medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract (Words): 350
Manuscript (Words): 4500
Main Text, Figures: 4
Main Text, Tables: 0
Supplementary Materials, Figures: 21
Supplementary Materials, Tables: 6
References: 52

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]

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

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Declaration of Interests and Sources of Funding

The research described in this paper was funded by the W. M. Keck Foundation (to LRMP), the White House Brain Research Through Advancing Innovative Technologies (BRAIN) Initiative (NSFNCS-FR 1926781 to LRMP), and the Baszucki Brain Research Fund (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.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 separately 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 the identified neurocognitive effects.

Findings: The UK Biobank dataset included cognitive and neuroimaging data (N=20,314) including 1,012 T2DM and 19,302 HC, aged between 50 and 80 years. Duration of T2DM ranged from 0–31 years (mean 8.5±6.1 years); 498 were treated with metformin alone, while 352 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=1,088). As compared to age, sex, education, and hypertension-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* and *premotor cortex* and increased in the *subgenual area*, *orbitofrontal cortex*, *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 26% ± 14%; 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.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Approximately 6.6% of the global population carries 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 reveal themselves 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, education and hypertension-matched healthy controls (HC). To do so, we leverage the robust statistical power made possible by UK Biobank [13], the largest (N=20,314, ages 50-80y) neurocognitive lifespan dataset to date. The UK Biobank results are 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 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=20,314)

General Overview: UK Biobank data were analyzed for both cognitive and neuroimaging data. Datafield identifiers for all utilized features are shown in **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 exact pairwise matched for not only age, but also sex, education, and hypertension status. T2DM status was assessed based on self-reported diagnosis by doctor. We considered education as a binary variable based on possession of a college degree. Hypertension was quantified using measured blood pressure values: all individuals with systolic blood pressure >140 mmHg or diastolic blood pressure >90 mmHg were labeled with hypertension [14]. To exclude potential confounding

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

effects due to menopausal transition, particularly relevant to the youngest age group in our sample (age cutoff >50y), we excluded all female subjects who did not report menopause or if they reported ongoing hormone therapy [15]. To minimize the number of individuals with Type 1 diabetes in our sample, rather than Type 2 diabetes, we only included diabetic individuals with a self-reported age of onset >=40 years [16].

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 linear continuous factor with increments in years. The latter was justified given the linear trends we observed across age in all modalities (**SI Figs. 3, 4, and 5**). Regression models were fit using the Statsmodels Python library[17]. To account for multiple comparisons, we applied Bonferroni correction to cognitive and structural results, and adjusted for false discovery rate in our brain activation results; we report adjusted *p* values exclusively.

To determine whether T2DM neurocognitive effects suggested an acceleration of typical aging trajectories, versus non-aging-related degenerative pathways specific to T2DM, 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 18,829 participants (T2DM: N=914, HC: N=17,915) were extracted from the UK Biobank dataset, including *abstract reasoning*, *executive function*, *processing speed*, *reaction time*, and *numeric memory* (~2-3 seconds). Exact sample sizes varied across cognitive domains based on data availability, and therefore are noted separately for each result (**SI Figs. 1C and 2C**). 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. 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=821, HC: N=821, **SI Fig. 1A**) as compared to non-T2DM-specific age-related effects (N=4,775, **SI Fig. 2A**) on atrophy of gray matter volume; these findings were available in units of mm^3 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 normalized 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 of 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=646; HC: N=646, **SI Fig. 1B**) as compared to non-T2DM-specific age-related effects (N=2,250, **SI Fig. 2B**) on resting-state brain activity. Data were accessed already preprocessed by UK Biobank according to their standard pipelines[18]. 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 activation by calculating the *amplitude of low-frequency fluctuation*[19] (ALFF). We used the program *3dRSFC*, which is a component of Analysis of Functional NeuroImages[20, 21] (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).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 self-reported values, we averaged the reported age of onset values 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 [22] 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 activation. 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, education, and hypertension, and 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 included BMI as a regressor since it was the only available proxy measure for disease severity [23, 24].

Implications of sex: To determine whether results detected in the sample might be driven by sex-specific factors, we additionally performed analyses separately in females and males. For these, we evaluated our cognitive and neurobiological measures in association with age and T2DM and quantified the overlap separately for the two subsamples.

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 [25].

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.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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, executive function, processing speed, numeric memory, visual memory, verbal memory, verbal fluency, visuospatial reasoning, and working memory* (**SI Table 2**).

Statistical analyses were performed using R version 3.6.1 [26] and the Metafor package version 2.4-0 [27]. 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 [28]; 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 [29]. 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[30] to conduct a meta-analysis of published neurobiological results associated with T2DM and age. NeuroQuery is an automated Coordinate-Based Meta-Analysis (CBMA)[31-33] tool based on a database of z-scores collected by crawling through texts and tables of published research articles by an automated algorithm[34]. 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[34]. 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-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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 1.9% ± 0.1% decrease in performance per year (N=2,450, T= –17.2, p<1e–10) and *processing speed*, which showed 1.5% ± 0.2% decrease in performance per year (N=2,525, T= –22.8, p<1e–10;). Our analyses identified further cognitive deficits associated with T2DM that were consistent with accelerated age-related cognitive decline (**Fig. 1B)**. As with aging, the strongest T2DM effects were also observed in *executive function*, which showed a further 13.1% ± 6.9% decrease in performance, beyond age-related effects (T2DM: N=446; HC: N=446; T= –3.7, p=0.001), and *processing speed*, which showed a further 6.7% ± 3.2% decrease in performance, beyond age-related effects. (T2DM: N=454; HC: N=454; T= –4.1, p=0.0002). A more modest decline (3.7% ± 2.3%) was observed in *numeric memory* (~2-3 seconds) (T2DM: N=483; HC: N=483; T= –3.2, p=0.007), whereas *abstract reasoning* (T2DM: N=886; HC: N=886; T= –2.4, p=0.08) and *reaction time* (T2DM: N=914; HC: N=914; T= –1.0, p=0.32) were not statistically significant. Our meta-analysis confirmed that individuals with T2DM exhibited markedly lower performance when compared to age and education-matched controls, over an even broader set of domains (**Fig. 1C**). These again included *executive function* (K=18, d= –0.40, p=0.009), *processing speed* (K=31, d= –0.34, p=5e–8), and *numeric memory* (~2-3 seconds) (K=16, d= –0.21, p=0.05), as well as *abstract reasoning* (K=8, d= –0.36, p=1e–7), *immediate (~30 seconds) verbal memory* (K=23, d= –0.39, p=0.001), *delayed* (~20 minute) *verbal memory* (K=21, d= –0.21, p=0.005), *verbal fluency* (K=25, d= –0.37, p=2e–8), *visuospatial reasoning* (K=13, d= –0.32, p=4e–7), and *working memory* (K=12, d= –0.36, p=0.002 (**SI Table 4)**.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

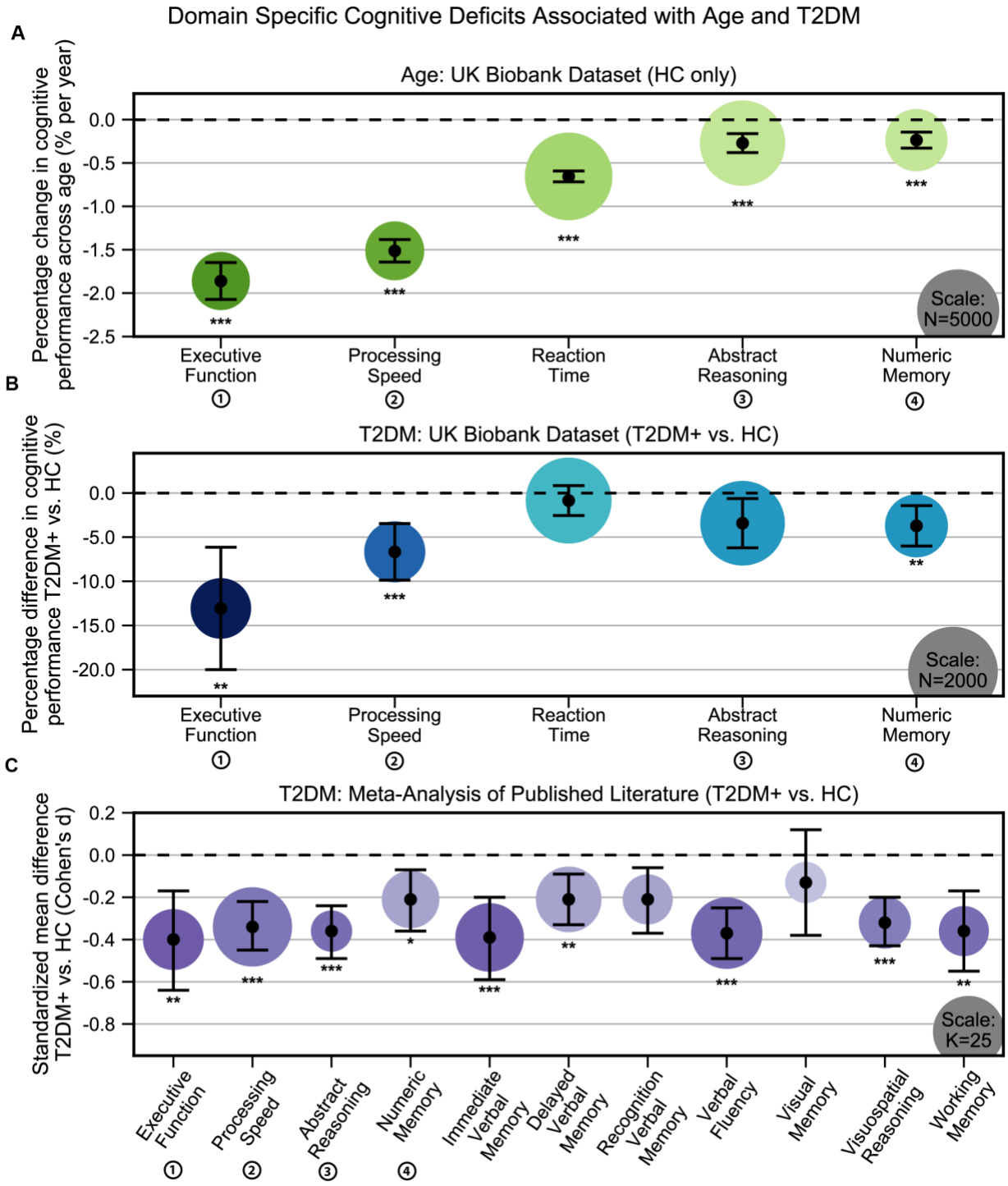


Figure 1: Cognitive deficits are apparent with respect to both age and T2DM diagnosis. A: Using the UK Biobank dataset, we performed a quantitative analysis of the effects related to age on cognitive performance across five cognitive domains. Associated changes were derived from estimated regression 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 and education-matched HC. Numbers next to labels identify domains common across panels. Marker sizes represent sample sizes scaled (per area) as indicated in the bottom right corner of each panel. On Panel C, sample size indicates the number of individual studies. Underlying sample size distributions can be found in **SI Figures 1C and 2C**. Error bars are 95% CI. *$P \leq 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$, Bonferroni corrected.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Neurobiological Correlates with Age and T2DM

Brain Atrophy: HC (N =4,775) showed a linear decrease in brain gray matter with age. This was most pronounced in the *ventral striatum*, which showed a 1.0% ± 0.06% decrease per year (T= –30.4, p<1e–10) and *Heschl's gyrus*, which also showed a 0.9% ± 0.06% decrease per year (T= –30.4, 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=821; HC: N=821). These included both cortical and subcortical regions, with the most severe atrophy observed in the *ventral striatum*, which showed on average a 6.2% ± 1.6% further decrease in volume, beyond age-related effects (T= –7.5, p<1e-10), in the *cerebellum* with an additional 4.9% ± 1.1% decrease in volume, beyond age-related effects (T= –8.8, p<1e-10), and in the *putamen*, which showed a 4.7% ± 2.3% further decrease, beyond age-related effects (T= –4.1, p=0.002) **(Fig. 2B).**

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

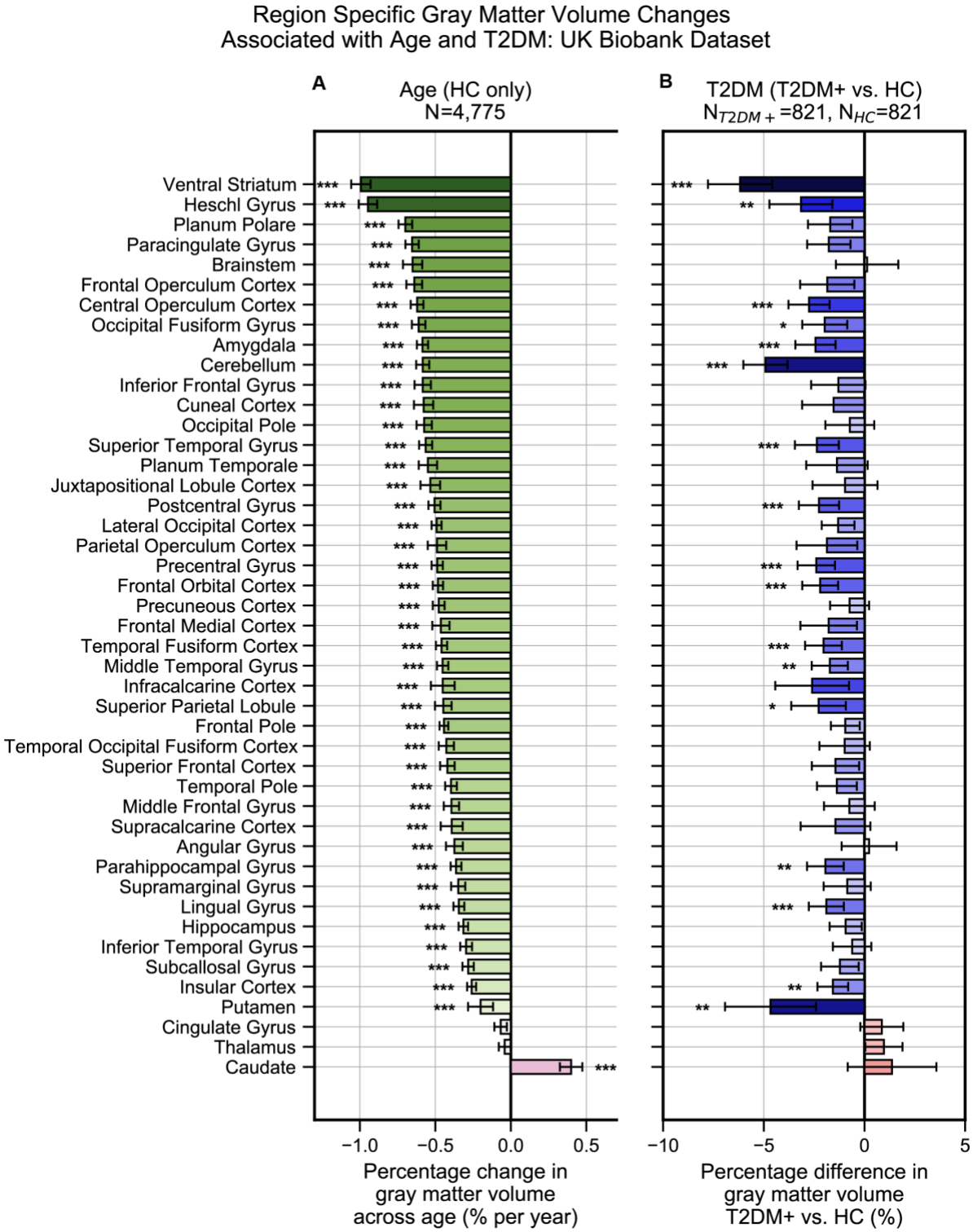


Figure 2: Widespread gray matter atrophy can be observed with respect to both age and T2DM diagnosis status. Using the UK Biobank dataset, we measured gray matter atrophy across 45 anatomical regions. Associated changes were derived from estimated regression coefficients as percentages and are shown 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*. Underlying sample size distributions can be found in **SI Figures 1A and 2A**. Error bars are 95% CI. $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$, Bonferroni corrected.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Brain Activity: Age was associated with functional reorganization of brain activation (ALFF), rather than global decrease or increase. Brain activation in T2DM showed similar reorganization. Normalized to whole brain activation, both age (HC: N=2,250) and T2DM (T2DM: N=646, HC: N=646) were associated with decreased activation in the *caudate and premotor cortex*, and with increased brain activity in the *subgenual area, orbitofrontal cortex, posterior cingulate cortex* 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 (*premotor cortex, caudate, posterior cingulate gyrus*), but additionally identified clusters of decreased activity in *Broca area* and the *frontal eye fields* and increased activity in the *thalamus*, and *inferior temporal gyrus* (**Fig. 3B**).

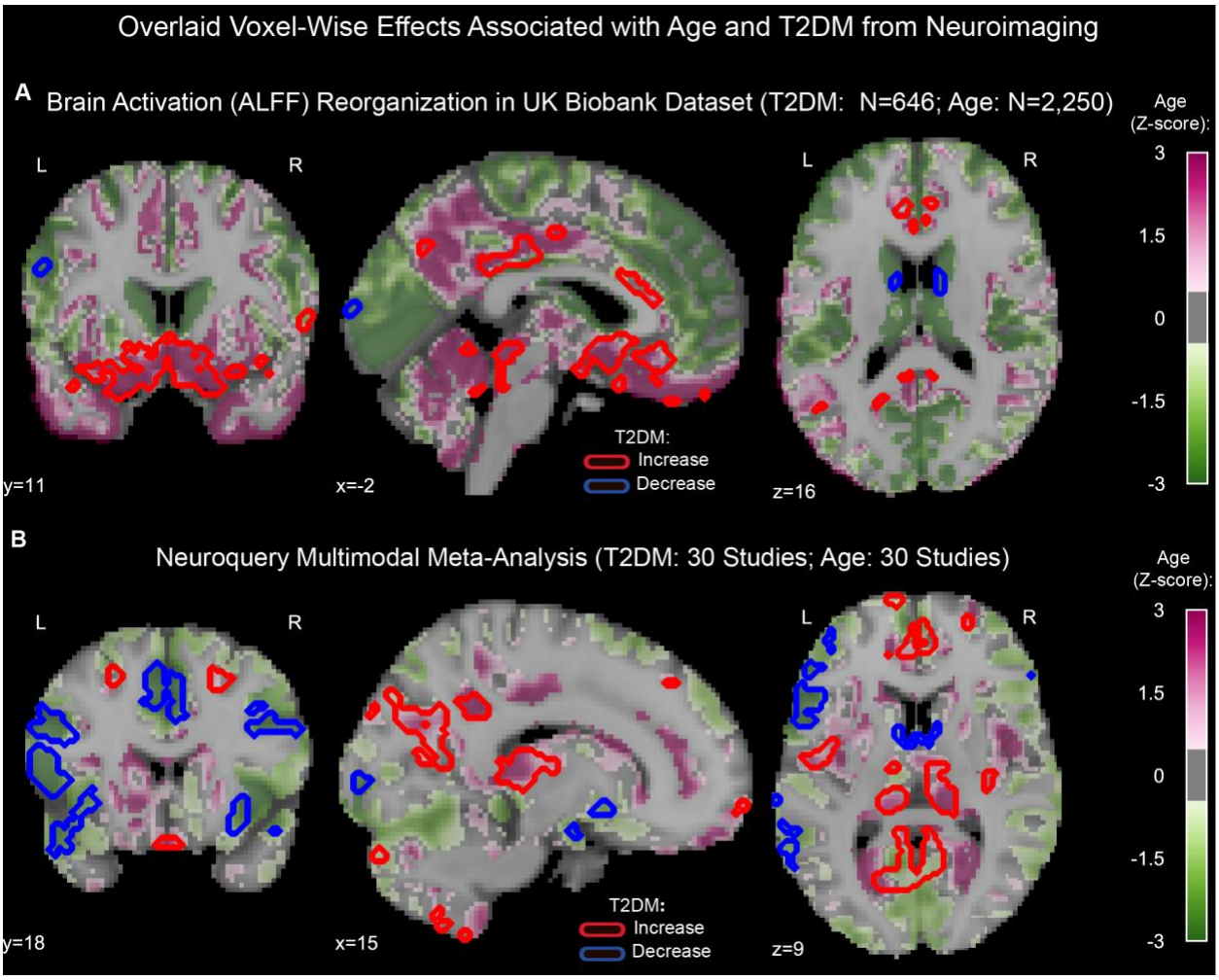


Figure 3: Age and T2DM associated reorganization patterns in brain activity are significant and overlap. A: For functional MRI data obtained from the UK Biobank dataset, we used the amplitude of low-frequency fluctuation (ALFF) to quantify brain activation. Effects linked to age are shown in the form 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) *orbitofrontal cortex* (increased), *the premotor cortex* (decreased) and the *caudate* (decreased). The highlighted regions were similarly impacted across age, indicating substantial overlap between the two contrasts. Underlying sample size distributions can be found in **SI Figures 1B and 2B**. **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 present in several regions, most importantly in the *posterior cingulate gyrus, thalamus, caudate, and premotor cortex*. These results support the hypothesis that neurodegeneration in both T2DM and aging may be associated with common mechanistic pathways.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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. Even after controlling for education, cognitive deficits remained statistically significant. Both age and T2DM implicated the same areas of greatest vulnerability: for brain atrophy, these were the *ventral striatum, Heschl's gyrus and cerebellum*; 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=0.60, p=0.0002) (**SI Figs. 7 and 8A**). Similarly, T2DM-related changes in brain activation (ALFF) also exhibited significant overlap with those associated with age (r=0.64, p=0.00004) (**SI Figs. 7 and 8B**). 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 Figs. 7 and 8C**).

T2DM Chronicity Exacerbates Neurocognitive Symptoms

Neurocognitive effects were more severe with increased disease duration, particularly for structural changes (T= –3.8, p=0.0002) (**Fig. 4**). T2DM progression was associated with 26% ± 14% acceleration of typical neurogenerative age-related effects, as per the linear shift along the horizontal time axis shown in **Fig 4**.

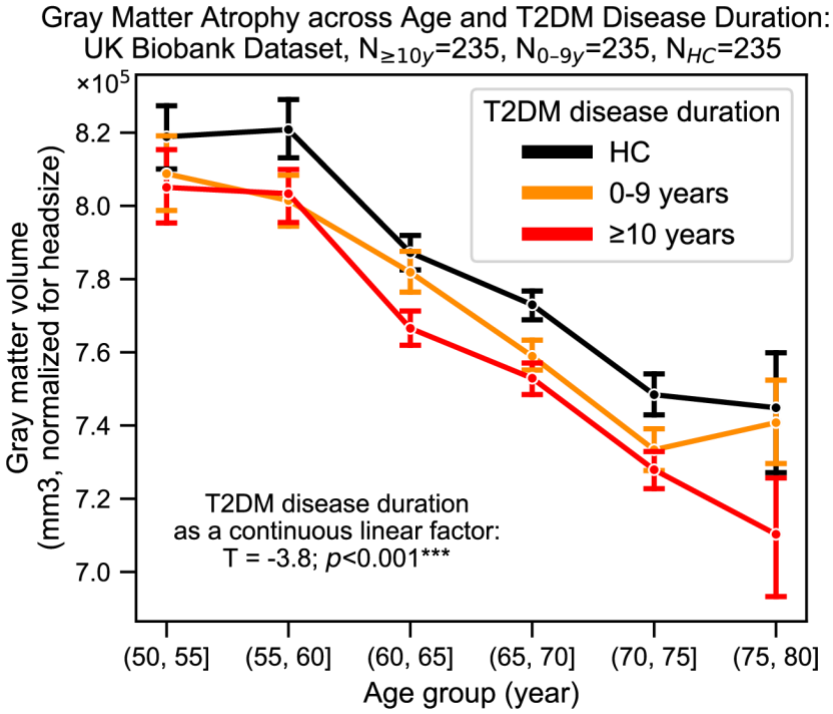


Figure 4: Progression of T2DM disease is significantly associated with gray matter atrophy, accelerating neurodegenerative effects seen in brain aging. For a quantitative evaluation of the impact of T2DM progression on gray matter volume, we considered time since T2DM diagnosis as the main factor of 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, and hypertension 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 ~0.26 ± 0.14 years of brain aging beyond that of age-matched T2DM–. Underlying sample size distributions can be found in **SI Figure 3**. Error bars are standard error of the mean. *$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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=498) did not differ with respect to cognition or brain atrophy compared to T2DM patients who were unmedicated (N=352) (**SI Fig. 21**). Likewise, treatment status showed no significant impact on resting-state brain activity.

Age and T2DM Associated Effects were Consistent in Females and Males

For the UK Biobank dataset, we identified marked sex-related differences both in neurobiological and cognitive measures, consistent with the literature (**SI Figs. 9, 11, and 16**) [35, 36]. All reported age and T2DM effects were seen for both males and females, and strongly correlated between them **(SI Figs. 14, 18**). However, age and T2DM effects were stronger in males (**SI Figs. 10, 13, and 17**), and neurodegeneration (brain atrophy, in particular) was more similar across sex for age (r_{age}=0.93, Bonferroni corrected $p \leq 0.001$) than for T2DM (r_{T2DM}=0.51, Bonferroni corrected $p \leq 0.01$). Consistent with the latter, the overlap between effects associated with age and T2DM was statistically significant in males (gray matter volume: r=0.77, p=4e–9; brain activation (ALFF): r=0.51, p=0.002) but not in the female-only subsample (**SI Figs. 14, 15, 18, and 19**). All females in our UK Biobank sample were menopausal and not on hormone replacement therapy. For our meta-analyses, we were unable to perform a sex-based comparison, as the underlying articles from which our data were derived from did not always perform sex-matching across their subject pools and did not provide access to the individual subject-level data that would be required to control for sex in our analyses.

Discussion

The UK Biobank dataset confirms that T2DM patients show deficits in cognitive performance compared to HC, even after controlling for age, sex, education, and hypertension. These findings were supported by meta-analysis of the published literature. Deficits in cognitive performance were accompanied by marked brain atrophy in the T2DM sample as compared to age-matched HC. The atrophy was most severe (6.4% 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 severity for all regions increased with longer disease duration. We detected qualitatively consistent results in females and males; however, males exhibited stronger effects in relation to T2DM (**SI Fig. 12**). This result is consistent with the well-established neuroprotective effects of female hormones such as estrogen[37]. This result also suggests that the T2DM neurological effects observed result from chronic degenerative processes which, for our female participants, may have been at least partially ameliorated prior to menopause.

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 far more statistically robust (p<2e–10) than either global cognitive measures (p=0.0001) or global brain activation (p=0.002). One additional advantage of structural MRI over cognitive testing is that the former avoids confounding associated with education and practice effects, and therefore may be more interpretable in real-world clinical settings in which matching for education and practice effects 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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

in the face of diminishing access to glucose, a pattern previously documented in aging[38-41]. 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.

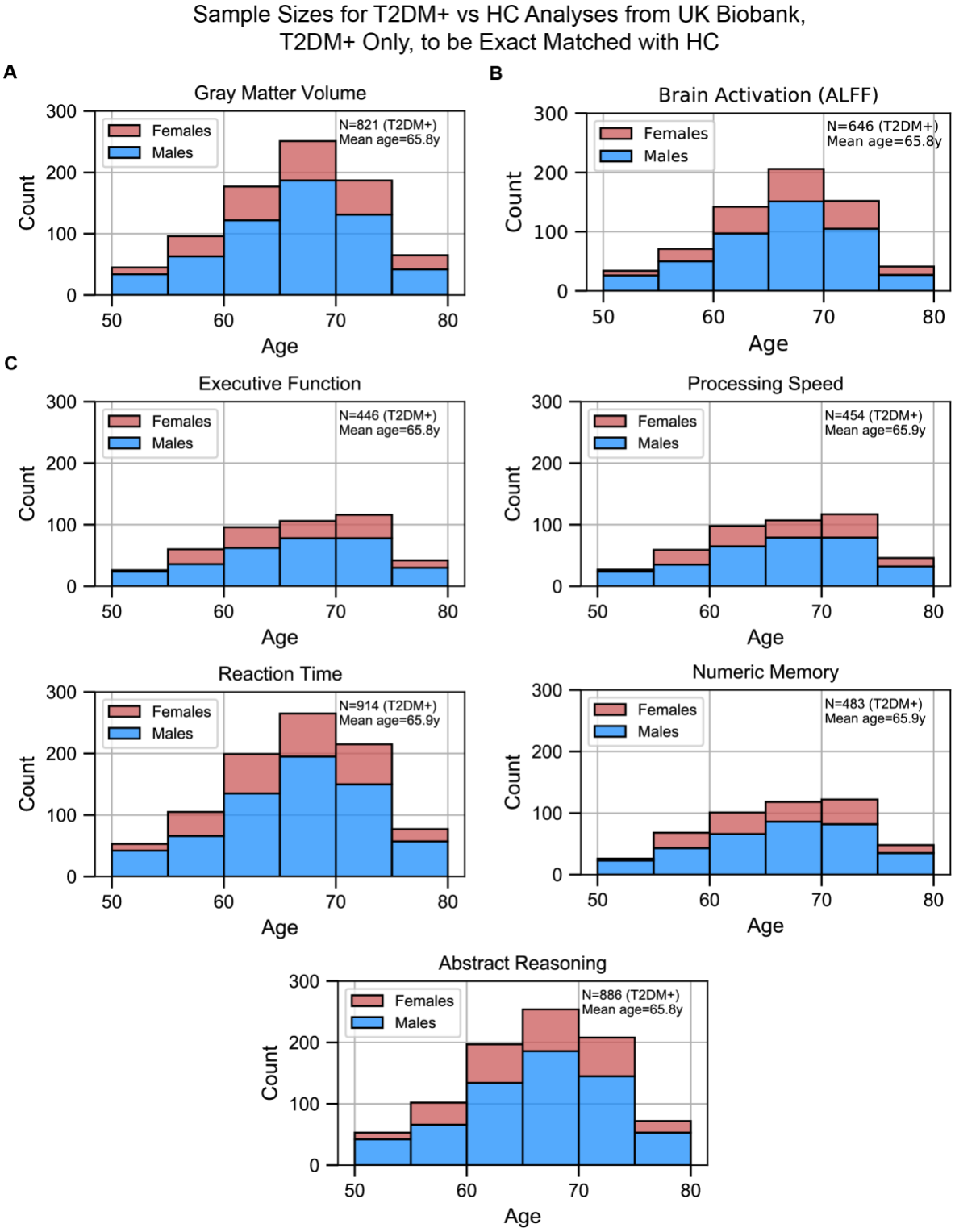
The localization of brain atrophy in T2DM to the *ventral striatum*, followed by the *cerebellum*, may reflect the fact that these two brain regions contain the densest concentrations of insulin-dependent GLUT-4[42-45], as compared to non-insulin-dependent isoforms GLUT-1 and GLUT-3. The *ventral striatum* 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[46], a neurotransmitter severely reduced in dementia[47] and a target for its pharmaceutical treatment[48, 49]. Release of nitric oxide is insulin dependent and reduced in T2DM[50]. 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 on one hand suggests that neurocognitive changes seen in T2DM may progress via a common mechanistic trajectory as normal brain aging, but which is accelerated. From another perspective, the overlap implies that brain aging itself is largely a metabolic syndrome driven by impaired insulin signaling and glucose metabolism, the same processes that are well-established with respect to T2DM.

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[23, 24]), 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 [51, 52], 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].

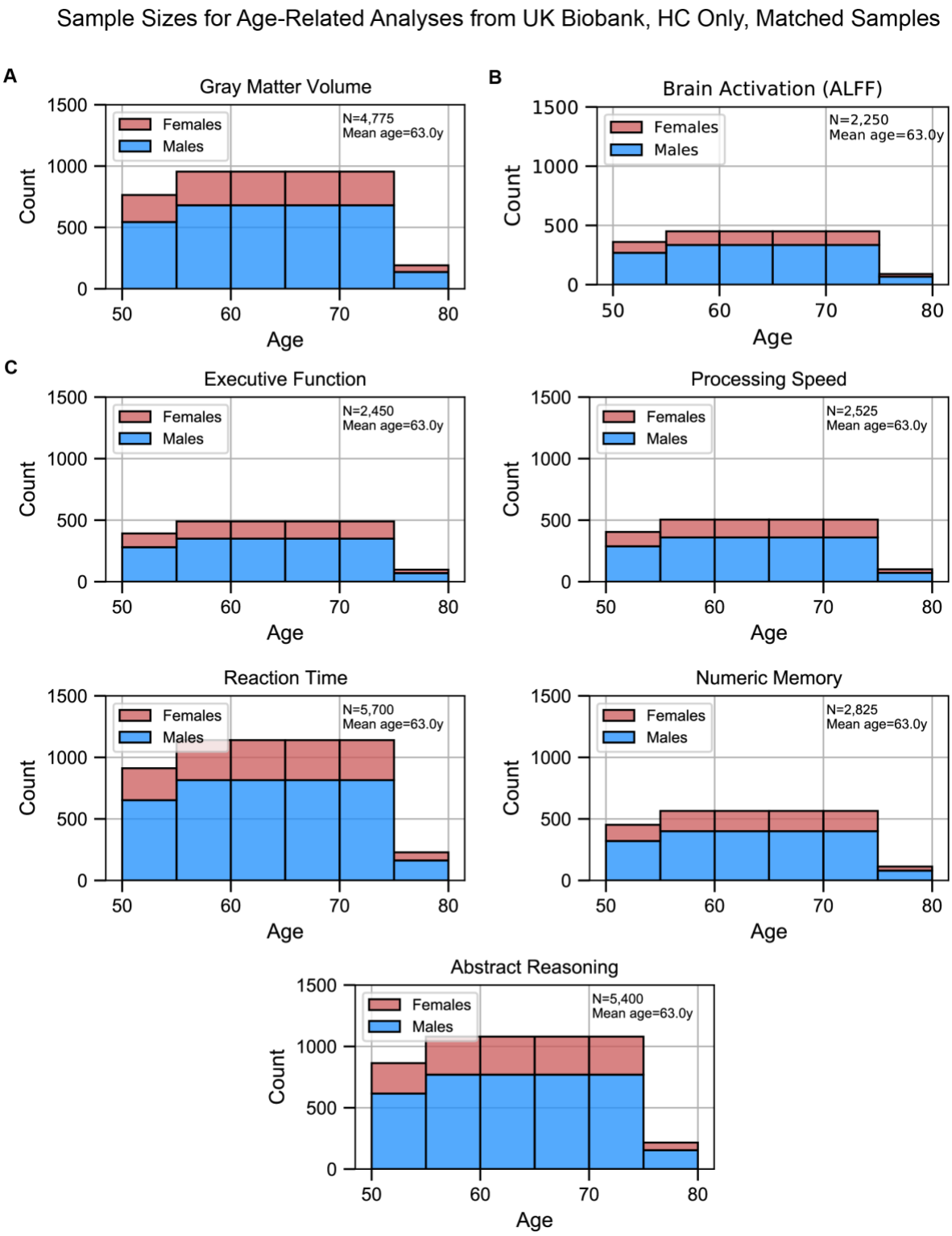
medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary Figures



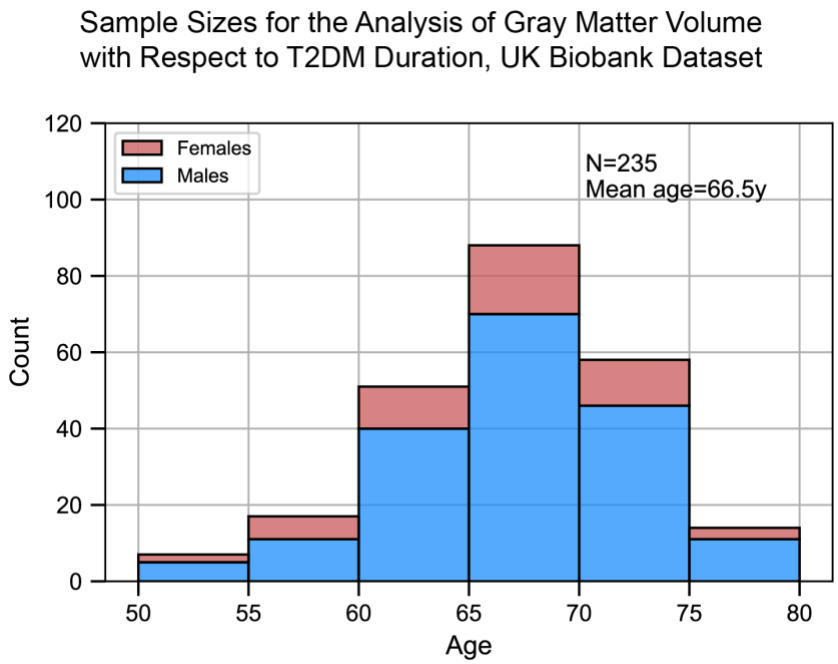
Supplementary Figure 1: Sample size distributions in T2DM vs HC analyses. The values shown represent T2DM+ only, T2DM+ were matched with an equivalent number of HCs, resulting in twice the number of samples in every analysis. The bars are stacked. **A:** Gray matter volume analyses **B:** Brain activation (ALFF) analyses **C:** Cognition analyses (5 domains).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



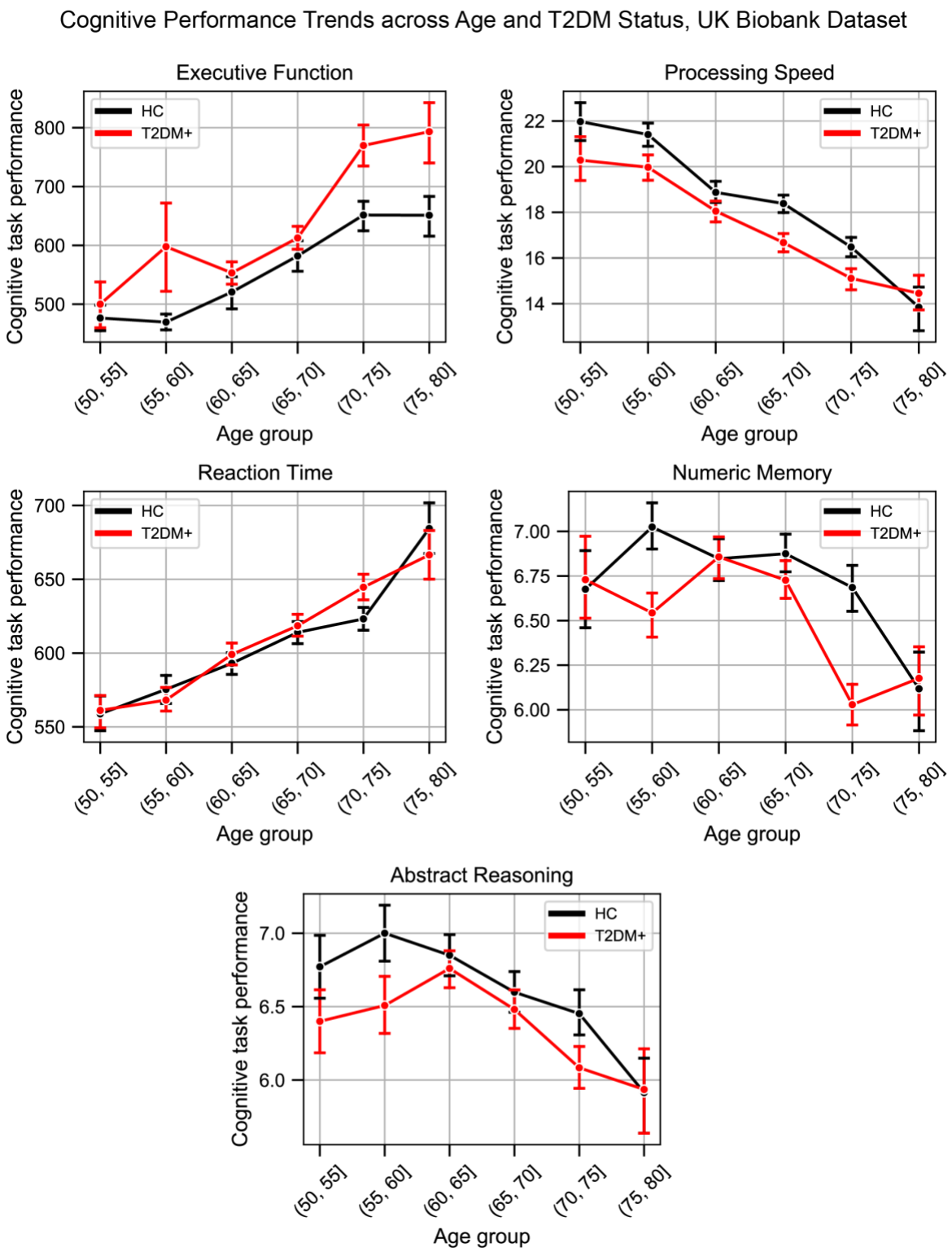
Supplementary Figure 2: Sample size distributions of the analyses which investigated Age as the factor of interest. Samples included HC only and were matched for sex, education, and hypertension across age. The matching method resulted in "sample chains" across age: within each chain, samples were equivalent in terms of sex, education, and hypertension. The bars are stacked. **A:** Gray matter volume analyses **B:** Brain activation (ALFF) analyses **C:** Cognition analyses (5 domains).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



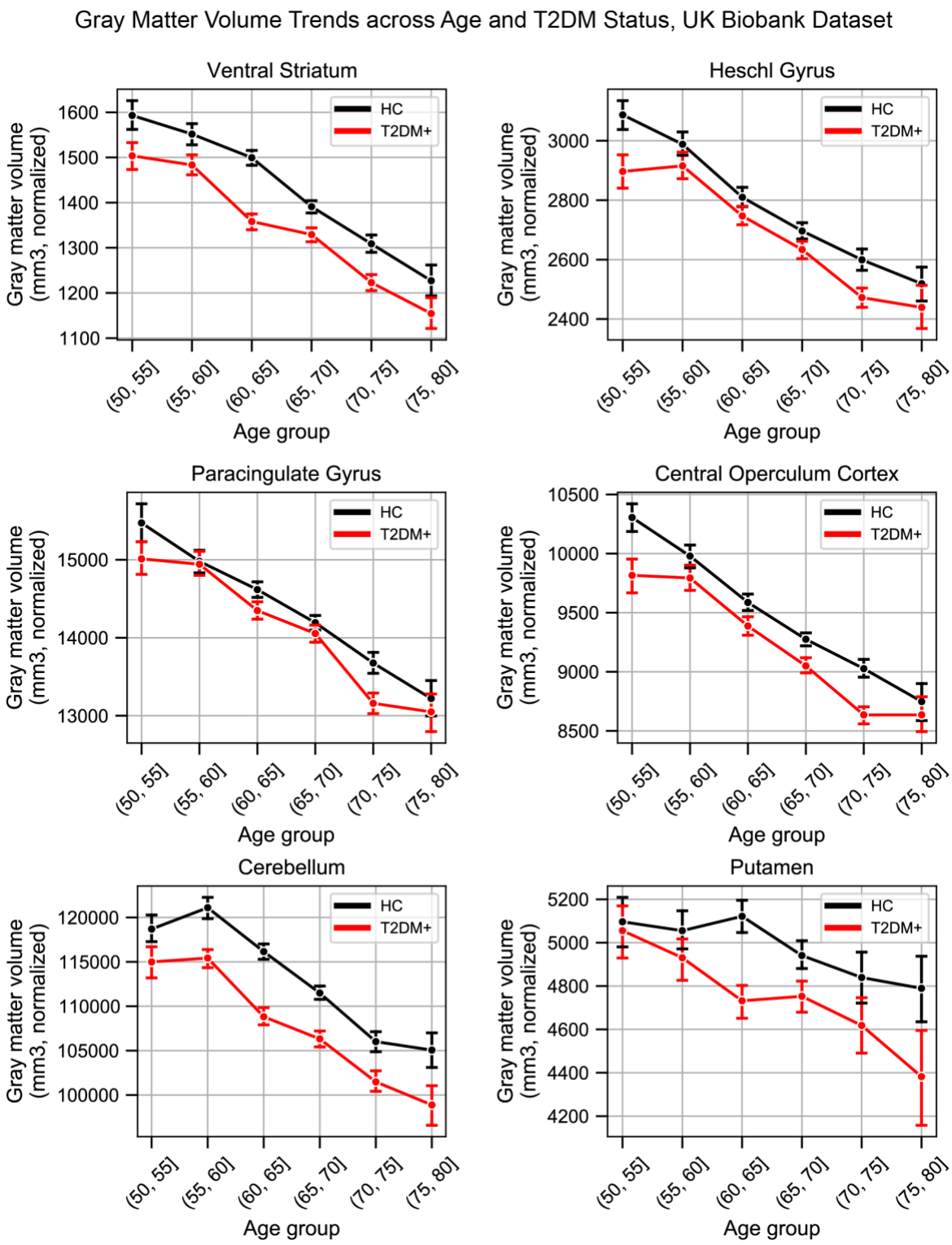
Supplementary Figure 3: Sample size distribution in the analysis which investigated T2DM duration as the main factor. The values shown represent one out of the three duration groups (HC, T2DM+ for 0-9y, and T2DM+ for 10+y). Samples across the three duration groups were matched for age, sex, education, and hypertension. Given the three groups, the total sample size was three times what is shown here. The statistical analysis included T2DM+ samples only (for which duration was defined), HC were included for visual representation only. The bars are stacked.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



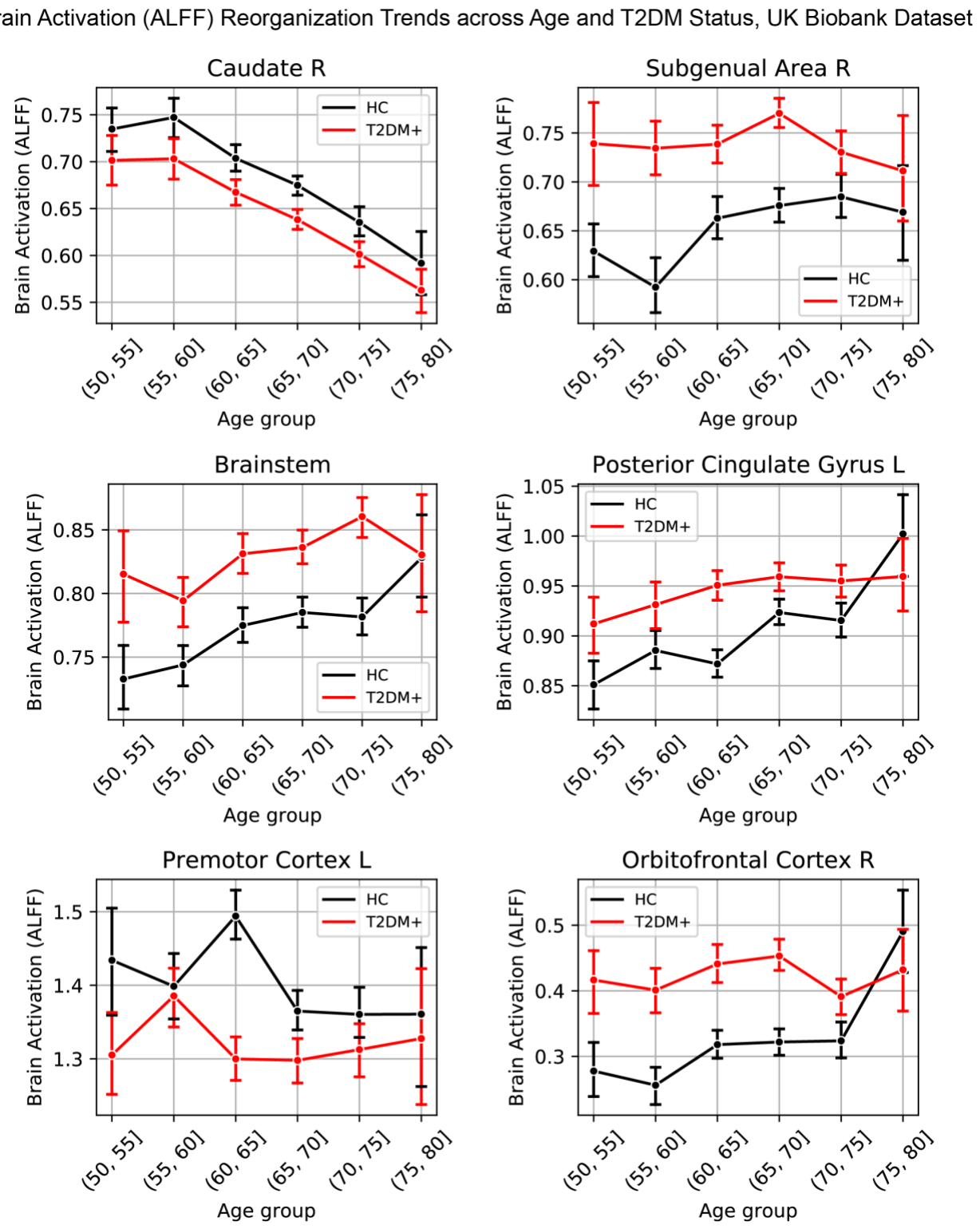
Supplementary Figure 4: Lineplots representing trends across Age and T2DM for each of the five investigated cognitive domains. For *executive function* and *reaction time*, lower scores represented better performance. The linear trends across age were robust and justified modeling age as a linear factor. The apparent deviation from linear trends in the youngest and oldest age groups are explained by drastically lower sample sizes in these age groups (see underlying sample sizes in **SI Figure 1C**).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



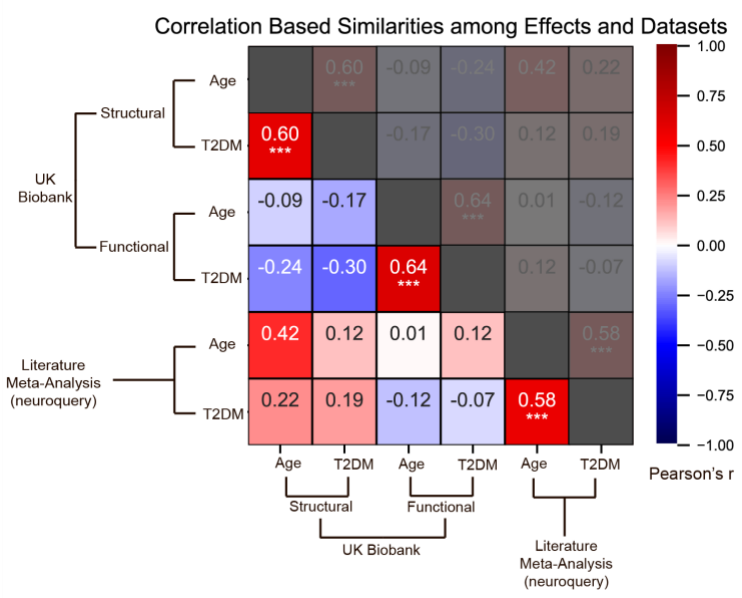
Supplementary Figure 5: Lineplots representing trends across Age and T2DM for five regions which exhibited the strongest trends in gray matter volume changes. The linear trends across age were robust and justified modeling age as a linear factor. The apparent deviation from linear trends in the youngest and oldest age groups are explained by markedly lower sample sizes in these age groups (see underlying sample sizes in **SI Figure 1A**).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



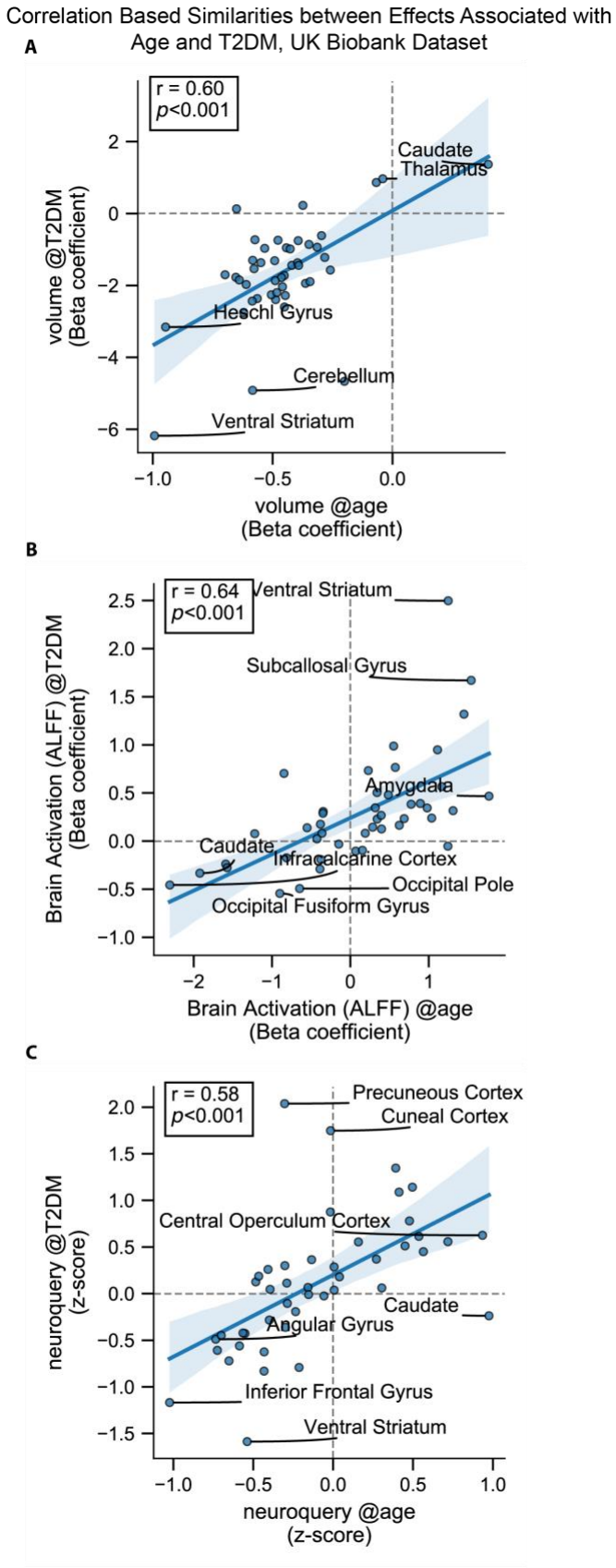
Supplementary Figure 6: Lineplots representing trends across Age and T2DM for five regions which exhibited the strongest trends in brain activation (ALFF). The linear trends across age were robust and justified modeling age as a linear factor. The apparent deviation from linear trends in the youngest and oldest age groups are explained by drastically lower sample sizes in these age groups (see underlying sample sizes in **SI Figure 1B**).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 7: Effects of age and T2DM exhibited strong correlations within datasets and modalities. We considered six sets of changes we previously characterized 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. See relevant scatterplots in **SI Figure 8**. $^{*}P \leq 0.05$; $^{**}P \leq 0.01$; $^{***}P \leq 0.001$, Bonferroni corrected.

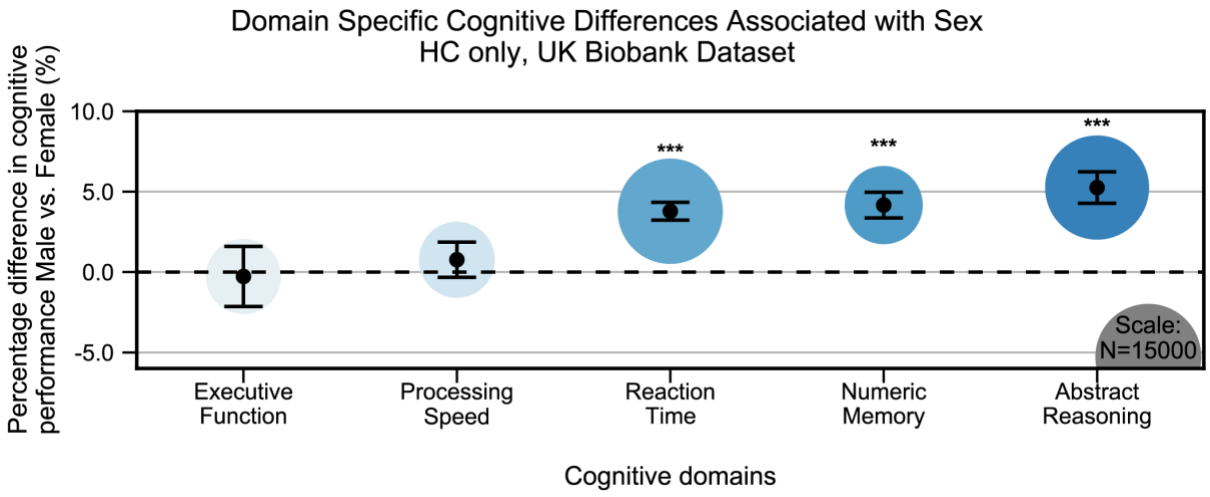
medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 8: Scatterplots corresponding to the statistically significant cells in **SI Figure 7**. Statistics shown are Pearson's correlations. Regions on the extremes are labeled. **A:** Gray matter volume results, **B:** Brain activation (ALFF) results, **C:** Neuroquery-based meta-analysis results.

481

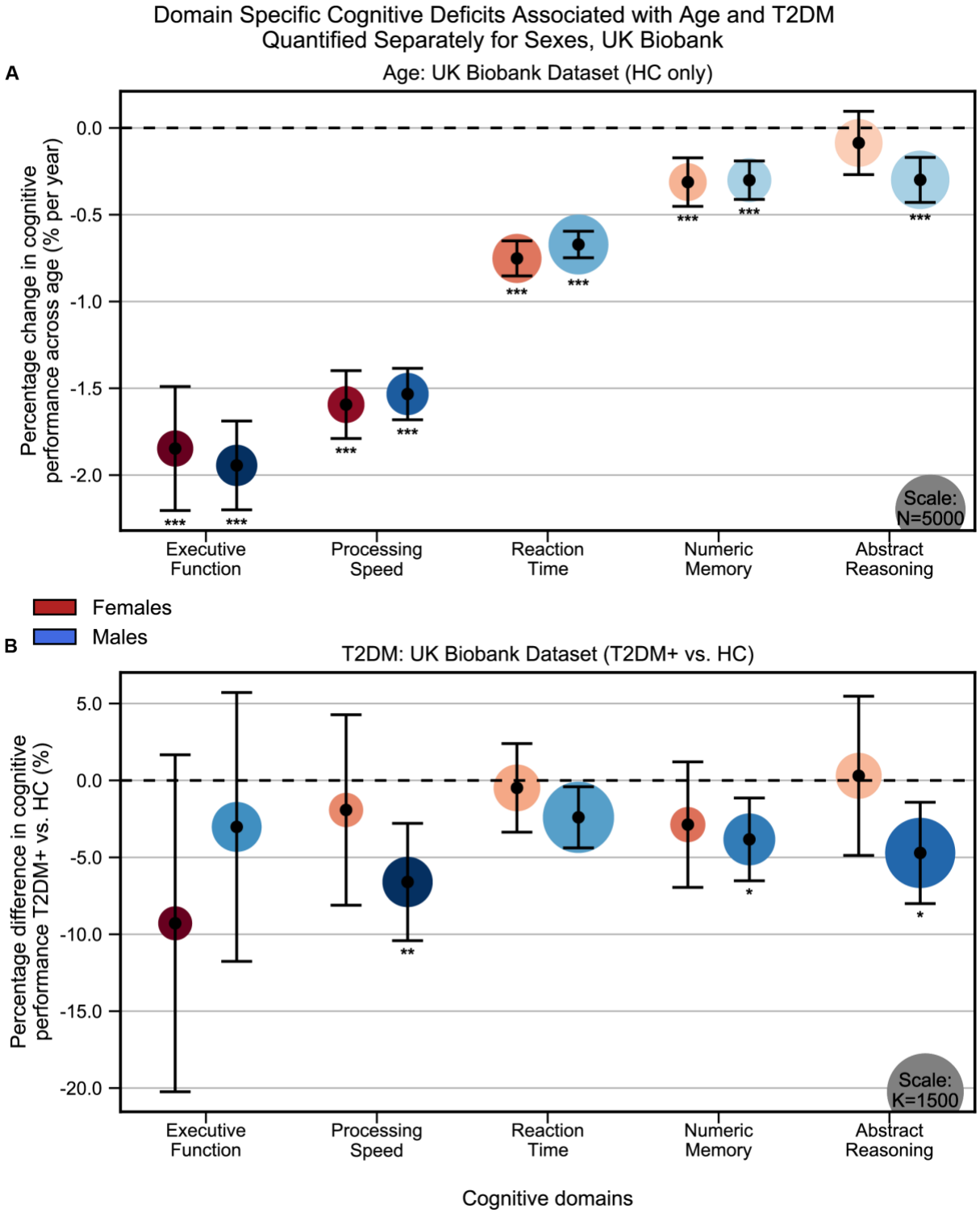
medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



482

Supplementary Figure 9: Cognitive differences associated with sex in the UK Biobank dataset across the five cognitive domains. Samples were HC only and were matched for age, education, and hypertension. Values on the y axis represent % difference in performance compared to the combined average (from both F and M). A positive % represents better performance in males. Marker sizes represent sample sizes scaled (per area) as indicated in the bottom right corner. Error bars are 95% CI. $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$, Bonferroni corrected.

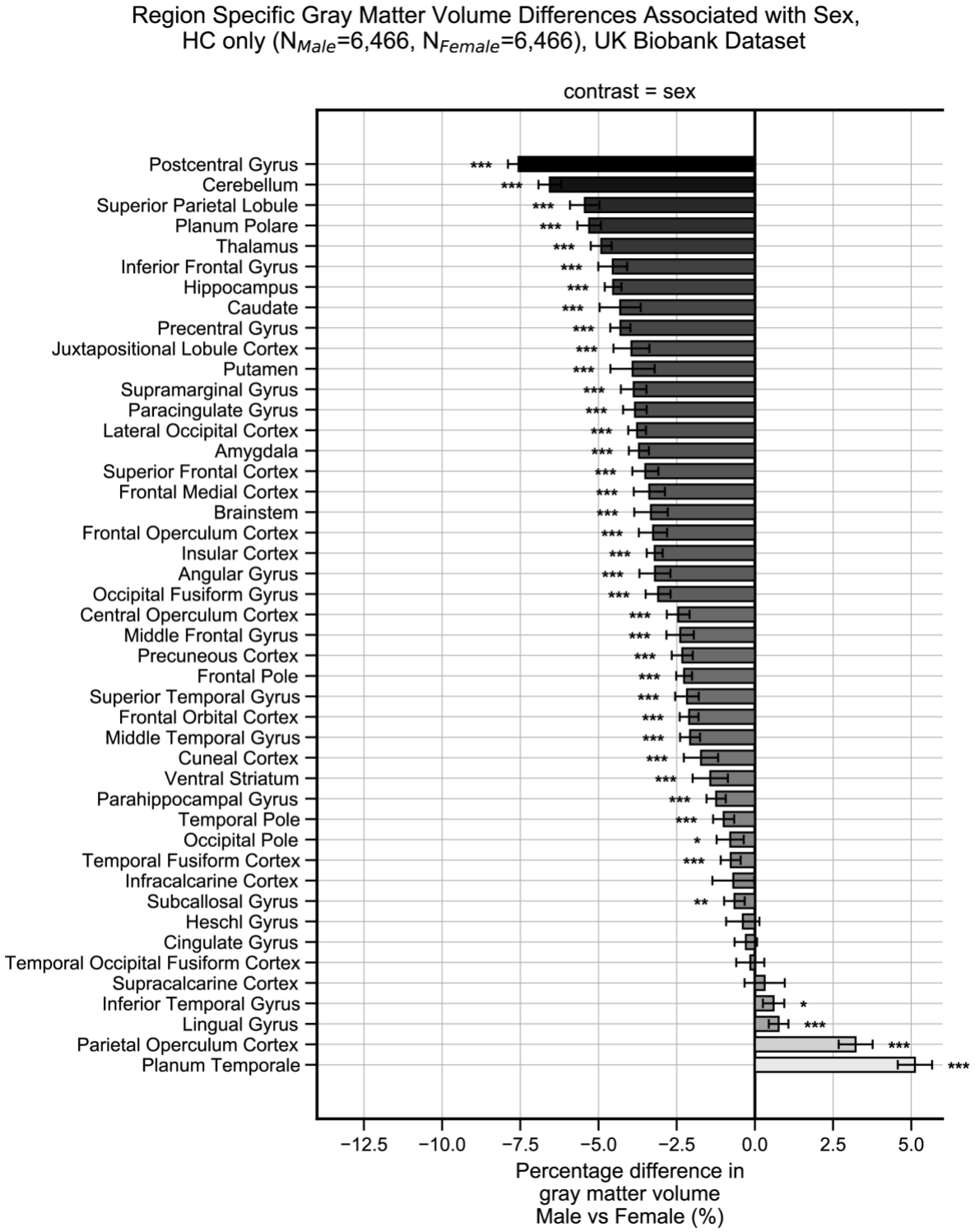
medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



483

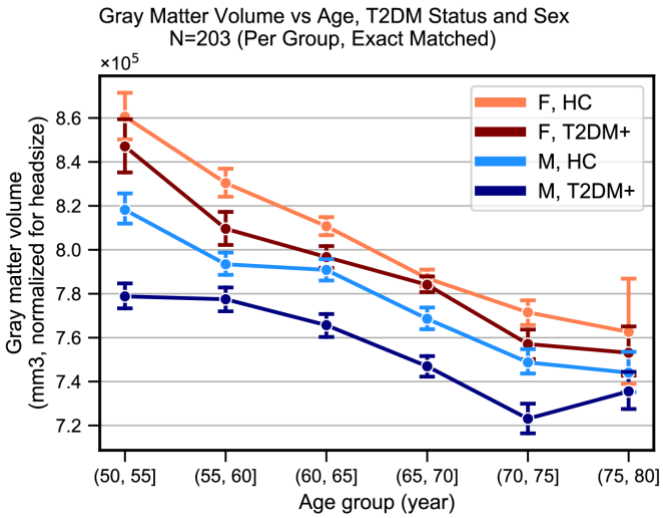
Supplementary Figure 10: Cognitive differences associated with Age (**A**) and T2DM (**B**) in the UK Biobank dataset across the five cognitive domains, analyzed separately within Males and Females. Samples were matched for age, education, and hypertension. The associated effects were consistent across males vs females, but with respect to T2DM, effects were generally stronger in males. Note that sample sizes were significantly larger for males compared to females. Marker sizes represent sample sizes scaled (per area) as indicated in the bottom right corner of each panel. Error bars are 95% CI. $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$, Bonferroni corrected.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 11: Gray matter volume (normalized for headsize) differences associated with sex in the UK Biobank dataset across the 45 anatomical regions. Samples were HC only and were matched for age, education, and hypertension. Values on the x axis represent % difference in volume compared to the combined average (from both F and M). A negative % represents larger volume in females. Error bars are 95% CI. $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$, Bonferroni corrected.

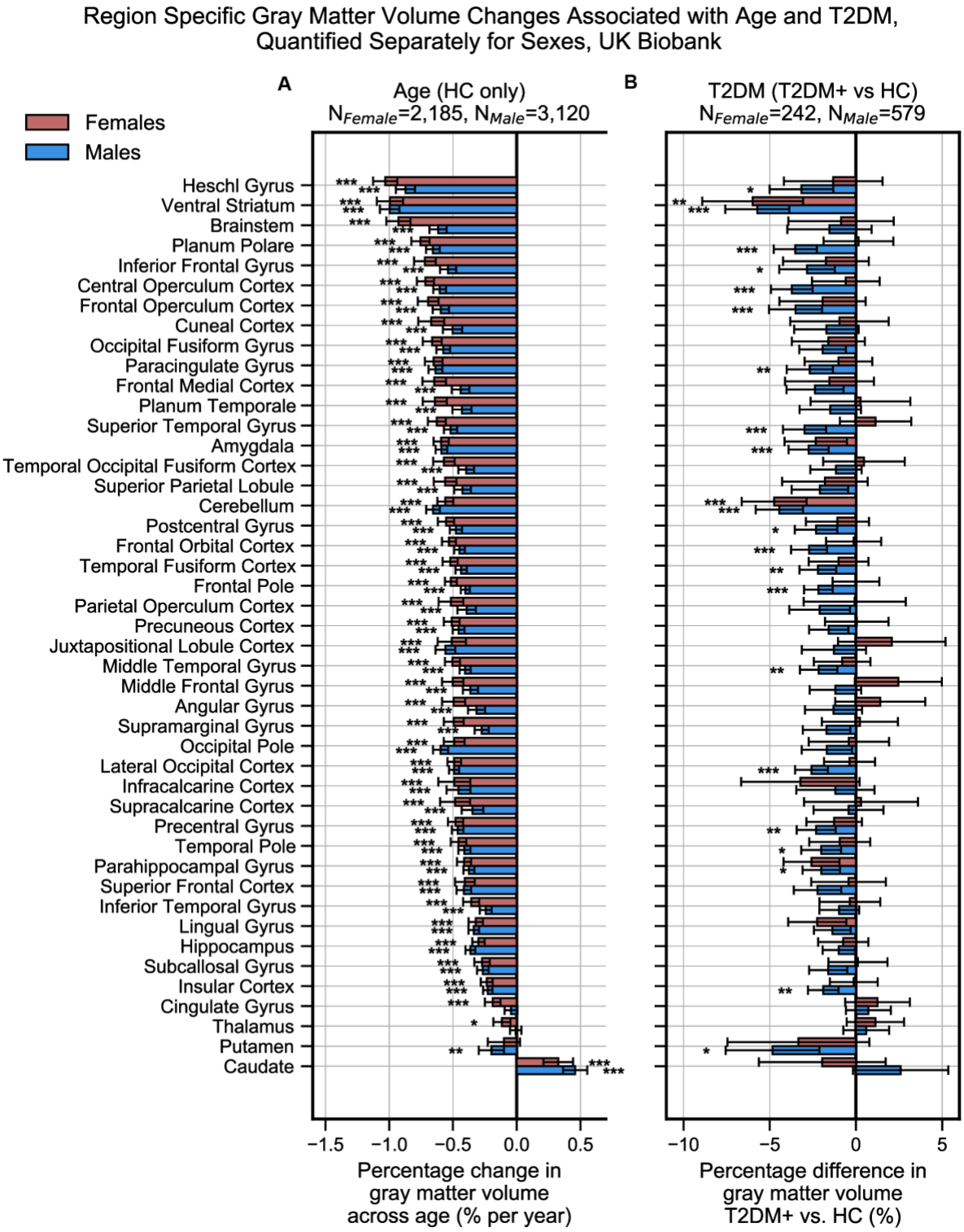
medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



485

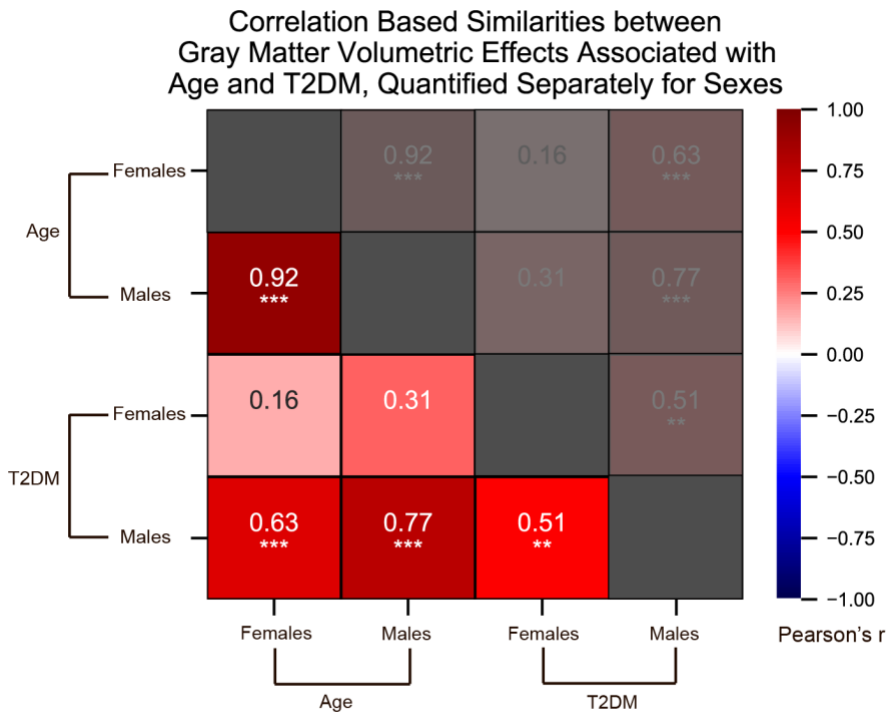
Supplementary Figure 12: Lineplots representing trends in total gray matter volume (normalized for headsize) across Age, Sex and T2DM. Samples were matched for age, education, and hypertension. The age-related decline appeared equivalent in the two sexes but T2DM associated deficits were stronger in males. With respect to disease duration (not shown), equivalent effects were observed magnitude-wise in males (24% ± 17%) and females (27% ± 32%), but these results were only significant in the male-only group (p=0.002), for which sample sizes were considerably larger (N_M=366, N_F=104).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 13: Region-specific gray matter volume deficits associated with Age (**A**) and T2DM (**B**) in the UK Biobank dataset, analyzed separately within females and males. Samples were matched for age, education, and hypertension. The associated effects were correlated across males vs females (see **SI Figures 14**, **15A**, and **15B**), With respect to T2DM, effects were generally stronger in males. Age and T2DM effects significantly correlated in males but not in females (see **SI Figures 14**, **15C**, and **15D**). Note that sample sizes were significantly larger for males compared to females. Error bars are 95% CI. $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$, Bonferroni corrected.

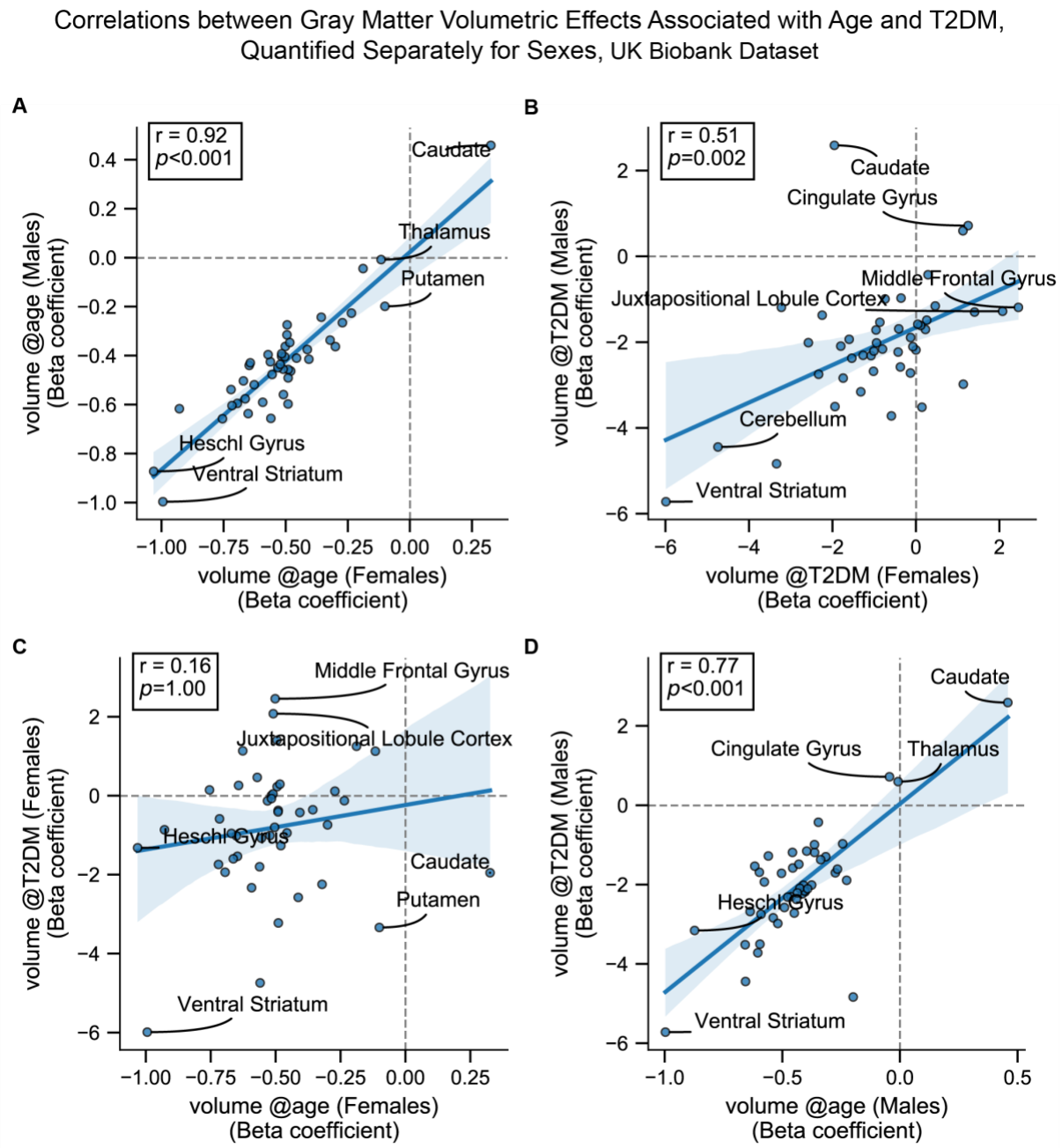
medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



487

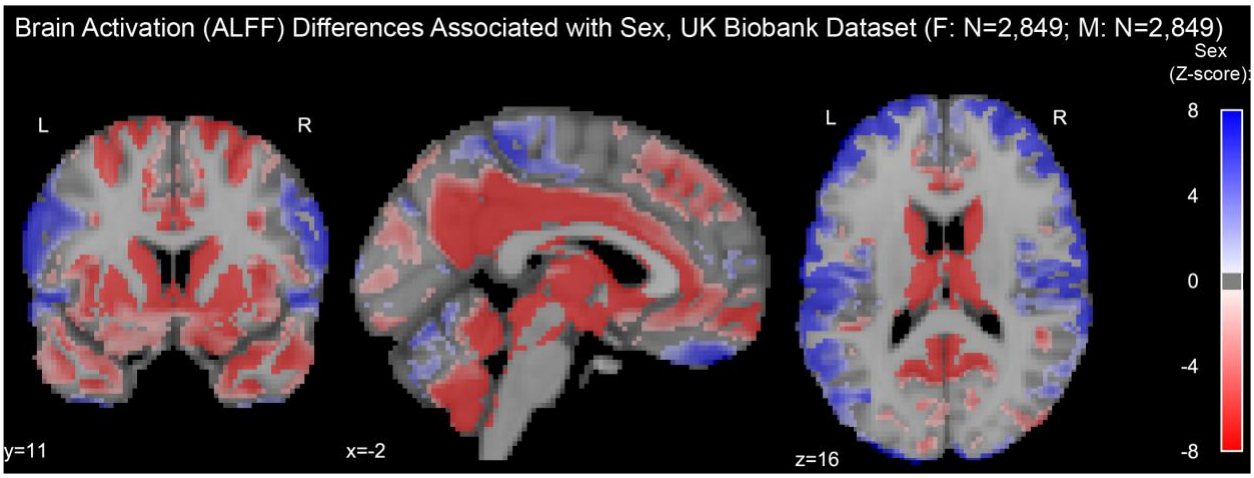
Supplementary Figure 14: We quantified correlations (Pearson's r) among the region-specific gray matter volume deficits associated with Age and T2DM, which we quantified in the UK Biobank dataset separately for females and males. See relevant scatterplots in **SI Figure 15.** $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$, Bonferroni corrected.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



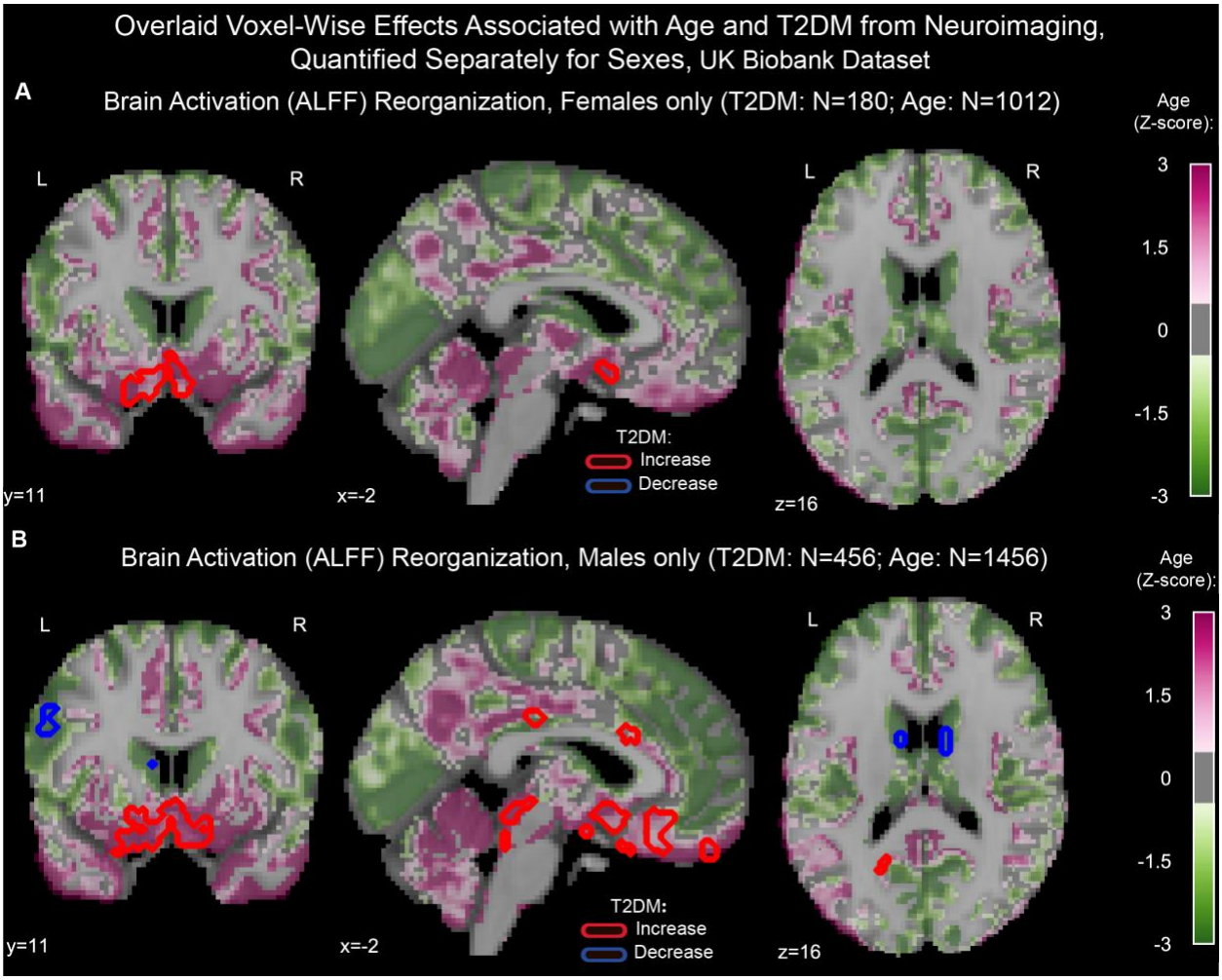
Supplementary Figure 15: Scatterplots corresponding to the most relevant cells in **SI Figure 14**. Statistics shown are Pearson's correlations. Regions on the extremes are labeled. All panels contain gray matter volume related effects only. **A:** Age-related effects in males vs females, **B:** T2DM-related effects in males vs females, **C:** Age vs T2DM effects in females only, **D:** Age vs T2DM effects in males only.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



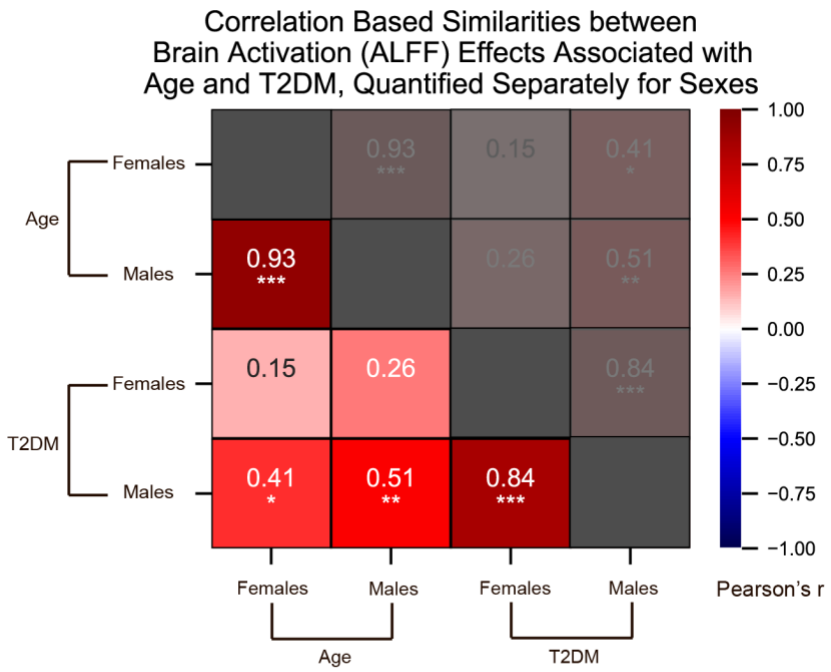
Supplementary Figure 16: Resting-state brain activation (ALFF) differences associated with sex in the UK Biobank dataset. Samples were HC only and were matched for age, education, and hypertension. Only significantly different voxels are shown following false discovery rate correction. Positive (labeled with blue) voxels represent stronger activation in males compared to females.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



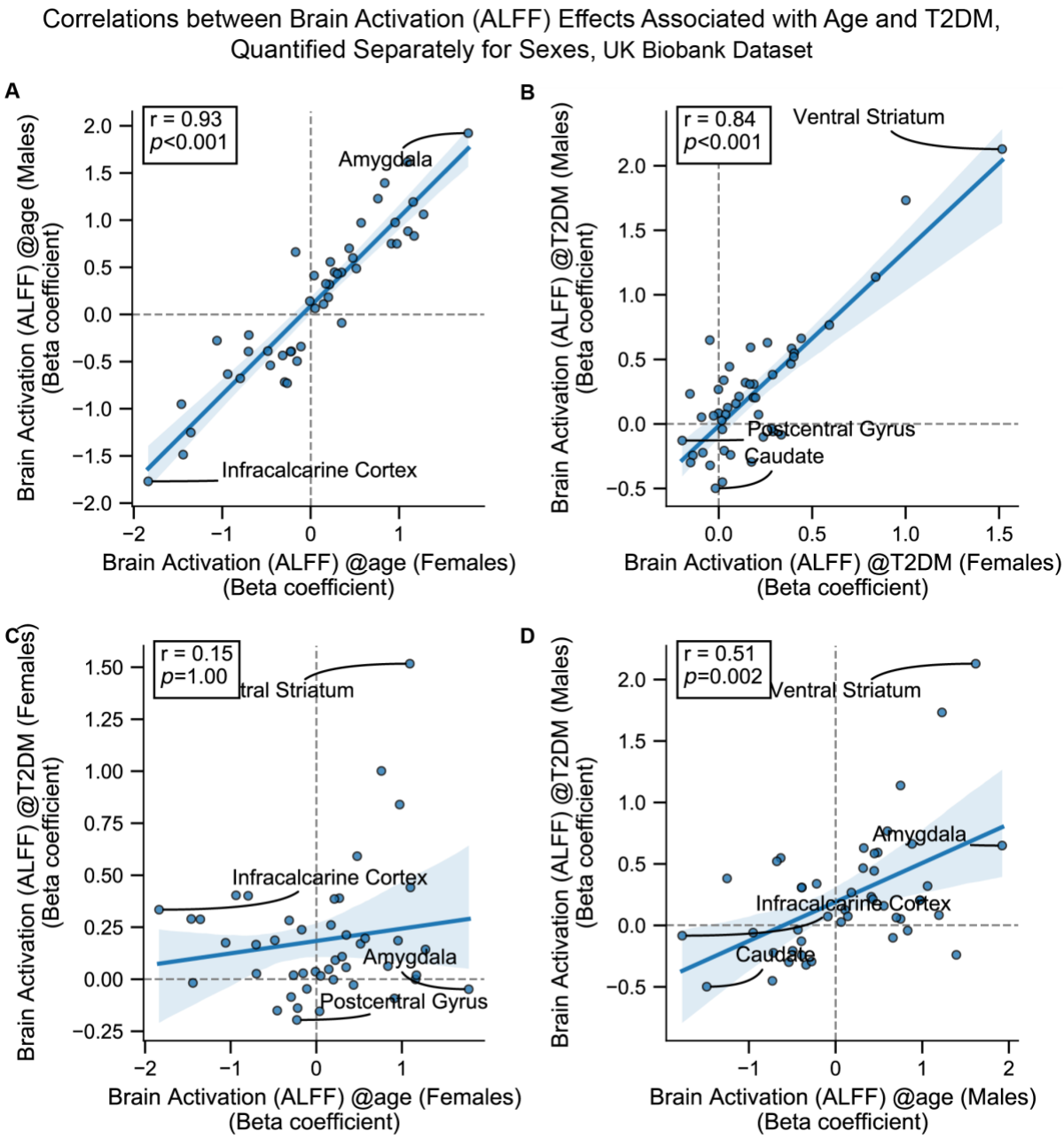
Supplementary Figure 17: Age and T2DM associated reorganization patterns in brain activation (ALFF) were overlaid separately for females (**A**) and males (**B**). Samples were matched for age, education, and hypertension. Effects linked to age are shown in the form of an unthresholded z-map represented 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. As also seen with gray matter volume, the associated effects were correlated across males vs females (see **SI Figures 18**, **19A**, and **19B**), With respect to T2DM, effects were generally stronger in males. Age and T2DM effects significantly correlated in males but not in females (see **SI Figures 18**, **19C**, and **19D**). Note that sample sizes were significantly larger for males compared to females.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



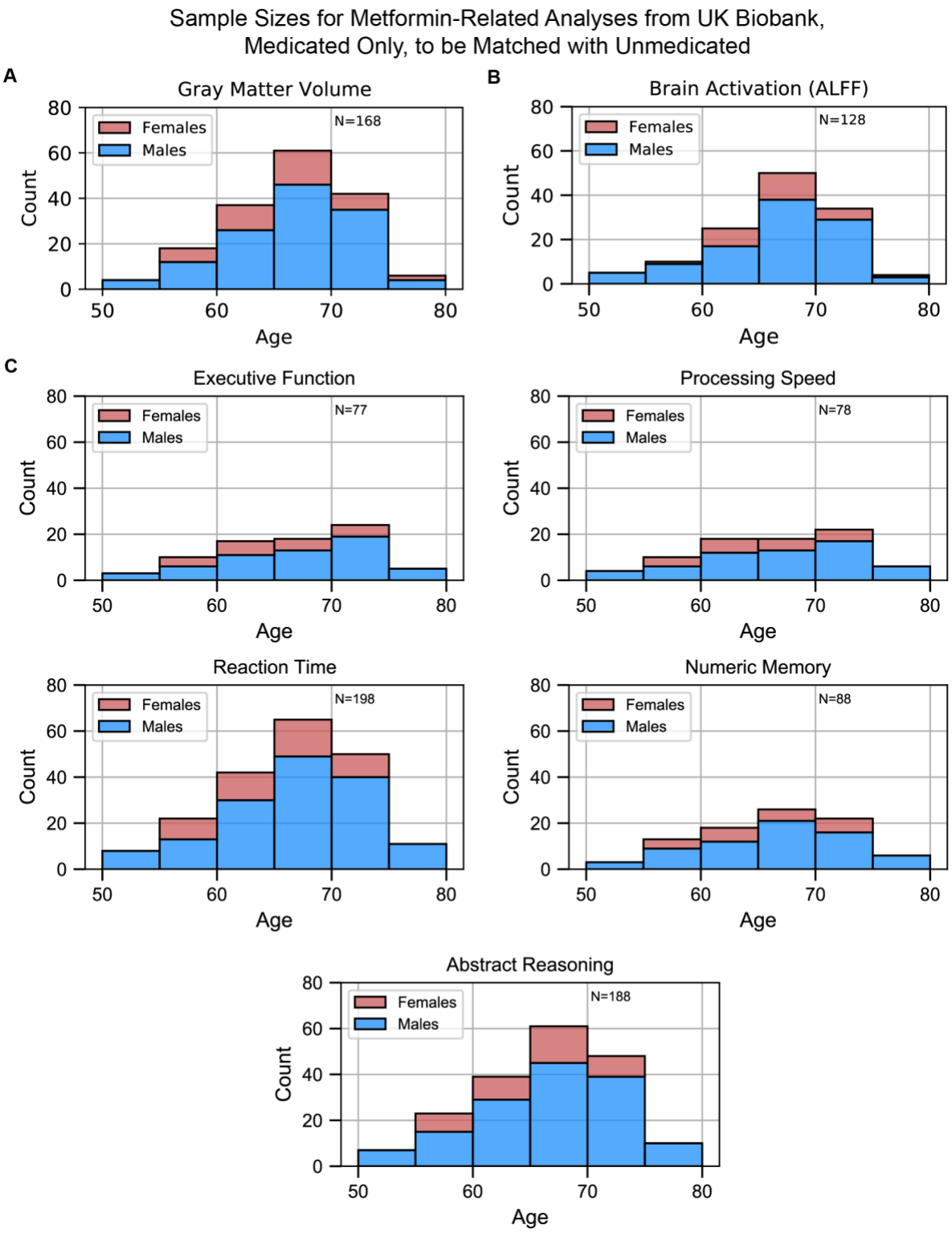
Supplementary Figure 18: We quantified correlations (Pearson's r) among the region-specific changes in brain activation (ALFF) patterns associated with Age and T2DM, which we quantified in the UK Biobank dataset separately for females and males. See relevant scatter plots in **SI Figure 19.** $*P \leq 0.05$; $**P \leq 0.01$; $***P \leq 0.001$, Bonferroni corrected.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



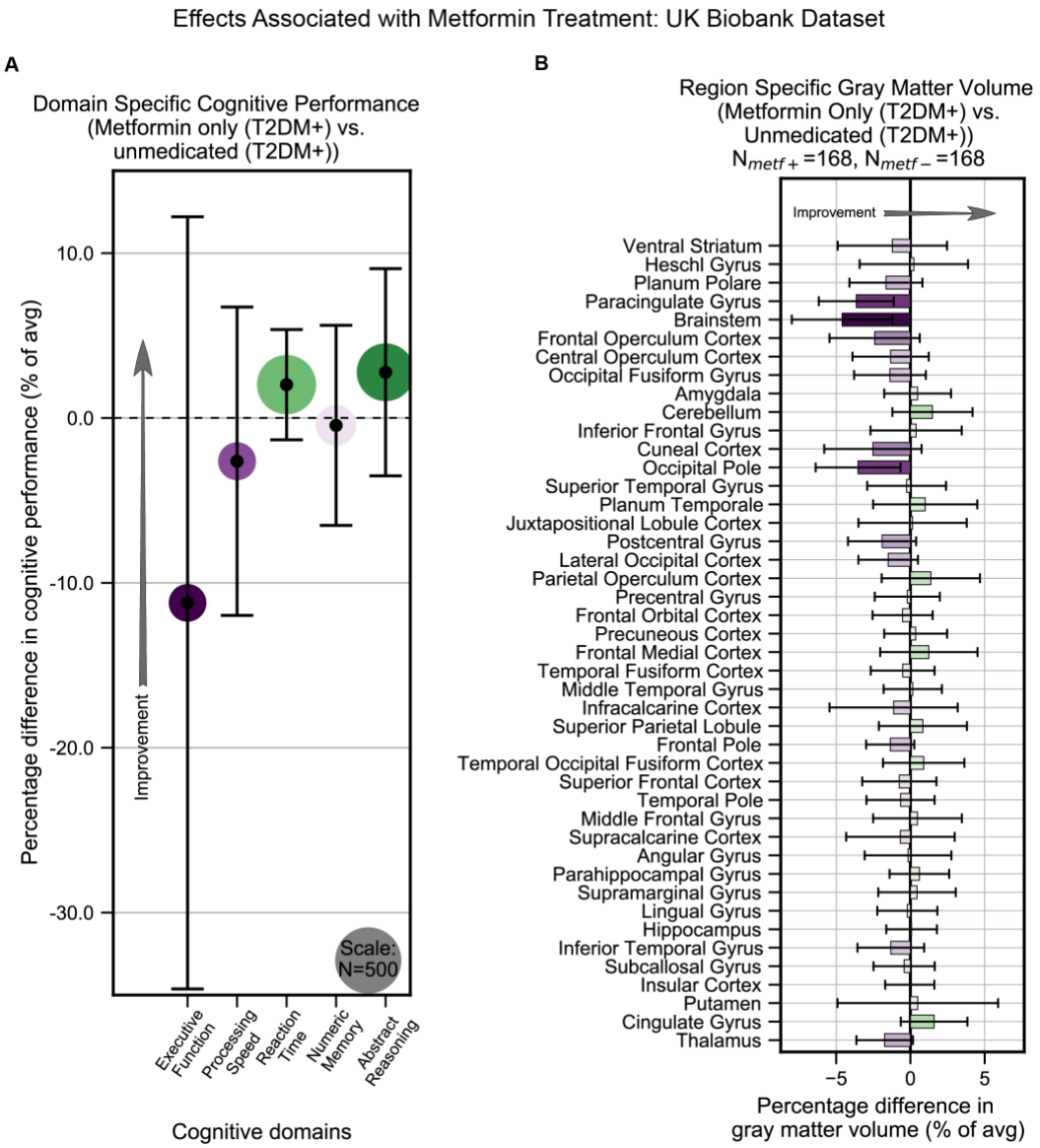
Supplementary Figure 19: Scatterplots corresponding to the primary cells in **SI Figure 18**. Statistics shown are Pearson's correlations. Regions on the extremes are labeled. All panels contain brain activation (ALFF) effects only. **A:** Age-related effects in males vs females, **B:** T2DM-related effects in males vs females, **C:** Age vs T2DM effects in females only, **D:** Age vs T2DM effects in males only.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 20: Sample size distributions in metformin medication status related analyses. The values shown represent metformin medicated T2DM+ only, these samples were matched with an equivalent number of unmedicated T2DM+ samples, resulting in twice the amount of samples in every analysis. Samples were matched for age, sex, education, and hypertension. The bars are stacked. **A:** Gray matter volume analyses **B:** Brain activation analyses **C:** Cognition analyses (5 domains).

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .



Supplementary Figure 21: 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 arrows. **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. Underlying sample size distributions can be found in **SI Figure 20**. Error bars are 95% CI.

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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
Blood Pressure	93, 94, 4079, 4080	2
Had Menopause	2724	0-2
Had hormone replacement therapy	3546	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 (Abstract Reasoning)	20016	2
Alphanumeric Trail-Making Test (Executive Function)	6350	2
Symbol-Digit Substitution (Processing Speed)	23324	2
Snap Game (Reaction Time)	20023	0, 2
Numeric Memory Test (Numeric Memory)	4282	2

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

Domains	Common Tests	Description
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.
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.
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.
Numeric Memory	(Forwards) Digit Span, Digit Vigilance Test	Short-term (~2-3 s) recall of numeric information.
Visual Memory	Location Learning, Weschler Visual Memory Subtest, Rey-Osterreith Delayed Recall, Face Recognition Test	Short and long-term recall of visually encoded information.
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.
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").
Visuospatial Reasoning	Rey-Osterreith Figure Copy, Taylor Complex Figure, Weschler Object Assembly	Manipulation or reconstruction of spatial 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.

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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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
Immediate 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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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%)**

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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
Numeric 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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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*

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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. doi:10.2337/db18-202-LB
2. Moheet A, Mangia S, Seaquist ER. Impact of diabetes on cognitive function and brain structure. *Ann N Y Acad Sci*. Sep 2015;1353:60-71. doi:10.1111/nyas.12807
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*. 2019/03/01 2019;62(3):448-458. doi:10.1007/s00125-018-4778-9
4. Gold SM, Dziobek I, Sweat V, et al. Hippocampal damage and memory impairments as possible early brain complications of type 2 diabetes. *Diabetologia*. Apr 2007;50(4):711-9. doi:10.1007/s00125-007-0602-7
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*. 09/01 2018;35doi:10.1111/dme.13800
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*. Apr 2006;55(4):1106-13. doi:10.2337/diabetes.55.04.06.db05-1323
7. Rawlings AM, Sharrett AR, Schneider AL, et al. Diabetes in midlife and cognitive change over 20 years: a cohort study. *Ann Intern Med*. Dec 2 2014;161(11):785-93. doi:10.7326/m14-0737
8. Zilliox LA, Chadrasekaran K, Kwan JY, Russell JW. Diabetes and Cognitive Impairment. *Current diabetes reports*. Sep 2016;16(9):87. doi:10.1007/s11892-016-0775-x
9. Pelimanni E, Jehkonen M. Type 2 Diabetes and Cognitive Functions in Middle Age: A Meta-Analysis. *J Int Neuropsychol Soc*. Dec 21 2018:1-16. doi:10.1017/s1355617718001042
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. doi:doi:10.1016/S1474-4422(05)70284-2
11. Ninomiya T. Diabetes mellitus and dementia. *Curr Diab Rep*. 2014;14(5):487. doi:10.1007/s11892-014-0487-z
12. Tuligenga RH. Intensive glycaemic control and cognitive decline in patients with type 2 diabetes: a meta-analysis. *Endocr Connect*. Jun 2015;4(2):R16-24. doi:10.1530/EC-15-0004
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*. Mar 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779
14. Chobanian AV, Bakris GL, Black HR, et al. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *hypertension*. 2003;42(6):1206-1252.
15. Maki PM, Henderson VW. Cognition and the menopause transition. *Menopause*. 2016;23(7):803-805.
16. Thomas NJ, Jones SE, Weedon MN, Shields BM, Oram RA, Hattersley AT. Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *The Lancet Diabetes & Endocrinology*. 2018;6(2):122-129.
17. Seabold S, Perktold J. Statsmodels: Econometric and statistical modeling with python. 2010:92-96.
18. 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*. Feb 1 2018;166:400-424. doi:10.1016/j.neuroimage.2017.10.034
19. Zang YF, He Y, Zhu CZ, et al. Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain & development*. Mar 2007;29(2):83-91. doi:10.1016/j.braindev.2006.07.002

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

20. Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. Jun 1996;29(3):162-73. doi:10.1006/cbmr.1996.0014

21. Cox RW, Hyde JS. Software tools for analysis and visualization of fMRI data. *NMR in biomedicine*. Jun-Aug 1997;10(4-5):171-8. doi:10.1002/(sici)1099-1492(199706/08)10:4/5<171::aid-nbm453>3.0.co;2-l

22. Fieller EC. Some problems in interval estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1954;16(2):175-185.

23. 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*. Jun 1 2017;1(6):600-608. doi:10.1210/js.2017-00046

24. 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*. Mar 2016;30(2):212-20. doi:10.1016/j.jdiacomp.2015.11.016

25. Page MJ, Moher D, Bossuyt PM, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *bmj*. 2021;372

26. Team RC. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2018;

27. Viechtbauer W. Conducting Meta-Analyses in R with The metafor Package. *Journal of Statistical Software*. 08/01 2010;36doi:10.18637/jss.v036.i03

28. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. L. Erlbaum Associates; 1988:xxi, 567 p.

29. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj*. Sep 6 2003;327(7414):557-60. doi:10.1136/bmj.327.7414.557

30. Dockès Jm, Poldrack RA, Primet R, et al. NeuroQuery, comprehensive meta-analysis of human brain mapping. *Elife*. 2020;9:e53385.

31. 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.

32. Van Horn JD, Grafton ST, Rockmore D, Gazzaniga MS. Sharing neuroimaging studies of human cognition. *Nature neuroscience*. 2004;7(5):473-481.

33. 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.

34. 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.

35. Sanchis-Segura C, Ibañez-Gual MV, Adrián-Ventura J, et al. Sex differences in gray matter volume: how many and how large are they really? *Biology of sex Differences*. 2019;10(1):1-19.

36. Jain A, Bansal R, Kumar A, Singh K. A comparative study of visual and auditory reaction times on the basis of gender and physical activity levels of medical first year students. *International Journal of Applied and Basic Medical Research*. 2015;5(2):124.

37. Behl C. Oestrogen as a neuroprotective hormone. *Nature reviews Neuroscience*. Jun 2002;3(6):433-42. doi:10.1038/nrn846

38. 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*. Aug 13 2013;110(33):13642-7. doi:10.1073/pnas.1303346110

39. Weistuch C, Mujica-Parodi LR, Amgalan A, Dill KA. Metabolism modulates global synchrony in the aging brain. *bioRxiv*. 2020;

40. Tomasi D, Volkow ND. Aging and functional brain networks. Research Support, N.I.H., Extramural. *Mol Psychiatry*. May 2012;17(5):471, 549-58. doi:10.1038/mp.2011.81

41. 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*. Mar 17 2020;117(11):6170-6177. doi:10.1073/pnas.1913042117

medRxiv preprint doi: https://doi.org/10.1101/2021.05.23.21257682; this version posted March 24, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

42. 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. *Journal of chemical neuroanatomy*. Nov 2002;24(4):225-42. doi:10.1016/s0891-0618(02)00058-3

43. Kobayashi M, Nikami H, Morimatsu M, Saito M. Expression and localization of insulin-regulatable glucose transporter (GLUT4) in rat brain. *Neuroscience letters*. Aug 2 1996;213(2):103-6. doi:10.1016/0304-3940(96)12845-7

44. 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. *The Journal of comparative neurology*. Oct 5 1998;399(4):492-512. doi:10.1002/(sici)1096-9861(19981005)399:4<492::aid-cne4>3.0.co;2-x

45. Vannucci SJ, Koehler-Stec EM, Li K, Reynolds TH, Clark R, Simpson IA. GLUT4 glucose transporter expression in rodent brain: effect of diabetes. *Brain research*. Jun 22 1998;797(1):1-11. doi:10.1016/s0006-8993(98)00103-6

46. 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*. Jul 1995;352(1):67-73. doi:10.1007/BF00169191

47. Kurotani S, Umegaki H, Ishiwata K, Suzuki Y, Iguchi A. The age-associated changes of dopamine-acetylcholine interaction in the striatum. *Exp Gerontol*. Sep 2003;38(9):1009-13. doi:10.1016/s0531-5565(03)00162-1

48. Ferreira-Vieira TH, Guimaraes IM, Silva FR, Ribeiro FM. Alzheimer's disease: Targeting the Cholinergic System. *Curr Neuropharmacol*. 2016;14(1):101-15. doi:10.2174/1570159x13666150716165726

49. Muramatsu I, Uwada J, Yoshiki H, et al. Novel regulatory systems for acetylcholine release in rat striatum and anti-Alzheimer's disease drugs. *Journal of neurochemistry*. Jun 2019;149(5):605-623. doi:10.1111/jnc.14701

50. Tessari P, Cecchet D, Cosma A, et al. Nitric oxide synthesis is reduced in subjects with type 2 diabetes and nephropathy. *Diabetes*. Sep 2010;59(9):2152-9. doi:10.2337/db09-1772

51. 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.

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