Q=114.3, I^2=77.5%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Arvanitakis et al., 2006	-0.01	-0.21, 0.18
Bangen et al., 2015	-0.07	-0.19, 0.05
Brands et al., 2007	-0.28	-0.60, 0.04
Cholerton et al., 2019	0.04	-0.15, 0.23
Cosway et al., 2001	-0.24	-0.73, 0.24
Cui et al., 2014	-0.08	-0.60, 0.45
Dai et al., 2017	-0.75	-1.22, -0.27
Liu et al., 2018	-0.62	-1.13, -0.12
Lowe et al., 1994	-0.10	-0.41, 0.21
Mehrebian et al., 2012	-0.76	-1.31, -0.22
Mogi et al., 2004	0.18	-0.26, 0.63
Moran et al., 2013	0.20	0.05, 0.35
Rawlings et al., 2015	-0.40	-0.45, -0.35
Reijmer et al., 2016	-0.13	-0.60, 0.34
Ryan & Geckle, 2008	-0.22	-0.61, 0.17
Takeuchi et al., 2012	-0.36	-0.82, 0.10
van den Berg et al., 2010	-0.40	-0.80, 0.00
van Harten et al., 2007	-0.42	-0.78, -0.05
Xia et al, 2015	-0.15	-0.59, 0.30
Yeung et al., 2009	-0.07	-0.39, 0.25
Zhou et al., 2010	-0.96	-1.62, -0.31
Short-Term Memory	**d= –0.21, *p*=0.005, K=16, Q=38.1, I^2=70.2%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Arvanitakis et al., 2006	-0.13	-0.33, 0.06
Atiea et al., 1995	-0.47	-1.01, 0.08
Biessels et al., 2001	-1.05	-1.83, -0.27
Brands et al., 2007	-0.08	-0.40, 0.24
Dai et al., 2017	-0.66	-1.14, -0.19
Lindeman et al., 2001	-0.17	-0.33, 0.00
Liu et al., 2018	-0.53	-1.03, -0.03
Lowe et al., 1994	0.14	-0.17, 0.45
Mankovsky et al., 2018	0.17	-0.34, 0.67
Mattei et al., 2019	-0.11	-0.22, 0.01
Moran et al., 2013	0.08	-0.07, 0.22
Naseer et al., 2014	0.04	-0.57, 0.66
Solanki et al., 2009	-0.92	-1.39, -0.44
Takeuchi et al., 2012	-0.33	-0.80, 0.13
van den Berg et al., 2010	-0.12	-0.52, 0.27
Yau et al., 2010	-0.30	-0.96, 0.36
Recognition Verbal Memory	**d= –0.21, *p*=0.01, K=12, Q=37.4, I^2=78.8%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Arvanitakis et al., 2006	0.00	-0.20, 0.20
Bangen et al., 2015	-0.06	-0.17, 0.06
Brands et al., 2007	-0.38	-0.70, -0.06

Dai et al., 2017	-0.61	-1.08, -0.14
Liu et al., 2018	-0.53	-1.02, -0.03
Lowe et al., 1994	-0.07	-0.38, 0.24
Mattei et al., 2019	-0.17	-0.28, -0.05
Mehrebian et al., 2012	-0.46	-1.00, 0.07
Moran et al., 2013	0.16	0.01, 0.31
Takeuchi et al., 2012	0.12	-0.34, 0.58
van den Berg et al., 2010	-0.51	-0.92, -0.11
Zhou et al., 2010	-0.89	-1.55, -0.24
Visual Memory	**d= –0.13, *p*=0.32, K=8, Q=19.9, I^2=66.6%**	
Studies	**Effect Size (d)**	**Confidence Interval (95%)**
Aberle et al., 2008	-0.16	-0.49, 0.18
Brands et al., 2007	0.19	-0.13, 0.51
Cosway et al., 2001	-0.10	-0.59, 0.39
Lowe et al., 1994	0.20	-0.11, 0.51
Solanki et al., 2009	-0.96	-1.43, -0.48
Takeuchi et al., 2012	-0.27	-0.73, 0.19
van den Berg et al., 2010	0.00	-0.39, 0.40
Yau et al., 2010	-0.13	-0.79, 0.52

Supplementary Table 5. Characteristics of patients who underwent cognitive testing in studies included in our meta-analysis.

Studies		*N*		Age (Mean)		Education (Years)		Female (%)
Author	**Year**	**T2DM**	**HC**	**T2DM**	**HC**	**T2DM**	**HC**	
Aberle et al.	2008	38	421	62.9	62.97	9.94	9.93	48.5
Arvanitakis et al.	2006	116	766	78	80.9	13.7	14.5	78
Atiea et al.	1995	40	20	69.05	68.1	-	-	31
Bangen et al.	2015	378	1115	75.4	76.3	9.9	11.2	67
Biessels et al.	2001	13	16	57.7	57.9	11.2	11.4	41.3
Brands et al.	2007	119	55	65.9	65.2	4*	4	49.4
Cholerton et al.	2019	185	261	53	51.3	12	12.5	70.4
Cosway et al.	2001	33	32	57.7	55.9	11.2	11.8	59.2
Cui et al.	2014	29	27	58.3	57.8	10.4	10.2	55.4
Dai et al.	2017	41	32	65.51	67.28	15.35	16.05	52
Garcia-Casares et al.	2014	25	25	60	57.8	18.3	18.9	38
Kanaya et al.	2004	118	632	73.55	69.2	-	-	57.2
Lindeman et al.	2001	188	476	73.4	73.8	10.9	12.3	-
Liu et al.	2018	32	32	58.09	56.88	9	12	42.2
Lowe et al.	1994	80	81	59.3	55.1	-	-	63.9
Mankovsky et al.	2018	93	18	62.3	59.5	14.7	14.3	70.2
Mattei et al.	2019	465	711	58.9	56	-	-	73.1
Mehrebian et al.	2012	37	22	56	56	14	14	56.5
Mogi et al.	2004	69	27	71.6	73.4	10.4	11.4	64.5
Moran et al.	2013	350	363	67.8	72.1	11.3	10.9	43.2
Naseer et al.	2014	20	20	53.3	-	-	-	-
Rawlings et al.	2015	1779	11572	58.2	56.8	-	-	55.6
Redondo et al.	2016	20	23	70.82	70.92	6.79	7.36	46
Reijmer et al.	2016	35	35	71.1	71	4*	4	41.4
Ryan & Geckle	2008	50	50	50.8	50.5	14.4	14	73
Solanki et al.	2009	50	30	-	-	-	-	-
Takeuchi et al.	2012	42	32	62.4	63.8	13.7	14.5	40
van den Berg et al.	2010	68	38	65.6	64.8	4*	4	48.1
van Harten et al.	2007	92	44	73.2	72.9	4*	4.4	55.8
Xia et al.	2015	38	40	56	57.1	9.6	10.3	51.3
Yau et al.	2010	18	18	16.46	17.16	10.75	11.15	-
Yeung et al.	2009	41	424	68.59	67.84	15.12	15.33	68
Zhou et al.	2010	21	19	68	69.16	12.48	13.84	50
Zihl et al	2010	12	19	42.45	36.8	10.5	10.9	-

T2DM, type-2 diabetes mellitus; HC, healthy control
**median education*

Supplementary Table 6. List and justification for studies excluded from our cognitive meta-analysis.

Studies	Justification for Exclusion
Asimakopoulou et al., 2002	Did not match for education
Brands et al., 2007	Identical sample of study already included
Bruehl et al., 2009	Authors did not provide requested data
Callisaya et al., 2018	Identical sample of study already included
Chen et al., 2014	Inadequate cognitive testing
Chen et al., 2017	Inadequate cognitive testing
Christman et al., 2010	Did not match for education
Cooray et al., 2011	Authors did not provide requested data
Cui et al., 2015	Identical sample of study already included
Cui et al., 2017	Identical sample of study already included
Cui et al., 2017	Identical sample of study already included
Degen et al., 2016	Authors did not provide requested data
Dey et al., 1997	Inadequate cognitive testing
Dore et al., 2009	Authors did not provide requested data
Elias et al., 1997	No baseline data in longitudinal design
Grodstein et al., 2001	Inadequate cognitive testing
Hassing et al., 2004	Unclear sizes of sample sub-groups
Helkala et al., 1995	Did not match for education
Kinga & Anett, 2016	Authors unable to be reached
Kumari et al., 2005	Did not match for education
Liu et al., 2016	Identical sample of study already included
Liu et al., 2018	Identical sample of study already included
Liu et al., 2020	Identical sample of study already included
Manschot et al., 2006	Identical sample of study already included
Mooradian et al., 1988	Inadequate cognitive testing
Nazaribadie et al., 2013	Authors did not provide requested data
Nealon et al., 2017	Did not match for education
Nooyens et al., 2010	Authors did not provide requested data
Perlmuter et al., 1984	Inadequate cognitive testing
Ravona-Springer et al., 2018	Authors did not provide requested data
Reijmer et al., 2011	Identical sample of study already included
Robertson-Tchabo et al., 1986	Inadequate cognitive testing
Ruis et al., 2009	Authors did not provide requested data
Scott et al., 1998	Identical sample of study already included
Sinclair et al., 2000	Inadequate cognitive testing
Smith et al., 2009	Identical sample of study already included
Spauwen et al., 2015	Authors did not provide requested data
van Gemert et al., 2018	Authors did not provide requested data
Watari et al., 2006	Authors did not provide requested data
Xia et al., 2013	Identical sample of study already included
Xia et al., 2015	Identical sample of study already included

References

1. Kaiser AB, Zhang N, Der Pluijm WV. Global Prevalence of Type 2 Diabetes over the Next Ten Years (2018-2028). *Diabetes.* 2018;67(Supplement 1):202-LB.
2. Moheet A, Mangia S, Seaquist ER. Impact of diabetes on cognitive function and brain structure. *Ann N Y Acad Sci.* 2015;1353:60-71.
3. Callisaya ML, Beare R, Moran C, Phan T, Wang W, Srikanth VK. Type 2 diabetes mellitus, brain atrophy and cognitive decline in older people: a longitudinal study. *Diabetologia.* 2019;62(3):448-458.
4. Gold SM, Dziobek I, Sweat V, et al. Hippocampal damage and memory impairments as possible early brain complications of type 2 diabetes. *Diabetologia.* 2007;50(4):711-719.
5. Mankovsky B, Zherdova N, Van den berg E, Biessels G, de Bresser J. Cognitive functioning and structural brain abnormalities in people with Type 2 diabetes mellitus. *Diabetic Medicine.* 2018;35.
6. Manschot SM, Brands AM, van der Grond J, et al. Brain magnetic resonance imaging correlates of impaired cognition in patients with type 2 diabetes. *Diabetes.* 2006;55(4):1106-1113.
7. Rawlings AM, Sharrett AR, Schneider AL, et al. Diabetes in midlife and cognitive change over 20 years: a cohort study. *Ann Intern Med.* 2014;161(11):785-793.
8. Zilliox LA, Chadrasekaran K, Kwan JY, Russell JW. Diabetes and Cognitive Impairment. *Curr Diab Rep.* 2016;16(9):87.
9. Pelimanni E, Jehkonen M. Type 2 Diabetes and Cognitive Functions in Middle Age: A Meta-Analysis. *J Int Neuropsychol Soc.* 2018:1-16.
10. Biessels G, Staekenborg S, Brunner E, Brayne C, Scheltens P. Risk of dementia in diabetes mellitus: a systematic review [published correction appears in Lancet]. *Lancet Neurol.* 2006;5(2):113.
11. Ninomiya T. Diabetes mellitus and dementia. *Curr Diab Rep.* 2014;14(5):487.
12. Tuligenga RH. Intensive glycaemic control and cognitive decline in patients with type 2 diabetes: a meta-analysis. *Endocr Connect.* 2015;4(2):R16-24.
13. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med.* 2015;12(3):e1001779.
14. Seabold S, Perktold J. Statsmodels: Econometric and statistical modeling with python. Paper presented at: Proceedings of the 9th Python in Science Conference2010.
15. Alfaro-Almagro F, Jenkinson M, Bangerter NK, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *Neuroimage.* 2018;166:400-424.
16. Zang YF, He Y, Zhu CZ, et al. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev.* 2007;29(2):83-91.
17. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res.* 1996;29(3):162-173.
18. Cox RW, Hyde JS. Software tools for analysis and visualization of fMRI data. *NMR Biomed.* 1997;10(4-5):171-178.
19. Fieller EC. Some problems in interval estimation. *Journal of the Royal Statistical Society: Series B (Methodological).* 1954;16(2):175-185.
20. Bower JK, Meadows RJ, Foster MC, Foraker RE, Shoben AB. The Association of Percent Body Fat and Lean Mass With HbA1c in US Adults. *J Endocr Soc.* 2017;1(6):600-608.
21. Bae JP, Lage MJ, Mo D, Nelson DR, Hoogwerf BJ. Obesity and glycemic control in patients with diabetes mellitus: Analysis of physician electronic health records in the US from 2009-2011. *J Diabetes Complications.* 2016;30(2):212-220.
22. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *bmj.* 2021;372.

23. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2018.
24. Viechtbauer W. Conducting Meta-Analyses in R with The metafor Package. *Journal of Statistical Software.* 2010;36.
25. Cohen J. *Statistical power analysis for the behavioral sciences.* 2nd ed. Hillsdale, N.J.: L. Erlbaum Associates; 1988.
26. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj.* 2003;327(7414):557-560.
27. Dockès Jm, Poldrack RA, Primet R, et al. NeuroQuery, comprehensive meta-analysis of human brain mapping. *Elife.* 2020;9:e53385.
28. Fox PT, Parsons LM, Lancaster JL. Beyond the single study: function/location metanalysis in cognitive neuroimaging. *Current opinion in neurobiology.* 1998;8(2):178-187.
29. Van Horn JD, Grafton ST, Rockmore D, Gazzaniga MS. Sharing neuroimaging studies of human cognition. *Nature neuroscience.* 2004;7(5):473-481.
30. Yarkoni T, Poldrack RA, Van Essen DC, Wager TD. Cognitive neuroscience 2.0: building a cumulative science of human brain function. *Trends in cognitive sciences.* 2010;14(11):489-496.
31. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. *Nature methods.* 2011;8(8):665-670.
32. Tomasi D, Wang GJ, Volkow ND. Energetic cost of brain functional connectivity. *Proceedings of the National Academy of Sciences of the United States of America.* 2013;110(33):13642-13647.
33. Weistuch C, Mujica-Parodi LR, Amgalan A, Dill KA. Metabolism modulates global synchrony in the aging brain. *bioRxiv.* 2020.
34. Tomasi D, Volkow ND. Aging and functional brain networks. *Molecular psychiatry.* 2012;17(5):471, 549-458.
35. Mujica-Parodi LR, Amgalan A, Sultan SF, et al. Diet modulates brain network stability, a biomarker for brain aging, in young adults. *Proceedings of the National Academy of Sciences of the United States of America.* 2020;117(11):6170-6177.
36. El Messari S, Ait-Ikhlef A, Ambroise DH, Penicaud L, Arluison M. Expression of insulin-responsive glucose transporter GLUT4 mRNA in the rat brain and spinal cord: an in situ hybridization study. *J Chem Neuroanat.* 2002;24(4):225-242.
37. Kobayashi M, Nikami H, Morimatsu M, Saito M. Expression and localization of insulin-regulatable glucose transporter (GLUT4) in rat brain. *Neurosci Lett.* 1996;213(2):103-106.
38. El Messari S, Leloup C, Quignon M, Brisorgueil MJ, Penicaud L, Arluison M. Immunocytochemical localization of the insulin-responsive glucose transporter 4 (Glut4) in the rat central nervous system. *J Comp Neurol.* 1998;399(4):492-512.
39. Vannucci SJ, Koehler-Stec EM, Li K, Reynolds TH, Clark R, Simpson IA. GLUT4 glucose transporter expression in rodent brain: effect of diabetes. *Brain Res.* 1998;797(1):1-11.
40. Prast H, Fischer H, Werner E, Werner-Felmayer G, Philippu A. Nitric oxide modulates the release of acetylcholine in the ventral striatum of the freely moving rat. *Naunyn Schmiedebergs Arch Pharmacol.* 1995;352(1):67-73.
41. Kurotani S, Umegaki H, Ishiwata K, Suzuki Y, Iguchi A. The age-associated changes of dopamine-acetylcholine interaction in the striatum. *Exp Gerontol.* 2003;38(9):1009-1013.
42. Ferreira-Vieira TH, Guimaraes IM, Silva FR, Ribeiro FM. Alzheimer's disease: Targeting the Cholinergic System. *Curr Neuropharmacol.* 2016;14(1):101-115.
43. Muramatsu I, Uwada J, Yoshiki H, et al. Novel regulatory systems for acetylcholine release in rat striatum and anti-Alzheimer's disease drugs. *J Neurochem.* 2019;149(5):605-623.
44. Tessari P, Cecchet D, Cosma A, et al. Nitric oxide synthesis is reduced in subjects with type 2 diabetes and nephropathy. *Diabetes.* 2010;59(9):2152-2159.
45. Sokoloff L, Mangold R, Wechsler RL, Kennedy C, Kety SS. Effect of Mental Arithmetic on Cerebral Circulation and Metabolism. *Journal of Clinical Investigation.* 1955;34(7):1101-1108.

46. Clark DD, Sokoloff L. In: Siegel GJ, Agranoff BW, Albers RW, Risher SK, Uhler MD, eds. *Basic Neurochemistry: Molecular, Cellular and Medical Aspects.* Philadelphia: Lippincott; 1999:637-670.