



HHS Public Access

Author manuscript

Lancet Healthy Longev. Author manuscript; available in PMC 2024 October 08.

Published in final edited form as:

Lancet Healthy Longev. 2024 September ; 5(9): 100623. doi:10.1016/j.lanhl.2024.07.006.

Association between surgical admissions, cognition, and neurodegeneration in older people: a population-based study from the UK Biobank

Jennifer Taylor,
Kristy P Robledo,
Vicente Medel,
Gillian Heller,
Thomas Payne,
Jordan Wehrman,
Cameron Casey,
Phillip F Yang,
Bryan M Krause,
Richard Lennertz,
Sharon Naismith,
Armando Teixeira-Pinto,
Robert D Sanders

Department of Anaesthetics, Royal Prince Alfred Hospital, Sydney Local Health District, Sydney, NSW, Australia (J Taylor PhD, T Payne MD, J Wehrman PhD, Prof R D Sanders MBBS PhD); NHMRC Clinical Trials Centre, The University of Sydney, Sydney, NSW, Australia (J Taylor, K P Robledo PhD, Prof G Heller PhD, T Payne, J Wehrman, Prof R D Sanders); Central Clinical School (J Taylor, T Payne, J Wehrman, Prof R D Sanders), Surgical Outcomes Research Centre (SOuRCe) (P F Yang MBBS MS), Healthy Brain Ageing Program, Brain and Mind Centre (Prof S Naismith DPsych), and School of Public Health (Prof A Teixeira-Pinto PhD), Faculty of Medicine and Health, The University of Sydney, Sydney, NSW, Australia; School of Psychology, Faculty of Science, The University of Sydney, Sydney, NSW, Australia (Prof S Naismith); Latin American Brain Health Institute (BrainLat), Universidad Adolfo Ibáñez, Santiago, Chile (V Medel PhD); Department of Anesthesiology, School of Medicine and Public Health, University of Wisconsin, Madison, WI, USA (C Casey PhD, B M Krause PhD, R Lennertz MD PhD); South West Sydney Clinical Campus, Faculty of Medicine and Health, University of New South Wales, Sydney, NSW, Australia (P F Yang)

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Correspondence to: Dr Jennifer Taylor, Central Clinical School, Faculty of Medicine and Health, The University of Sydney, Sydney, NSW 2006, Australia, jennifer.taylor1@sydney.edu.au.

Contributors

RDS and JT planned and submitted the original UK Biobank application. RDS, JT, SN, and RL contributed to study design and resourcing. RDS and PFY classified surgeries for complexity. RDS and JT verified the data provided by UK Biobank. TP assisted with code review. JT performed statistical analyses supervised by RDS, AT-P, BMK, KPR, and GH. Additional imaging data analyses were conducted by JW, VM, and CC. All authors contributed to manuscript revision. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

See **Online** for appendix

Summary

Background—Previous studies have shown that major surgical and medical hospital admissions are associated with cognitive decline in older people (aged 40–69 years at recruitment), which is concerning for patients and caregivers. We aimed to validate these findings in a large cohort and investigate associations with neurodegeneration using MRI.

Methods—For this population-based study, we analysed data from the UK Biobank collected from March 13, 2006, to July 16, 2023, linked to the National Health Service Hospital Episode Statistics database, excluding participants with dementia diagnoses. We constructed fully adjusted models that included age, time, sex, *Lancet* Commission dementia risk factors, stroke, and hospital admissions with a participant random effect. Primary outcomes were hippocampal volume and white matter hyperintensities, both of which are established markers of neurodegeneration, and exploratory analyses investigated the cortical thickness of Desikan–Killiany–Tourville atlas regions. The main cognitive outcomes were reaction time, fluid intelligence, and prospective and numeric memory. Surgeries were calculated cumulatively starting from 8 years before the baseline evaluation.

Findings—Of 502 412 participants in the UK Biobank study, 492 802 participants were eligible for inclusion in this study, of whom 46 706 underwent MRI. Small adverse associations with cognition were found per surgery: reaction time increased by 0.273 ms, fluid intelligence score decreased by 0.057 correct responses, prospective memory (scored as correct at first attempt) decreased (odds ratio 0.96 [95% CI 0.95 to 0.97]), and numeric memory maximum correct matches decreased by 0.025 in fully adjusted models. Surgeries were associated with smaller hippocampal volume ($\beta = -5.76 \text{ mm}^3$ [-7.89 to -3.64]) and greater white matter hyperintensities volume ($\beta = 100.02 \text{ mm}^3$ [66.17 to 133.87]) in fully adjusted models. Surgeries were also associated with neurodegeneration of the insula and superior temporal cortex.

Interpretation—This population-based study corroborates that surgeries are generally safe but cumulatively are associated with cognitive decline and neurodegeneration. Perioperative brain health should be prioritised for older and vulnerable patients, particularly those who have multiple surgical procedures.

Funding—The Australian and New Zealand College of Anaesthetists (ANZCA) Foundation and the University of Sydney.

Introduction

Patient outcomes after surgery and anaesthesia have been greatly improved by evidence-based advances in clinical care, leading to an expansion of treatments offered across the age spectrum. In the UK, the proportion of people aged 75 years or older undergoing surgery has increased from 14.9% in 1999 to 22.9% in 2015.¹ However, increased cumulative surgical exposure could have unintended consequences. Patients, families, and health-care providers have valid concerns about the secondary, adverse effects of anaesthesia and surgery on short-term and long-term cognition, particularly for older people. In our previous survey,² most respondents expressed a fear of permanent cognitive deficits after surgery.

The negative cognitive effects of surgery and anaesthesia are supported by a moderate level of evidence from animal models³ and within the surgical,^{4,5} anaesthetic,⁶ and geriatric literature, although some researchers contend that these associations are subtle, if they exist at all.^{7,8} We previously analysed data from 7532 older adults (aged 35–55 years at recruitment) from the UK Whitehall II cohort study, in whom up to five cognitive assessments had been conducted over 19 years.^{9,10} We found that major surgery was associated with additional cognitive decline equivalent to 5 months of ageing on average, after adjusting for age and other health determinants such as medical admissions, which had a larger association with changes in cognition than surgery. We also showed that substantial cognitive decline occurred in 12.7% of individuals after major hospital admissions and 5.5% after surgical admissions, compared with 2.5% of those without major admissions. The aim of this study is to validate and extend these findings using the UK Biobank—a large-scale, prospective, longitudinal cohort study of people aged 40–69 years at baseline with extensive biological phenotyping, cognitive assessment, imaging, and medical record linkage. Now a large-scale biomedical database, the UK Biobank aims to improve understanding, diagnosis, prevention, and treatment of serious and life-threatening illnesses.¹¹

Small studies have also shown that surgery is associated with short-term increases in neuronal injury biomarkers^{12–15} and markers of neurodegeneration¹⁶ as well as covert strokes.^{17,18} Possible mechanisms include inflammation,^{19,20} thromboembolic disease, and cellular energetic stress.²¹ Given the importance of neurodegeneration in determining age-related and dementia-related cognitive decline,²² a secondary aim of this study is to establish whether cumulative surgeries are associated with chronic neurodegeneration using longitudinal imaging data within the UK Biobank. Two important markers of neurodegeneration-related cognitive decline are reduced hippocampal volume²³ and cerebrovascular disease, denoted by white matter hyperintensities.²⁴ Given that a previous small cross-sectional study found that surgical exposure was associated with cortical atrophy but not white matter hyperintensities,²⁵ we sought to address the potential role of both of these pathologies in determining surgery-related cognitive decline in a large, population-based cohort.

Our primary outcome was to identify associations between hippocampal volume and white matter hyperintensities and surgery. We hypothesised that surgical exposure would be associated with decreased hippocampal volume and a greater volume of white matter hyperintensities. To provide face validity to our results we first verified that a change in cognition is predicted by cumulative surgeries, consistent with our previous report,¹⁰ and that changes in cognition would be associated with brain imaging markers of neurodegeneration, consistent with dementia literature.^{18,19} Finally, we conducted further analyses to investigate regional cortical neurodegeneration associated with surgery, given that degeneration occurs in a regional manner across many preclinical dementia phenotypes.²⁶

Methods

Study design and participants

For this population-based analysis, we used de-identified data from the UK Biobank in accordance with STROBE reporting guidelines. Ethics are managed by the UK Biobank Ethics and Governance Council and this study was approved under application 82574. The UK Biobank is a prospective, longitudinal, cohort study of 502 412 people living in the UK, aged 40–69 years at recruitment between 2006 and 2010, who consented in writing to long-term follow-up approximately every 4 years.¹¹ At each visit, participants attend an assessment centre to answer detailed questions about their health and lifestyle, undergo cognitive tests, and provide biofluids. A subset of participants (n=46 706) also undergo various imaging studies, including brain MRI.

We analysed data collected by the UK Biobank from March 13, 2006, to July 16, 2023. Participants who had withdrawn from the study by Aug 21, 2023, and participants with any dementia-related hospital inpatient diagnosis codes were excluded (appendix pp 6–7). Field references are in the appendix (pp 2–4) and further details are available in the UK Biobank data showcase.¹¹

Cognitive outcomes

Primary cognitive outcomes were reaction time (ms), fluid intelligence (correct responses), prospective memory (dichotomous, correct first time), and numeric memory (maximum correct matches) tests collected at instance 0 (baseline), instance 2 (imaging 1; the time of the first MRI scan), and instance 3 (imaging 2; the time of the second MRI scan). We also excluded cognitive data from instance 1, which were collected using web-based home assessment instead of assessment centre kiosks, owing to difficulty in controlling for variance as acknowledged in the literature.²⁷ The median time between instance 0 and instance 2 was 10.3 years (IQR 8.5–13.0) and between instance 2 and instance 3 was 2.3 years (2.2–4.0). Because we did not have an a priori hypothesis that a specific cognitive domain would be vulnerable to surgery, we analysed each test individually, expecting convergent results across all four tests. Secondary cognitive outcomes were the trail making test A (numeric; TMTA) and the trail making test B (alphanumeric; TMTB; both scored in deciseconds), the symbol digit substitution test (SDST; scored by mean correct matches), and paired associate learning (scored by correct matches). Cognitive tests were conducted in supervised kiosks at UK Biobank assessment centres.

The reaction time (Snap) test is a test of processing speed, measuring the mean duration in milliseconds until the button is pressed in rounds with correct paired matches. The fluid intelligence raw score is an unweighted sum of correct answers out of 13 language and mathematics questions. The prospective memory (Shape) test is a memory and inhibition test, dichotomised as either correct on the first attempt or not. The numeric memory raw score is the maximum number of digits remembered correctly in 12 rounds. Introduced from June, 2016, the TMTA and TMTB raw scores are the durations required to complete the tests in deciseconds. The SDST score was the mean number of correct matches. The

paired associate learning score was the number of word pairs, out of 12, that were correctly associated.

Brain imaging outcomes

Brain imaging was conducted at UK Biobank assessment centres using a standard Siemens Skyra 3T MRI machine running VD13A SP4 (as of October, 2015), with a standard Siemens 32-channel radiofrequency receive head coil in accordance with UK Biobank imaging acquisition protocols. We used two primary brain imaging outcomes preprocessed by the UK Biobank: the volume of white matter hyperintensities (from T1 and T2 fluid-attenuated inversion recovery images) and the average volume of left and right hippocampi (from T1 structural brain MRI, subcortical volumes; measured with the Oxford Centre for Functional Magnetic Resonance Imaging of the Brain Integrated Registration and Segmentation tool²⁸). These variables were selected because of known associations with dementia and cognition.^{22–24} Outcomes are expressed in cubic millimetres to provide a clinical perspective on the extent of injury.

Secondary brain imaging outcomes were generated by the UK Biobank using the FreeSurfer image analysis suite²⁹ from T1 structural volumes with parcellation of the white surface using the Desikan–Killiany–Tourville (DKT31) cortical labelling protocol.³⁰ We used regional mean thickness variables in millimetres.

Primary exposure and risk factors

Two clinician authors (PFY and RDS) categorised surgical complexity as minor, intermediate, major, major plus, complex major, or excluded (such as diagnostic day procedures and intracranial surgery), using the National Health Service medical record main procedure code (OPCS-4) and the BUPA Schedule of Procedures as of June 15, 2022. Cumulative surgeries were the sum of eligible procedures dated from 8 years before the baseline date of assessment (instance 0) to the instance date of assessment; surgeries from 1998 onwards were therefore considered. To adjust for confounding from other acute illnesses, hospital medical admissions were derived using ICD-10 main diagnosis codes.

Statistical analysis

Statistical analyses were conducted using R statistical software (version 4.2.3). Analyses were carried out in a multi-step manner. First, data were summarised using descriptive statistics. Next, the associations between surgery and cognitive outcomes and between brain imaging findings and cognitive outcomes were separately modelled. After verifying the associations, we investigated the association of surgery with brain imaging outcomes (primary outcome) and the possible mechanisms. Baseline characteristics were compared between participants who returned for further assessments and those who did not using Wilcoxon rank sum and Pearson's χ^2 tests.

We constructed linear mixed effects regression models for cognitive outcomes to assess the association with cumulative surgeries at each timepoint, adjusted for age and sex with random effect for participant (1|participant) using R packages *lme4* (version 1.1.34) and *mgcv* (version 1.8.42). To assess the appropriate relationship between outcomes, age, and

time in the study (years), we compared this linear model to linear mixed effects regression models with a quadratic term for baseline age (normalised/1000), quadratic age and time (years), and a generalised additive model with a cubic regression spline for age. We selected the model with quadratic baseline age (cognitive outcome \sim surgeries + age + age² + time + sex + 1|participant) as our foundational model on the basis of the lowest Bayesian information criterion (appendix p 10). Model conditional pseudo-R² values were calculated using the R package *MuMIn* (version 1.47.5).

We then constructed fully adjusted models that included as many of the 12 *Lancet* Commission³¹ dementia risk factors and Charlson index³² factors, such as solid tumour, as were available. Hearing difficulty, leukaemia, and lymphoma were then removed from the model for not contributing to the model with lower Bayesian information criteria in backwards selection. We further tested two-way interactions between surgeries and other medical admissions in line with our previous work, as well as between time and deprivation to address the considerable loss to follow-up between instance 0 and instance 2. The final fully adjusted model was cognitive outcome \sim surgeries + stroke admissions + other medical admissions + age + age² normalised/1000 + time \times deprivation + sex + education + BMI + BMI² normalised/1000 + smoking + depression + alcohol consumption + physical activity + hypertension + diabetes + solid tumour + 1|participant. The interaction for surgeries \times other medical admissions was retained in reaction time, fluid intelligence, and prospective memory models, but not in other outcomes. The least absolute shrinkage and selection operator confirmed that there were no superfluous variables in the final model.

We also did sensitivity analyses to investigate the effects of hospital admission for stroke, the number of surgeries (0, 1–2, 3–5, or 6 surgeries), high-complexity surgeries, and cumulative surgeries aggregated from the study baseline rather than our definition of 8 years before the baseline. Additionally, we conducted sensitivity analyses of dichotomous cognitive decline groups (\pm 1 SD and \pm 2 SD) for changes in reaction time between instance 2 and instance 0 and between instance 3 and instance 2. For logistic regression we used a foundation model, cognitive decline group \sim cumulative surgeries + baseline age + sex. We also ran these models for surgeries as a factor (1–2, 3–5, and 6 surgeries) and for high-complexity surgeries.

For brain imaging models we used the same foundational and fully adjusted models without interaction terms. In further exploratory imaging analysis, we averaged the mean thicknesses of the 31 cortical regions across hemispheres and constructed foundational models for DKT cortical region \sim surgeries with false discovery rate-corrected p values for multiple comparisons.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

502 412 participants are included in the UK Biobank cohort study (figure 1). After excluding 53 participants who withdrew from the study and 9557 with a dementia diagnosis, 492 802 (98.1%) participants were available for analysis in our study at baseline. The mean age was 56 years (SD 8) and, of the 492 802 participants, 268 733 (54.5%) were female and 224 069 (45.5%) were male, 463 474 (94.1%) were White, and 487 271 (98.88%) completed reaction time testing (table 1). Demographic data for all cognitive measures are shown in the appendix (p 11).

Mean cognitive scores at baseline were reaction time 559 ms (SD 117), fluid intelligence score 5.99 (2.16), and numeric memory score 6.49 (1.82). Of the 168 755 participants who took part in prospective memory testing, 129 271 (76.6%) scored correct first time (table 1). Although many of these tests are non-standard, they have been shown to correlate with reference tests of general cognitive ability ($r=0.83$, $p<0.001$) with concurrent validity and test-retest reliability,²⁷ and there is no evidence that baseline results in this sample were lower than those of the general population. Of the 492 802 participants, 71 873 (14.6%) returned at instance 2 and 6324 (1.3%) returned at instance 3. 46 706 (65.0%) of the 71 873 participants at instance 2 had a brain MRI scan, and 4779 (10.2%) of these 46 706 participants had a follow-up MRI scan at instance 3 (figure 1).

Overall, instance 0 participants who returned for follow-up at instance 2 were younger (median age 55 years [IQR 49 to 61] vs 58 years [50 to 63]), less deprived (Townsend deprivation index -2.58 [-3.88 to -0.40] vs -2.05 [-3.59 to 0.70]), more educated, and had lower BMI (26.1 kg/m² [IQR 23.7 to 28.9] vs 26.9 kg/m² [24.2 to 30.1]) and lower mean surgical exposure at baseline (0.63 [SD 1.03] vs 0.88 [1.34] surgeries) than those who did not return (appendix pp 15–16, all $p<0.0001$). Moreover, those who returned were more likely to have no surgical exposure at baseline (44 497 [61.9%] of 71 873) than those who were lost to follow-up (229 067 [54.4%] of 420 929; $p<0.0001$; appendix pp 15–16). Similarly, those returning at instance 2 had better cognitive results on average at baseline than those who did not return.

In our foundation model, we found that surgery was associated with a small, adverse change in cognition on all primary and secondary cognitive outcomes (table 2). Actual and predicted cognitive outcomes for age were plotted for comparison (figure 2). Reaction time increased on average by 1.957 ms per additional surgery (model conditional pseudo- R^2 [R_c^2] = 0.55). Exploratory stratification analysis of increased number of surgeries ($\beta=2.53$ for 1–2 surgeries, $\beta=7.63$ for 3–5 surgeries, and $\beta=12.46$ for 6 surgeries; appendix p 18) or high-complexity surgeries ($\beta=2.97$; appendix p 19) showed similar results.

Lower fluid intelligence scores were associated with more surgeries ($\beta=-0.11$, $R_c^2 = 0.67$; table 2). The odds of correct first attempt matches in prospective memory tests were reduced by increasing numbers of cumulative surgeries (odds ratio [OR]=0.93 [95% CI 0.92–0.94]) and more so by high-complexity surgeries (0.89 [0.87–0.91]) than by no surgeries. Complete model outputs are shown in the appendix (pp 17–19).

Compared with the foundation model, the fully adjusted linear mixed effects regression models showed smaller associations of cumulative surgeries with adverse changes in cognition (table 2; all $p < 0.0001$ unless indicated), with an average increase in reaction time of 0.273 ms ($p=0.033$, $R_c^2 = 0.56$), a decrease in fluid intelligence score of 0.057 correct responses ($R_c^2 = 0.66$), lower odds of correct-first-time matches in prospective memory tests (OR 0.96, 95% CI 0.95–0.97), and lower numeric memory by 0.025 matches ($R_c^2 = 0.57$) per additional surgical procedure (see appendix pp 20–22 for complete model outputs). Similar relationships were observed in secondary cognitive measures (table 2).

Sensitivity analyses comparing participants who were admitted to hospital with stroke with participants who were not supported that this variable was not confounding in the model (appendix pp 42–47). Sensitivity analysis for participants who had a decline in reaction time showed that surgical exposure increased the likelihood of moving into a 1 SD cognitive decline group by 1.9% (OR 1.02, $p<0.0001$) and a 2 SD cognitive decline group by 2.5% (OR 1.03, $p=0.0048$; appendix p 51) per surgery. This likelihood increased with increasing numbers of surgeries and with high-complexity surgeries (appendix p 51).

Before testing the association of surgery and the imaging variables, we verified that the primary imaging outcomes were associated with cognition in foundational models for all available data in instance 2 and instance 3. An increase of 1 mm³ in white matter hyperintensities was associated with, on average, a 0.0008 ms increase in reaction time and a decrease in score of 0.00001 correct responses in the fluid intelligence test (all $p<0.0001$, table 2). Each 1 mm³ increase in hippocampal volume was associated with an average decrease in reaction time of 0.017 ms and an average increase in score of 0.0005 correct responses in the fluid intelligence test (table 2). These results were sustained in secondary cognitive measures and in fully adjusted models (table 2, appendix pp 23–26).

Surgeries were positively associated with white matter hyperintensities, with each surgery increasing white matter hyperintensities volume on average by 175.17 mm³ in foundational models (95% CI 144.88 to 205.47; appendix p 31). Surgeries were also negatively associated with average hippocampal volume ($\beta=-8.52$ mm³ [-10.43 to -6.62]; appendix p 31; all $p<0.0001$ unless indicated). These associations were moderated in fully adjusted models (white matter hyperintensities $\beta=100.02$ mm³ [66.17 to 133.87]; average hippocampal volume $\beta=-5.76$ mm³ [-7.89 to -3.64]; appendix p 34). For comparison, stroke admissions were associated with a greater change in white matter hyperintensities ($\beta=6271.29$ mm³) and average hippocampal volume reduction ($\beta=-161.29$ mm³; appendix p 34). In exploratory stratification analyses, greater associations were observed in participants with a higher number of surgeries and in those with more complex surgeries in fully adjusted models (appendix pp 36–39).

High-complexity surgeries were also associated with increased white matter hyperintensities ($\beta=247.73$ mm³ [161.70 to 333.76]) and reduced average hippocampal volume ($\beta=-17.05$ mm³ [-22.46 to -11.64]; appendix pp 38–39). Sensitivity analysis for participants with paired scans showed similar associations (data not shown).

Foundational linear mixed effects regression modelling showed that surgery was associated with reductions in mean cortical thickness in 23 (65%) of 31 DKT regions (figure 3, appendix p 40). In fully adjusted models, surgery was associated with neurodegeneration in the insula ($\beta=-0.001$ mm) and superior temporal ($\beta=-0.001$ mm) cortical regions after false discovery rate correction for multiple comparisons across 31 brain regions (figure 3, appendix p 40).

Discussion

This study of the UK Biobank cohort of approximately half a million participants showed small, adverse effects of surgery on cognition and neurodegeneration over time after adjustment for age, sex, other hospital admissions, and risk factors for dementia. The effects were small relative to stroke or other medical admissions but increased with cumulative surgical procedures and greater surgical complexity. These changes are quantified per procedure, with additional surgeries associated with compounding effects of poorer cognition, reduced hippocampal volume, and a greater volume of white matter hyperintensities. In exploratory analyses, regional cortical neurodegeneration in the insula and superior temporal regions was noted.

To provide an example that puts the cognitive changes in context, these data suggest that, on average per surgery, there is an incremental slowing in reaction time of 2 ms. The average risk of cognitive decline at thresholds of 1 SD or 2 SD also increases by approximately 2% per surgery; these thresholds are considered equivalent to minor (1 SD) and major (2 SD) cognitive change.³³ These data suggest that, on average, surgery is safe, but there are small risks of more profound deficits. The clinical importance of these results is underpinned by respective changes in brain structure and pathology. It is reassuring that the absolute effect on cognition per surgery was small, although the risks are cumulative, and therefore we consider that our work should not be interpreted as a deterrent to patients from undergoing worthwhile treatment. Nonetheless, we propose that we should seek to mitigate incremental detrimental risks of surgical exposure, particularly as many surgeries are elective and protective strategies, such as cognitive prehabilitation, could be instituted.

These results are in line with our previous work in the Whitehall II study^{9,10} and with the wider literature.^{7,8,34} Mechanisms proposed for adverse cognitive effects in older people after surgery include ischaemic brain injury and inflammation. The NeuroVISION study (n=1114) of non-cardiac surgery reported that perioperative covert stroke, which occurred in 7% of patients, increased the risk of cognitive decline 1 year after surgery compared with patients without perioperative covert stroke.^{17,35} Similar findings have been reported after cardiac surgery.^{4,7,8,34} A small cohort analysis of the Mayo Clinic Study of Aging²⁵ with imaging (n=1410) showed cortical thinning in regions associated with Alzheimer's disease (hippocampal, entorhinal, para-hippocampal, middle, and superior temporal regions) for participants exposed to surgery in the previous 20 years, but not an association with white matter hyperintensities or brain infarcts. This finding challenges associations with chronic cerebrovascular disease, prompting us to conduct a parallel study of neurodegeneration and cerebrovascular disease.

Studies of neuronal injury plasma biomarkers support that surgery is associated with contemporaneous brain injury,^{12–14,19} including both covert stroke¹⁸ and inflammation-based injury.^{12,14,19} Inflammation leads to breakdown of the blood–brain barrier, neuroinflammation, energetic stress, and increases in lactate concentration.^{12,19} Furthermore, slow clearance of the inflammatory biomarker IL-6 is associated with a slower cognitive recovery.³⁶ This finding suggests that anti-inflammatory therapies could improve perioperative care for older people, particularly if targeted to patients in whom perioperative inflammation resolves more slowly.

Notably, the cognitive changes observed in this study were associated with both reduced hippocampal volume and increased cerebrovascular disease.^{23–25} These outcomes were also associated with cumulative surgeries in our imaging models. Furthermore, in our exploratory analyses of DKT cortical regions, small, negative changes in the insula and superior temporal gyrus regions were associated with surgeries. These are known regions affected by oxidative stress,³⁷ impaired connectivity, and atrophy associated with Alzheimer’s disease and related dementias.

Our study has several strengths. First, the size of this dataset and the lengthy cognitive follow-up in the UK Biobank study provides the power to detect small associations between surgery and cognition. Second, controlling for stroke and other medical admissions avoids exaggeration of the associations between surgery and cognition. Third, our longitudinal linear mixed effects models with random effect for participant uses all available data and adjusts for the accelerated effects of increasing age on cognitive decline. Finally, our exclusion of intracranial surgeries and diagnostic procedures, and our sensitivity analyses including and excluding minor and intermediate surgeries, provide greater confidence in our findings.

Our study also has many limitations. First, we cannot exclude unmeasured confounders, nor can we ascribe causality in this observational data analysis. Furthermore, we cannot separate the pathology driving the need for an operation from the operation itself; our estimates are therefore likely to be at the upper limit of the estimate of the injury from surgery. Second, our classification of surgical severity using health insurance risk codes and main admissions diagnosis with ICD-10 codes cannot reflect the one-to-many nature of these data, ongoing code changes, and the complexity of many procedures and admissions, including postoperative complications. Third, the substantial loss to follow-up might bias our findings, and our data suggested that older participants with higher risk factors were more likely to not return; those returning at instance 2 were also younger and a greater proportion had no surgeries at baseline than those who did not return. Finally, the introduction of imaging in instances 2 and 3 only, although mostly due to the availability of funding and technology, means that the proportion of patients with imaging data was relatively small. Nonetheless, this is still the largest imaging study that we are aware of on this topic.

Overall, these results using a large, population-based dataset validate our previous research and suggest that cumulative surgeries, in number and complexity, are safe on average. However, an association of surgeries with small, adverse changes in cognition and neurodegeneration persists. This risk should be communicated to patients and families and

weighed against the benefits of surgery. Perioperative brain health remains an important goal in the clinical care of older and more vulnerable patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was funded by the ANZCA Foundation and the University of Sydney.

Declaration of interests

RDS and RL received support from US National Institutes of Health (NIH) grant R01 AG063849-01. CC received support from NIH R01 HD098202-02 and NIH R01 NS117901-01. SN declares consulting fees from Eisai, Roche, and Nutrica Pharmaceuticals. All other authors declare no competing interests.

Data sharing

This research has been conducted with UK Biobank resources under application 82574. Data are available via the UK Biobank on application and with permission of the UK Biobank Ethics and Governance Council.

References

1. Fowler AJ, Abbott TEF, Prowle J, Pearse RM. Age of patients undergoing surgery. *Br J Surg* 2019; 106: 1012–18. [PubMed: 31115918]
2. Rowley P, Boncyk C, Gaskell A, et al. What do people expect of general anaesthesia? *Br J Anaesth* 2017; 118: 486–88. [PubMed: 28403409]
3. Guo LY, Kaustov L, Brenna CTA, et al. Cognitive deficits after general anaesthesia in animal models: a scoping review. *Br J Anaesth* 2023; 130: e351–60. [PubMed: 36402576]
4. Kunicki ZJ, Ngo LH, Marcantonio ER, et al. Six-year cognitive trajectory in older adults following major surgery and delirium. *JAMA Intern Med* 2023; 183: 442–50. [PubMed: 36939716]
5. Greaves D, Psaltis PJ, Ross TJ, et al. Cognitive outcomes following coronary artery bypass grafting: a systematic review and meta-analysis of 91,829 patients. *Int J Cardiol* 2019; 289: 43–49. [PubMed: 31078353]
6. Evered LA, Silbert BS, Scott DA, Maruff P, Ames D. Prevalence of dementia 7.5 years after coronary artery bypass graft surgery. *Anesthesiology* 2016; 125: 62–71. [PubMed: 27127919]
7. Whitlock EL, Diaz-Ramirez LG, Smith AK, Boscardin WJ, Avidan MS, Glymour MM. Cognitive change after cardiac surgery versus cardiac catheterization: a population-based study. *Ann Thorac Surg* 2019; 107: 1119–25. [PubMed: 30578068]
8. Whitlock EL, Diaz-Ramirez LG, Smith AK, et al. Association of coronary artery bypass grafting vs percutaneous coronary intervention with memory decline in older adults undergoing coronary revascularization. *JAMA* 2021; 325: 1955–64. [PubMed: 34003225]
9. Krause BM, Manning HJ, Sabia S, Singh-Manoux A, Sanders RD. Association of major surgical admissions with quality of life: 19-year follow-up of the Whitehall II longitudinal prospective cohort study. *JAMA Surg* 2022; 157: 275–77. [PubMed: 35080594]
10. Krause BM, Sabia S, Manning HJ, Singh-Manoux A, Sanders RD. Association between major surgical admissions and the cognitive trajectory: 19 year follow-up of Whitehall II cohort study. *BMJ* 2019; 366: 14466. [PubMed: 31391161]
11. Sudlow C, Gallacher J, Allen N, et al. UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med* 2015; 12: e1001779. [PubMed: 25826379]

12. Casey CP, Lindroth H, Mohanty R, et al. Postoperative delirium is associated with increased plasma neurofilament light. *Brain* 2020; 143: 47–54. [PubMed: 31802104]
13. Evered L, Silbert B, Scott DA, Ames D, Maruff P, Blennow K. Cerebrospinal fluid biomarker for Alzheimer disease predicts postoperative cognitive dysfunction. *Anesthesiology* 2016; 124: 353–61. [PubMed: 26580833]
14. Ballweg T, White M, Parker M, et al. Association between plasma tau and postoperative delirium incidence and severity: a prospective observational study. *Br J Anaesth* 2021; 126: 458–66. [PubMed: 33228978]
15. Parker M, White M, Casey C, et al. Cohort analysis of the association of delirium severity with cerebrospinal fluid amyloid-tau neurodegeneration pathologies. *J Gerontol A Biol Sci Med Sci* 2022; 77: 494–501. [PubMed: 34260706]
16. Kant IMJ, de Bresser J, van Montfort SJT, et al. The association between brain volume, cortical brain infarcts, and physical frailty. *Neurobiol Aging* 2018; 70: 247–53. [PubMed: 30048892]
17. Mrkobrada M, Chan MTV, Cowan D, et al. Perioperative covert stroke in patients undergoing non-cardiac surgery (NeuroVISION): a prospective cohort study. *Lancet* 2019; 394: 1022–29. [PubMed: 31422895]
18. Taylor J, Eisenmenger L, Lindroth H, et al. Perioperative ischaemic brain injury and plasma neurofilament light: a secondary analysis of two prospective cohort studies. *Br J Anaesth* 2023; 130: e361–69. [PubMed: 36437124]
19. Taylor J, Parker M, Casey CP, et al. Postoperative delirium and changes in the blood–brain barrier, neuroinflammation, and cerebrospinal fluid lactate: a prospective cohort study. *Br J Anaesth* 2022; 129: 219–30. [PubMed: 35144802]
20. Kealy J, Murray C, Griffin EW, et al. Acute inflammation alters brain energy metabolism in mice and humans: role in suppressed spontaneous activity, impaired cognition, and delirium. *J Neurosci* 2020; 40: 5681–96. [PubMed: 32513828]
21. Medel V, Crossley N, Gajardo I, et al. Whole-brain neuronal MCT2 lactate transporter expression links metabolism to human brain structure and function. *Proc Natl Acad Sci USA* 2022; 119: e2204619119. [PubMed: 35939682]
22. Jack CR Jr, Bennett DA, Blennow K, et al. NIA-AA Research Framework: toward a biological definition of Alzheimer’s disease. *Alzheimers Dement* 2018; 14: 535–62. [PubMed: 29653606]
23. Ferreira D, Nordberg A, Westman E. Biological subtypes of Alzheimer disease: a systematic review and meta-analysis. *Neurology* 2020; 94: 436–48. [PubMed: 32047067]
24. Heinzinger N, Maass A, Berron D, et al. Exploring the ATN classification system using brain morphology. *Alzheimers Res Ther* 2023; 15: 50. [PubMed: 36915139]
25. Sprung J, Kruthiventi SC, Warner DO, et al. Exposure to surgery under general anaesthesia and brain magnetic resonance imaging changes in older adults. *Br J Anaesth* 2019; 123: 808–17. [PubMed: 31587833]
26. Dickerson BC, Stoub TR, Shah RC, et al. Alzheimer-signature MRI biomarker predicts AD dementia in cognitively normal adults. *Neurology* 2011; 76: 1395–402. [PubMed: 21490323]
27. Fawns-Ritchie C, Deary IJ. Reliability and validity of the UK Biobank cognitive tests. *PLoS One* 2020; 15: e0231627. [PubMed: 32310977]
28. Patenaude B, Smith SM, Kennedy DN, Jenkinson M. A Bayesian model of shape and appearance for subcortical brain segmentation. *NeuroImage* 2011; 56: 907–22. [PubMed: 21352927]
29. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis: i. segmentation and surface reconstruction. *NeuroImage* 1999; 9: 179–94. [PubMed: 9931268]
30. Klein A, Tourville J. 101 labeled brain images and a consistent human cortical labeling protocol. *Front Neurosci* 2012; 6: 171. [PubMed: 23227001]
31. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the *Lancet* Commission. *Lancet* 2020; 396: 413–46. [PubMed: 32738937]
32. Charlson ME, Carrozzino D, Guidi J, Patierno C. Charlson Comorbidity Index: a critical review of clinimetric properties. *Psychother Psychosom* 2022; 91: 8–35. [PubMed: 34991091]
33. Evered L, Silbert B, Knopman DS, et al. Recommendations for the nomenclature of cognitive change associated with anaesthesia and surgery—2018. *Br J Anaesth* 2018; 121: 1005–12. [PubMed: 30336844]

34. Schulte PJ, Roberts RO, Knopman DS, et al. Association between exposure to anaesthesia and surgery and long-term cognitive trajectories in older adults: report from the Mayo Clinic Study of Aging. *Br J Anaesth* 2018; 121: 398–405. [PubMed: 30032878]
35. Marcucci M, Chan MTV, Smith EE, Absalom AR, Devereaux PJ. Prevention of perioperative stroke in patients undergoing non-cardiac surgery. *Lancet Neurol* 2023; 22: 946–58. [PubMed: 37739575]
36. Taylor J, Wu JG, Kunkel D, et al. Resolution of elevated interleukin-6 after surgery is associated with return of normal cognitive function. *Br J Anaesth* 2023; 131: 694–704. [PubMed: 37385855]
37. Youssef P, Chami B, Lim J, Middleton T, Sutherland GT, Witting PK. Evidence supporting oxidative stress in a moderately affected area of the brain in Alzheimer’s disease. *Sci Rep* 2018; 8: 11553. [PubMed: 30068908]

Research in context

Evidence before this study

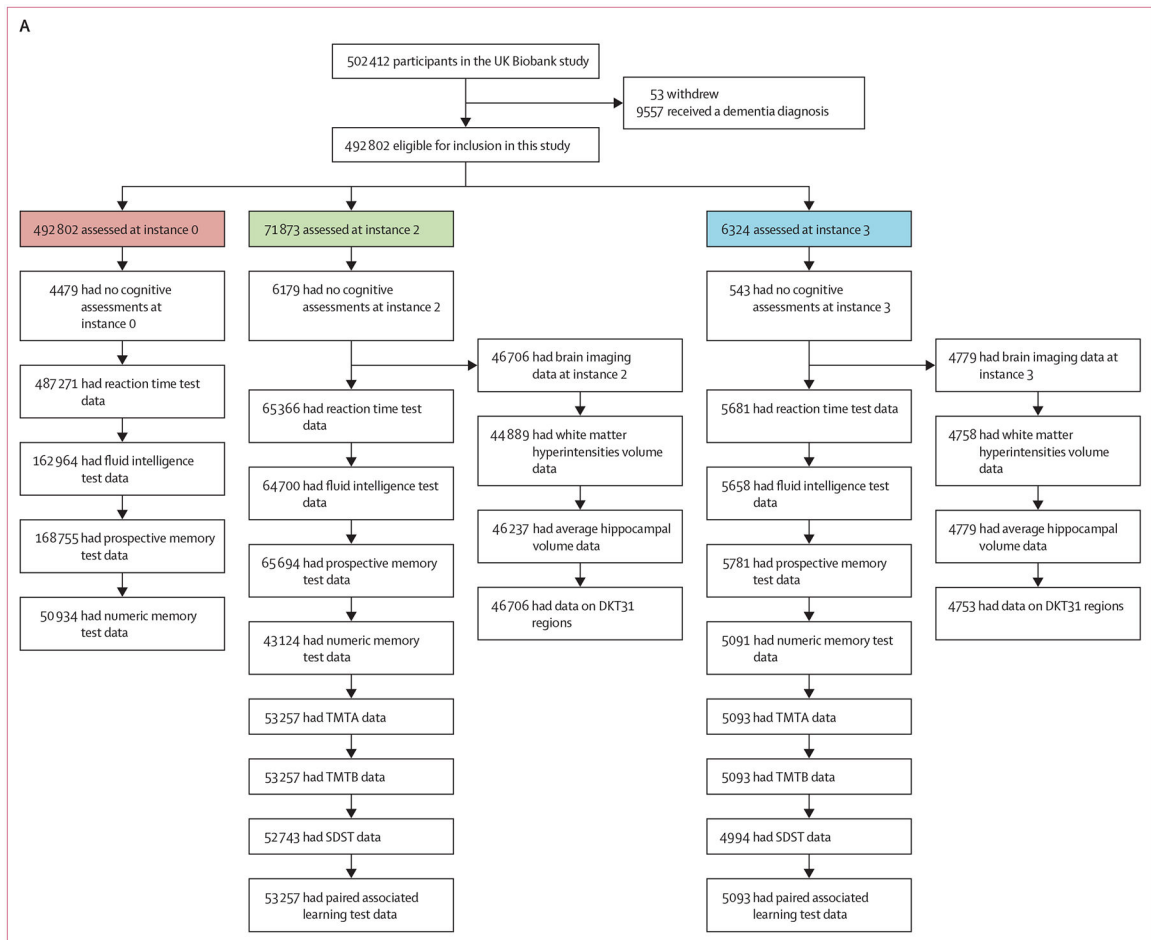
Previous work has found an association between cognitive decline and surgical admissions. We searched PubMed from database inception to July 31, 2024, using the terms (surg* OR anaesth* OR anesth*) AND (cognit* OR dementia OR neurodegenerat* or neurocognit*) with no language restrictions. Cohort studies have shown that, on average, there is a small but incremental cognitive decline associated with cumulative surgeries, for which vascular and inflammatory mechanisms have been suggested. However, data on the mechanisms underlying any cognitive change are scarce, and the role of neurodegeneration has been inadequately investigated to date.

Added value of this study

This study validates previous findings with thorough adjustment for dementia risk factors and provides a biological link between surgical exposure and neurodegeneration.

Implications of all the available evidence

This evidence substantiates that cognitive effects per surgical procedure are small but can be incrementally detrimental for older people. This finding should not deter treatment but suggests that the development of neuroprotective strategies should be prioritised in health care.



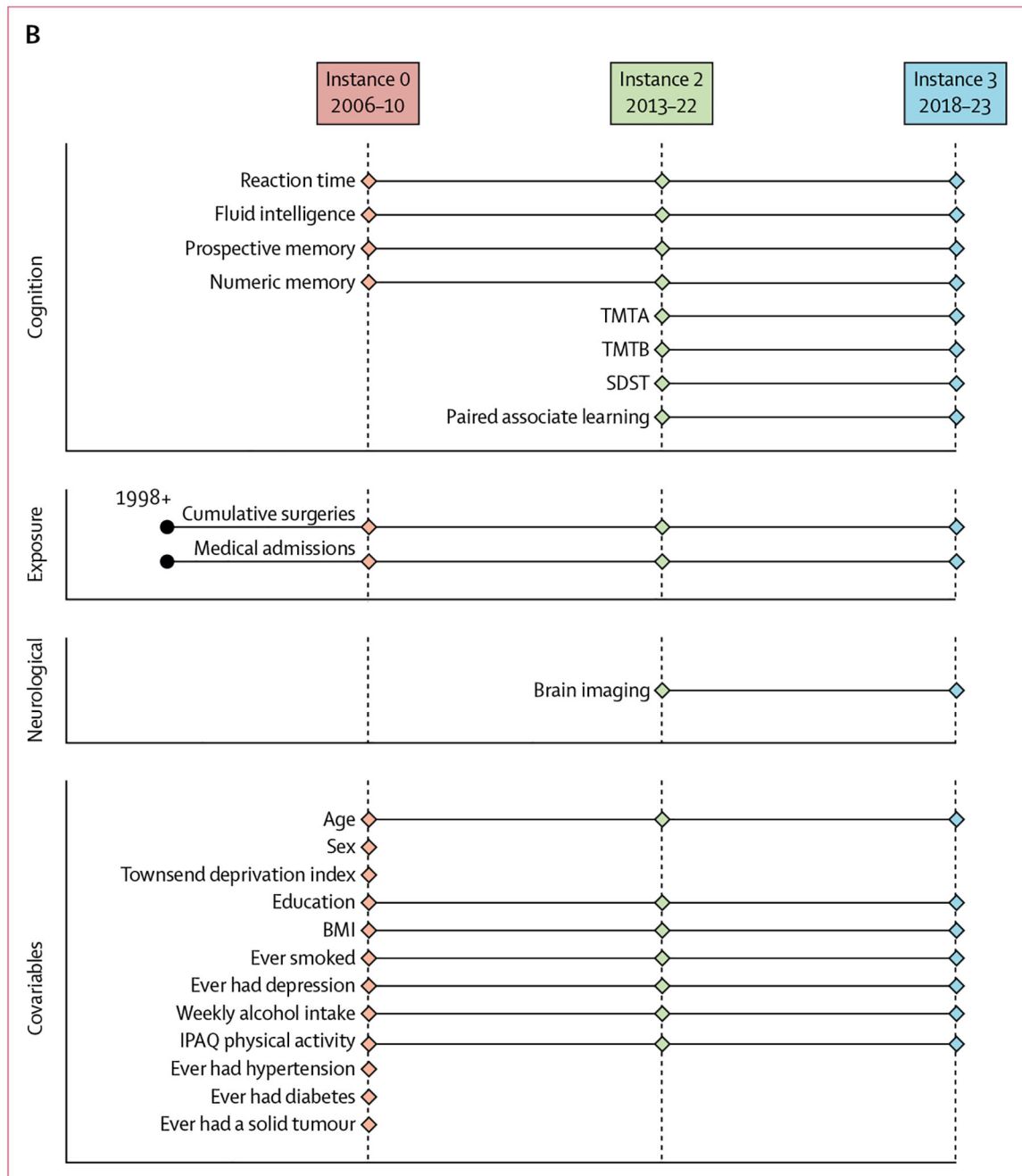


Figure 1: Study profile and timeline

(A) Study profile. (B) Study timeline. DKT=Desikan–Killiany–Tourville.

IPAQ=International Physical Activity Questionnaire. SDST=symbol digit substitution test.

TMTA=trail making test A (numeric). TMTB=trail making test B (alphanumeric).

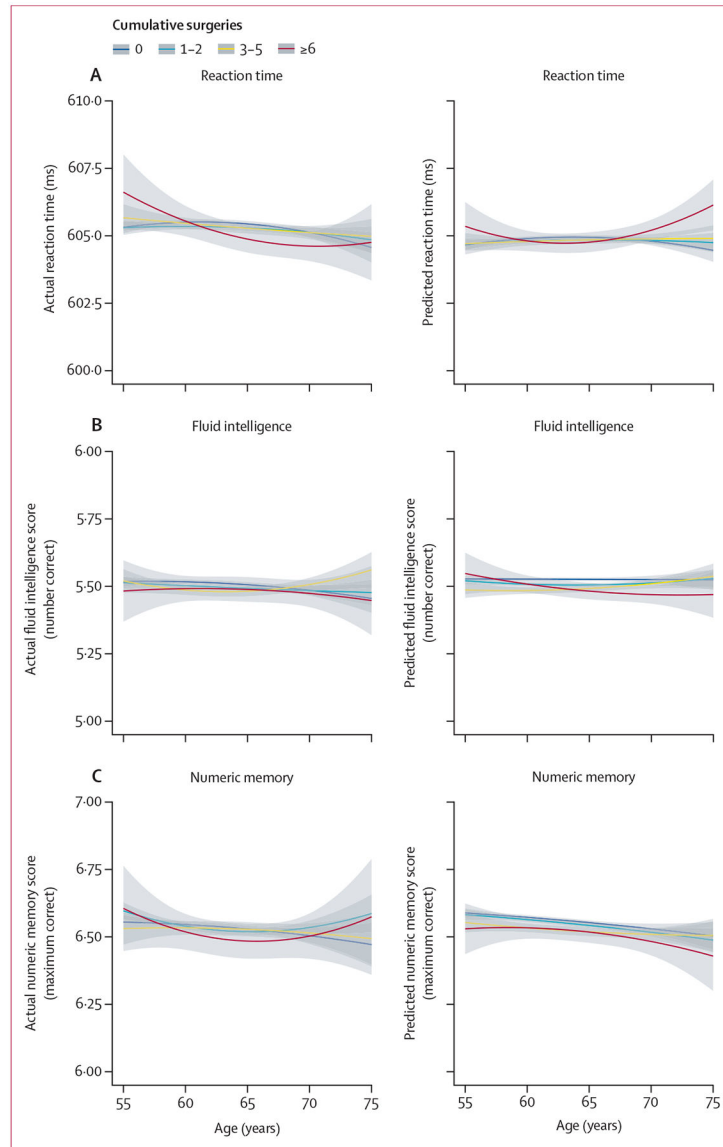


Figure 2: Actual and predicted values from fully adjusted models showing the association of cumulative surgeries with primary cognitive outcomes at specific ages
 (A) Reaction time. (B) Fluid intelligence. (C) Numeric memory. All analyses include instance 0, instance 2, and instance 3 data. All linear mixed effects models were fully adjusted for cognitive outcome \sim cumulative surgeries + cumulative stroke admissions + cumulative other medical admissions + baseline age + baseline age² normalised + time \times deprivation + female sex + education + BMI + BMI² normalised + smoking + depression + alcohol consumption + physical activity + hypertension + diabetes + solid tumour + (1|participant). Quadratic terms were normalised by variable²/1000. Interaction for surgeries \times other medical admissions is included in the reaction time and fluid intelligence fully adjusted models. Linear smoothing uses the formula $y \sim x + I(x^2)$.

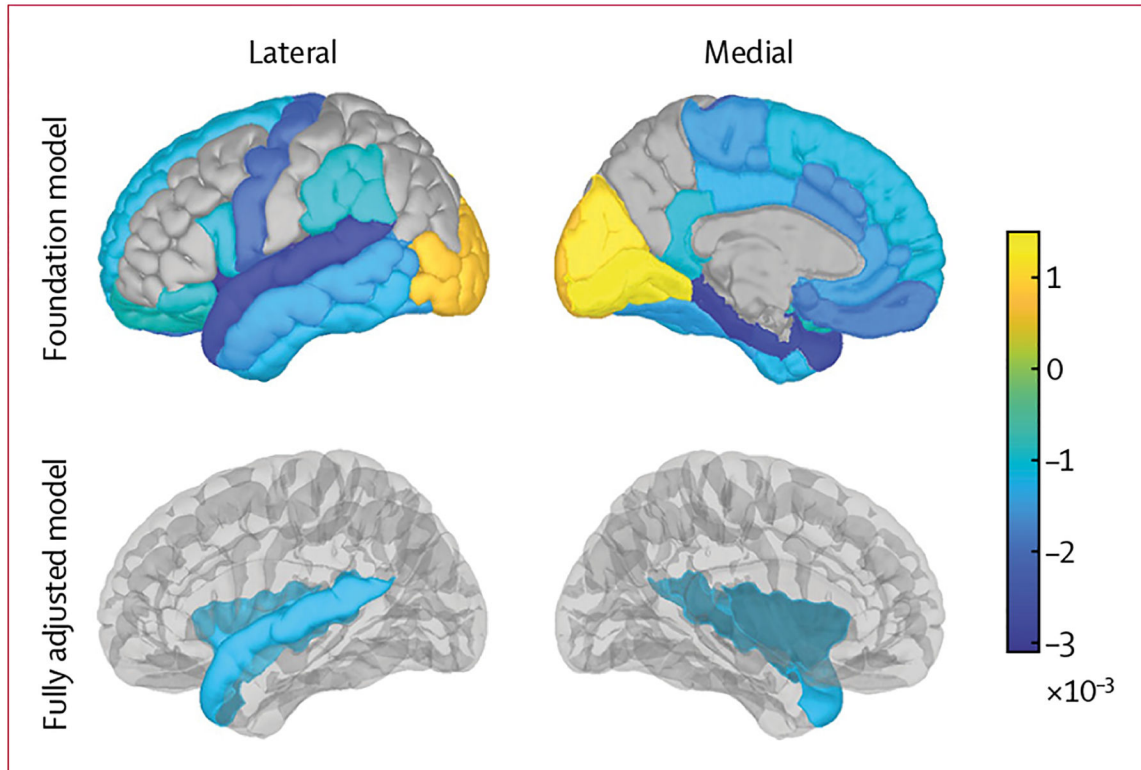


Figure 3: Brain maps of associations of DKT cortical regions with surgeries (β coefficients)
DKT=Desikan–Killiany–Tourville.

Table 1:

Participant demographics

	Instance 0 n=492 802	Instance 2 n=71 873	Instance 3 n=6324
Age, years	56 (8)	65 (8)	65 (7)
Sex			
Female	268 733 (54.5%)	37 118 (51.6%)	3256 (51.5%)
Male	224 069 (45.5%)	34 755 (48.4%)	3068 (48.5%)
Ethnic background			
Asian	11 296 (2.3%)	1038 (1.4%)	79 (1.3%)
Black	7914 (1.6%)	516 (0.7%)	38 (0.6%)
White	463 474 (94.1%)	69 322 (96.5%)	6144 (97.2%)
Do not know or prefer not to answer	1832 (0.4%)	193 (0.3%)	13 (0.2%)
Other	7417 (1.5%)	776 (1.1%)	48 (0.8%)
Missing	869 (0.2%)	28 (<0.1%)	2 (<0.1%)
Townsend deprivation index at instance 0	-1.30 (3.09)	-1.85 (2.75)	-1.92 (2.66)
Education			
Secondary (<12 years total)	134 863 (27.4%)	16 308 (22.7%)	1453 (23.0%)
Secondary (12 years total)	35 836 (7.3%)	5712 (8.0%)	523 (8.3%)
Tertiary (>12 years total)	230 355 (46.7%)	42 735 (59.5%)	3707 (58.6%)
None of the above	82 002 (16.6%)	4637 (6.5%)	292 (4.6%)
Prefer not to answer	5279 (1.1%)	297 (0.4%)	17 (0.3%)
Missing	4467 (0.9%)	2184 (3.0%)	332 (5.3%)
BMI, kg/m ²	27.4 (4.8)	26.7 (4.5)	26.4 (4.4)
Ever smoked	292 623 (59.4%)	43 628 (60.7%)	3805 (60.2%)
Hearing difficulty	118 962 (24.1%)	27 858 (38.8%)	2627 (41.5%)
Ever had depression	87 982 (17.9%)	38 520 (53.6%)	3918 (62.0%)
Weekly alcohol intake, standard units	15 (18)	13 (15)	13 (15)
IPAQ physical activity at instance 0			
Low	70 318 (14.3%)	10 929 (15.2%)	931 (14.7%)
Moderate	153 857 (31.2%)	25 001 (34.8%)	2202 (34.8%)
High	154 546 (31.4%)	23 927 (33.3%)	2222 (35.1%)
Missing	114 081 (23.2%)	12 016 (16.7%)	969 (15.3%)
Ever had hypertension	133 789 (27.2%)	19 458 (27.1%)	1655 (26.2%)
Ever had myocardial infarction	11 685 (2.4%)	1461 (2.0%)	132 (2.1%)
Ever had stroke	6773 (1.4%)	871 (1.2%)	82 (1.3%)
Ever had diabetes	22 107 (4.5%)	2807 (3.9%)	251 (4.0%)
Ever had lymphoma	1966 (0.4%)	341 (0.5%)	25 (0.4%)
Ever had a solid tumour	41 813 (8.5%)	8684 (12.1%)	740 (11.7%)
Mean number of cumulative surgeries	0.84 (1.30)	1.87 (2.09)	1.89 (2.01)
Number of cumulative surgeries			
0	273 564 (55.5%)	21 369 (29.7%)	1760 (27.8%)
1–2	171 644 (34.8%)	30 165 (42.0%)	2726 (43.1%)

	Instance 0 n=492 802	Instance 2 n=71 873	Instance 3 n=6324
3-5	41 813 (8.5%)	15 882 (22.1%)	1481 (23.4%)
6	5781 (1.2%)	4457 (6.2%)	357 (5.7%)
Reaction time, ms	559 (117)	601 (112)	594 (108)
Fluid intelligence score, number correct	5.99 (2.16)	6.48 (2.07)	6.74 (2.02)
Prospective memory score, number correct on first attempt	129 271/168 755 (76.6%)	53 693/65 694 (81.7%)	4973/5781 (86.0%)
Numeric memory score, maximum correct	6.49 (1.82)	6.62 (1.55)	6.67 (1.76)
TMTA score at instance 2, ds		230 (90)	218 (89)
TMTB score at instance 2, ds		570 (294)	528 (261)
SDST score at instance 2, correct matches		18.6 (5.3)	19.7 (5.3)
Paired associate learning score at instance 2, correct pairs		6.73 (2.66)	7.11 (2.59)
Volume of white matter hyperintensities at instance 2, mm ³		5216 (6877)	5176 (6553)
Volume of hippocampus at instance 2, mm ³		3818 (447)	3814 (439)

Data are n (%), n/N (%), or mean (SD). ds=deciseconds. IPAQ=International Physical Activity Questionnaire. SDST=symbol digit substitution test. TMTA=trail making test A (numeric). TMTB=trail making test B (alphanumeric).

Association of cumulative surgeries and mean brain volumes with different cognitive outcomes

Table 2:

Surgeries	Volume of white matter hyperintensities			Volume of average hippocampus		
	Foundation model, mean change per surgery (95% CI)*	Fully adjusted model, mean change per mm ³ ‡	Fully adjusted model, mean change per mm ³ §	Foundation model, mean change per mm ³ ‡	Fully adjusted model, mean change per mm ³ §	Fully adjusted model, mean change per mm ³ §
Model 1; reaction time, ms	1.957 (1.742 to 2.173)	0.273 (0.022 to 0.523)	0.0008 (0.0006 to 0.0009)	0.0007 (0.0006 to 0.0009)	-0.017 (-0.019 to -0.014)	-0.015 (-0.017 to -0.012)
Model 2; fluid intelligence score, number correct	-0.114 (-0.121 to -0.107)	-0.057 (-0.064 to -0.049)	-1.11 × 10 ⁻⁵ (-1.42 × 10 ⁻⁵ to -8.06 × 10 ⁻⁶)	-8.16 × 10 ⁻⁶ (-1.11 × 10 ⁻⁵ to -5.18 × 10 ⁻⁶)	0.0005 (0.0005 to 0.0006)	0.0004 (0.0004 to 0.0005)
Model 3; numeric memory score, maximum correct	-0.052 (-0.060 to -0.044)	-0.025 (-0.034 to -0.016)	-8.73 × 10 ⁻⁶ (-1.12 × 10 ⁻⁵ to -6.24 × 10 ⁻⁶)	-6.16 × 10 ⁻⁶ (-8.68 × 10 ⁻⁶ to -3.66 × 10 ⁻⁶)	0.0003 (0.0002 to 0.0003)	0.0002 (0.0001 to 0.0003)
Model 4; TMTA, deciseconds	1.681 (1.320 to 2.043)	0.686 (0.275 to 1.097)	0.0003 (0.0002 to 0.0005)	0.0003 (0.0001 to 0.0004)	-0.011 (-0.014 to -0.009)	-0.009 (-0.012 to -0.008)
Model 5; TMTB, deciseconds	8.562 (7.426 to 9.698)	4.351 (3.081 to 5.622)	0.0013 (0.0009 to 0.0017)	0.0012 (0.0007 to 0.0016)	-0.032 (-0.039 to -0.025)	-0.025 (-0.032 to -0.018)
Model 6; SDST score, correct matches	-0.193 (-0.214 to -0.173)	-0.094 (-0.117 to -0.070)	-4.82 × 10 ⁻⁷ (-6.44 × 10 ⁻⁷ to -3.2 × 10 ⁻⁷)	-3.53 × 10 ⁻⁷ (-5.2 × 10 ⁻⁷ to -1.85 × 10 ⁻⁷)	7.95 × 10 ⁻⁶ (5.41 × 10 ⁻⁶ to 1.05 × 10 ⁻⁵)	6.25 × 10 ⁻⁶ (3.67 × 10 ⁻⁶ to 8.84 × 10 ⁻⁶)
Model 7; paired associate learning score, correct matches	-0.099 (-0.110 to -0.088)	-0.050 (-0.062 to -0.038)	-1.67 × 10 ⁻⁵ (-2.08 × 10 ⁻⁵ to -1.26 × 10 ⁻⁵)	-1.15 × 10 ⁻⁵ (-1.56 × 10 ⁻⁵ to -7.3 × 10 ⁻⁶)	0.0003 (0.0003 to 0.0004)	0.0002 (0.0002 to 0.0003)

All available data were used across instances 0, 2, and 3. All comparisons are p < 0.0001, except the fully adjusted reaction time surgeries estimate (p = 0.033). All models were linear mixed effects models of all available data. The complete models are available in the appendix (pp 17–30). SDST = symbol digit substitution test, TMTA = trail making test A (numeric), TMTB = trail making test B (alphanumeric).

* Foundation surgeries models were minimally adjusted: cognitive outcome ~ surgeries + age + age² normalised / 1000 + time + female sex + (1 participant).

‡ Fully adjusted surgeries models were adjusted for cognitive outcome ~ surgeries + cumulative stroke admissions + cumulative other medical admissions + age + age² normalised + time * deprivation + female sex + education + BMI + normalised BMI² + smoking + depression + alcohol consumption + physical activity + hypertension + diabetes + solid tumour + (1 participant). Quadratic terms were normalised by variable² / 1000. Interaction for surgeries × other medical admissions and time × deprivation was included in reaction time, fluid intelligence, and prospective memory fully adjusted models. Cumulative surgical procedures were calculated from 8 years before the baseline (instance 0).

§ Foundation imaging models were adjusted for cognitive outcome ~ imaging variable + baseline age + baseline age² normalised + time + female sex + (1 participant). Quadratic terms were normalised by variable² / 1000.

¶ Fully adjusted imaging models were adjusted for cognitive outcome ~ imaging variables + cumulative stroke admissions + cumulative other medical admissions + baseline age + baseline age² normalised + time + female sex + deprivation + education + BMI + normalised BMI² + smoking + depression + alcohol consumption + physical activity + hypertension + diabetes + solid tumour + (1 participant). Quadratic terms were normalised by variable² / 1000. Interaction terms were not justified and therefore not used in imaging models.