POSTOPERATIVE cognitive dysfunction (POCD) involves pre- to postoperative reductions in memory, mental flexibility, and information processing. It is distinct from delirium and dementia. POCD occurs after noncardiac surgery to some extent in all age groups at hospital discharge (37% for those aged 18 to 39, 30% for 40 to 59 yr, and 41% for 60 and older), with longer-term POCD at 3 months for adults older than 60 yr (10 to 13%). POCD is associated with early retirement and dependency on social transfer payments. It is also associated with increased mortality 1 yr after surgery.

There are at least three types of POCD. Older nondemented adults with POCD can have isolated difficulties in learning/memory functions, in executive functions, or in a combination of memory and executive functions. Executive and combined impairments have been associated with functional limitations. To date, there have been no prospective investigations examining which cognitive/memory indices may best identify these POCD types, or whether there are predictive neuroanatomical risk factors for these

What We Already Know about This Topic

• Total knee arthroplasty improves knee function and physical activities in many patients but is associated with temporary or permanent cognitive dysfunction in some older adults
• Whether presurgery structural brain information would add to predictive models for this dysfunction is unknown

What This Article Tells Us That Is New

• In an exploratory, pilot study, memory and executive dysfunction occurred, but only brain markers of vascular disease associated with executive decline

ABSTRACT

Background: Total knee arthroplasty improves quality of life but is associated with postoperative cognitive dysfunction in older adults. This prospective longitudinal pilot study with a parallel control group tested the hypotheses that (1) nondemented adults would exhibit primary memory and executive difficulties after total knee arthroplasty, and (2) reduced preoperative hippocampus/entorhinal volume would predict postoperative memory change, whereas preoperative leukoaraiosis and lacunae volumes would predict postoperative executive dysfunction.

Methods: Surgery (n = 40) and age–education-matched controls with osteoarthritis (n = 15) completed pre- and postoperative (3 weeks, 3 months, and 1 yr) memory and cognitive testing. Hypothesized brain regions of interest were measured in patients completing preoperative magnetic resonance scans (surgery, n = 31; control, n = 12). Analyses used reliable change methods to identify the frequency of cognitive change at each time point.

Results: The incidence of postoperative memory difficulties was shown with delay test indices (i.e., story memory test: 3 weeks = 17%, 3 months = 25%, 1 yr = 9%). Postoperative executive difficulty with measures of inhibitory function (i.e., Stroop Color Word: 3 weeks = 21%, 3 months = 22%, 1 yr = 9%). Hierarchical regression analysis assessing the predictive interaction of group (surgery, control) and preoperative neuroanatomical structures on decline showed that greater preoperative volumes of leukoaraiosis/lacunae were significantly contributed to postoperative executive (inhibitory) declines.

Conclusions: This pilot study suggests that executive and memory declines occur in nondemented adults undergoing orthopedic surgery. Severity of preoperative cerebrovascular disease may be relevant for understanding executive decline, in particular. (ANESTHESIOLOGY 2014; 120:601-13)
imperfections. Such studies are necessary because at present there are no known mechanisms of POCD.

Specific preoperative neuroimaging markers may indicate risk for POCD. It is well established that the entorhinal cortex (ERC) and hippocampus are important in declarative memory processes6,7 as measured by delay memory tests.8 These structures change with Alzheimer disease.9 ERC volume predicts conversion to Alzheimer disease, with smaller volumes having greater conversion rates.10–12 Because of its involvement in the limbic–hypothalamic–pituitary–adrenal axis, the hippocampus is vulnerable to neuronal degeneration after severe biological/psychological stress.13,14 Posttraumatic stress disorder associated with reduced hippocampal volumes,15,16 and hippocampal degeneration also occurs in rats after anoxia and mild hypothermia.17 Preoperative leukoaraiosis and lacunae volume, by contrast, may indicate vulnerability to postoperative executive decline. Leukoaraiosis involves hypointense white matter regions on brain computed tomography/magnetic resonance images (MRIs)18 and occurs in 15 to 65% of adults.19 Leukoaraiosis was associated with demyelination and hyalinosis narrowing of small brain arterioles. This signifies microvascular burden to frontal-subcortical white matter pathways important for executive functions.20–22 Lacunae suggest ischemic stroke (often considered “silent”). Lacunae often occur in subcortical gray matter nuclei (e.g., thalamus, caudate) necessary for filtering/disenengaging attention.23 Leukoaraiosis and lacunae mark chronic small brain vessel disease24 and are contributors to insidious executive dysfunction.25

For this pilot investigation, we used a comprehensive neuropsychological protocol to assess whether the learning of new information (memory) and inhibitory functions (executive function) would be dominant forms of postoperative cognitive impairment.1 We then examined the hypothesis that specific presurgical neuroanatomical markers of early disease states (i.e., MRI-based hippocampus/ERC, leukoaraiosis/lacunae volumes) would differentially predict pre–postoperative memory and executive changes. Although individuals may not present with clinical signs of impairment preoperatively, we hypothesized that preoperative neuroimaging markers might serve as an indication of brain vulnerability to perioperative insult and resulting memory/executive decline. We secondarily examined intraoperative variables (e.g., emboli number, anesthesia duration) as contributors to decline.

Materials and Methods

Standard Protocol Approvals, Registrations, and Patient Consents

The University of Florida Institutional Review Board, Gainesville, Florida, approved this study, and all participants signed consents. Authors followed principles from the Declaration of Helsinki.

Participants

Total knee arthroplasty (TKA) and control participants were recruited through University of Florida orthopedic clinics, screened, and enrolled between the years of 2003 and 2005. Participants in the control group were selected from patients who had chosen to abstain from surgery for at least 1 yr. These two groups were recruited during the same time frame and were tested and scanned at the same time intervals. Participants had to meet the following inclusion/exclusion criteria: (1) aged 60 or older, (2) have English as a first language, (3) have osteoarthritis, (4) have intact activities of daily living, and (5) have baseline neuropsychological testing unsupportive for dementia criteria per Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition. Additional exclusion criteria included: another major surgery within the study timeline, history of head trauma/neurodegenerative illness, documented learning or seizure disorder, substance abuse in the last year, major cardiac disease, chronic medical illness known to induce encephalopathy, implantable device precluding an MRI, and an unwillingness to complete repeat testing. Two neuropsychologists reviewed the baseline data to confirm that test scores met the expected ranges for nondemented individuals.

Procedures

This was a prospective longitudinal study with TKA surgery and nonsurgery control groups. Participants completed a battery of tests that included a comprehensive history/systems interview, a comorbidity rating,26 activities of daily living,27 and neurocognitive and mood testing. Baseline MRI was performed on eligible participants. Cognitive and mood status were re-evaluated at 3 weeks, 3 months, and 1 yr postbaseline (postoperative for the TKA group). Delirium was ruled out postoperatively.28 Cognition was assessed 2 h after pain medication was given.

Anesthesia and Surgery Protocols

Protocols were standardized, with intravenous midazolam for anxiety and fentanyl, thiopental, and rocuronium for anesthesia induction and intubation. Patients were ventilated with an air/oxygen mixture to maintain an end-tidal carbon dioxide at 35 ± 5 mm; anesthesia was maintained with inhaled isoflurane and intravenous fentanyl and rocuronium. We attempted to record emboli incidents with transcranial Doppler probes over the transtemporal window placed by the same anesthesiologist and technician.29 The same surgeon and anesthesiaology teams completed all surgeries.

Neuropsychology Assessment

Baseline measures taken were within 1 week of the brain MRI. Baseline general cognitive30 and intellectual estimates31 helped equate groups. Measures of semantic/language fluency, visuo perception, and motor function were added to the protocol to examine the hypothesis that POCD primarily involved changes in memory and executive measures. Multiple measures were administered as part of the investigation to examine instruments and expected sensitivity to postoperative change. The same neuropsychologist administered all measures at each time point. Tests were scored by...
the neuropsychologist but were also rescored and double-data entered by two technicians who were blinded to group status. Alternate test forms were used, when applicable, with administration order based on a random generation list (see the study by Lezak et al.32 for test citations/descriptions). Primary neuropsychological measures of interest were theoretically grouped:

Memory: Auditory and visual stimuli with immediate and delay indices. Alternate versions were randomly administered for each period of testing.

1. Hopkins Verbal Learning Test—Revised: Twelve-word list-learning test with three immediate free recall learning trials, one 20-min delay free recall and one recognition trial. Dependent variables (DVs): Immediate and delay total correct.
2. Story Memory Test33: Paragraph recall test with immediate and 30-min delay recall indices. DVs: Immediate and delay total correct.
3. Brief Visuospatial Memory Test—Revised: Geometric figural memory with three learning trials, one 20-min delay free recall and one recognition trial. DVs: Immediate and delay total correct.

Executive Functions: Tests associated with working memory, word fluency, processing speed, and disengagement—response inhibition were applied:

1. Digit Span Backward Subtest from the Wechsler Memory Scale—Third Edition: Requires listening to an increasing digit series and then repeating those digits in forward and backward sequences. DV: Total backward.
4. Controlled Oral Word Association Test: Generating words beginning with a specific letter within 60 s, excluding numbers and proper nouns. DV: Total words minus errors.
5. Stroop Color-Word Test: Involves selective attention and cognitive control by expecting participants to suppress the automatic tendency to read aloud words rather than name the color of the ink in which the words are printed on the page. DV: Color-word score in 45 s.

Other Cognitive Domains: Semantic/language fluency, perceptual-spatial function, and motor measures were administered to better examine the hypothesis that memory and executive dysfunctions were primary POCD forms.

1. Category (Animal) Fluency: Requires generating animal names in 60 s. DV: Total words generated.
2. Judgment of Line Orientation: Involves matching two lines of varying degree to a spectrum of lines ranging from 0 to 180 degrees. DV = Total correct.
3. Finger Tapping Test: Involves rapid lever pressing for 10 s and 10 trials. DVs = Average dominant and nondominant taps across all trials.

Mood and Pain
Depression and anxiety were evaluated with the Geriatric Depression Scale34 and the State Trait Anxiety Inventory.35 A visual analog scale gauged pain severity.36

Neuroimaging
Total knee arthroplasty participants were scanned within 2 weeks of surgery, and control participants were scanned at their baseline assessment by using a Siemens 3T Allegra scanner (Erlangen, Germany). T1-weighted, three-dimensional magnetization prepared rapid acquisition gradient-echo sequence (repetition time = 2,500 ms, echo time = 4.38 ms, inversion time = 1,100 ms, flip angle = 8 degrees, matrix = 256 x 144) reconfigured to 160 gapless, 1-mm images provided gray and white matter segmentation. Leukoaraiosis volumetrics were acquired from two-dimensional fluid-attenuated inversion recovery sequences (repetition time range across scans = 8,402 to 12,800; echo time range = 125 to 147 ms; inversion time range = 1,800 to 2,200 ms, flip angle = 90 degrees, gap = 5 to 7 mm).

Magnetic Resonance Imaging Predictor Variables
Raters were blind to participant group. Hippocampi37 were manually segmented by using ITK-SNAP38 (www.itksnap.org) by one reliable rater with an excellent Dice Similarity Coefficient39 (grand Dice Similarity Coefficient = 80 ± 0.2; intrarater: grand Dice Similarity Coefficient = 0.81 ± 0.05; Pearson r range = 0.75 to 0.83; all P < 0.001) (Volume DVs = left, right, and total hippocampal volumes in cubed millimeters).

Entorhinal cortices40 were segmented by highly reliable raters (Intraclass Correlation Coefficient r >0.9333 using MEASURE).41 Tracings were made on oblique coronal slices with volumes calculated by compiling individual slice measurements (DV = left, right, and total hippocampal volumes in cubed millimeters).

Leukoaraiosis was measured by a reliable rater (Dice Similarity Coefficient interrater range of 0.84 to 0.93, intrarater range >0.99) by using ImageJ42 (http://imagej.nih.gov/ij/) and an in-house macro.43 By using published methods,44 leukoaraiosis voxels were thresholded and created into two-dimensional leukoaraiosis binary masks which were subsequently concatenated to form a single, three-dimensional binary mask for each brain (DV = total volume in cubed millimeters).

Lacunae were measured by a neuroradiologist with the use of volumetric T1 scans. Only well-defined, dark lesions with a diameter of 2 mm or greater that held a stationary
position between slices were graded as lacunae with their volume estimated using the formula of a sphere \((\frac{4}{3}\pi r^3)\) \(\delta\) (DV = total volume in cubed millimeters).

Imaging control variables included total brain volume corrected for group and total intracranial volume\(^45\) to correct for head size and age-related atrophy. Intracranial volume (brain plus cerebral spinal fluid) and supratentorial whole brain volume (gray and white tissue plus ventricular cerebrospinal fluid minus brainstem and cerebellum) were segmented by using Functional Magnetic Resonance Imaging of the Brain Software Library and BrainSuite methods.\(^46,47\)

**Statistical Analysis**

A power analysis addressed the primary hypothesis that a baseline neuroanatomical variable of interest would correlate to postoperative cognitive change. Given the pilot nature of the study and that we did not have prior data from which to draw upon to estimate the effect size, we used STPLAN software (1993, University of Texas M.D. Anderson Cancer Center, Houston, TX) and a Fisher Z transformation method for sample estimation. On the basis of a 0.05 level of significance, a power of 0.80, and a moderate correlation \(r = 0.32\), we expected to enroll 60 participants total for the investigation. All analyses were completed with SPSS, version 21.0 (IBM, New York, NY). The level of significance was set at 0.05.

**Assessing POCD Frequency**

Independent \(t\)- and chi-square tests examined group differences on baseline demographics (e.g., imaging, cognition, mood, and pain variables). A modification of Jacobson and Truax Reliable Change Index (see studies by Rasmussen et al.,\(^3\) Monk et al.,\(^5\) Lewis et al.,\(^48\) and Lewis M et al.\(^49\)) assessed of POCD by test: \(z\) scores

\[
\frac{\Delta X - \Delta X_c}{SD(\Delta X)}.
\]

Change was calculated by subtracting preoperative from the postoperative performance (\(\Delta X\)). The averaged control group change (\(\Delta X_c\), which was assumed to represent systematic error) was then subtracted from the individual change, with this value then divided by the control group’s SD of the change (SD\(\Delta X_c\)). Abnormal cognitive decline was a \(z\)-score \(\leq 1.96\). By using this method, we then examined tests for false positives, which is a consideration for POCD test sensitivity.\(^49\) Chi-square analyses assessed differences in POCD frequency. An overall doubly repeated longitudinal analysis was performed using the MIXED procedure in SAS, version 9.2 (Cary, NC). The repeated measures were \(z\)-scores observed over time (baseline, 2 weeks, 3 months, and 1 yr) and cognitive domain: (1) immediate learning and memory (Hopkins Verbal Learning Test I, Story I, and Brief Visuospatial Memory Test (BVMT) 1), (2) delayed learning and memory (Hopkins Verbal Learning Test D, Story D, and BVMT D), (3) attention, processing, and executive function (Dg Span, Sp Span, Dg Symbol, and Stroop C-W), and (4) language, visuospatial, and motor (animals, judgment of line orientation, finger D, and finger tapping nondominant hand). Norms were not available to form controlled oral word association \(z\)-scores thus controlled oral word association was omitted from this analysis.

**Predicting Cognitive Change**

Multivariate analyses were strategic to (1) confirm that there were certain domains that would change with surgery (we expected primary changes in memory and executive function); and (2) investigate the expected sensitivity of certain instruments to detect these changes (i.e., memory test delay indices word list memory measures \(v_s\), story paragraph memory measures; story memory measures are considered less dependent upon the processing speed).\(^32\) We limited the multivariate analyses to one memory and one executive measure, with this selection-based demonstration of change in the surgery group relative to the nonsurgery group, and knowledge of which tests may be most sensitivity to anterograde memory changes (e.g., delay index) and inhibitory decline. We did not consider using a composite score approach as we believed that would give an inaccurate assessment of POCD type in our sample, for less sensitive tests have the potential to introduce measurement error. Given the growing field of POCD research and the numerous concerns regarding which tests should be administered, we include information for the separate test measures within the provided tables. We believe that this information is relevant for guiding future investigations. Planned hierarchical regressions examined (1) hypothesized interaction of baseline hippocampal, ERC volumes, and group (surgery, control) on the delay index of the Story Memory Test, and (2) the baseline interaction of lacunae/leukoaraiosis volume and group (surgery, control) on the Stroop Color-Word Test Color-Word condition score. The first regression model always included the baseline cognitive variable of interest (e.g., baseline delay memory score) and the imaging control variable of total brain volume corrected. Including the preoperative cognitive score in the first regression model renders the DV into a “residual change score,” so that the effects of the predictors on the DV may be interpreted as predictors of change. Covariates of education, anesthesia duration, emboli count, and TKA type were analyzed and retained in the model if they were found to be significant. The second regression model included group type (surgery, control), and the third model included the interaction variable of interest (i.e., centered variable\(^50\) of group \(\times\) hippocampal volume).

**Results**

Total knee arthroplasty patients \((n = 40)\) and controls \((n = 15)\) were similar regarding general demographic variables, general cognitive status, mood, and baseline pain (table 1; all \(P > 0.05\)). Although not significant, the control group had on average 2 more years of education and were eight points higher on an abbreviated intellectual estimate.
All were considered healthy, with low comorbidity, and there were no statistically significant group baseline differences on the memory and cognitive measures (table 2; all P > 0.05). Emboli counts were acquired only on a subset of the surgery participants (n = 15; mean = 14.40; SD = 25.63; range = 1 to 100 emboli) because of difficulties maintaining transcranial Doppler placement throughout surgery. Surgeries included unilateral TKA (n = 28) and bilateral TKA (n = 12).

A subset of individuals was unable to complete the preoperative brain scanning because of new onset claustrophobia (n = 4), the size of the scanner bore, which limited patients with larger chests (n = 5), and poor image quality (n = 3). Because of pilot study timeline enrollment limitations, the final subgroup completing preoperative imaging included 31 surgery individuals and 12 control individuals who were similar regarding demographics, general cognitive status, mood, pain variables (table 1; P > 0.05), baseline neuropsychology...
variables (table 2; \( P > 0.05 \)), and baseline brain variables of interest (table 3; all \( P > 0.05 \)). This MRI subset of controls had approximately 2 more years in education, on average, and scored six points higher on the intellectual estimate. Emboli were acquired within a subset (\( n = 12; \) mean = 14.50; SD = 28.45; range = 1 to 100 emboli). This MRI subgroup also included unilateral (\( n = 20 \)) and bilateral (\( n = 11 \)) TKAs.

The rate of attrition for TKA patients was 0% at 3 weeks, 8% (3 of 40) at 3 months, and 15% (6 of 40) at 1 yr. Two patients were unavailable at the 3-month time point but were tested at 1 yr. Postbaseline testing was completed at 22±7 days, 3 months ± 29 days, and 1 yr ± 81 days.

### Table 3. Group Mean (M) and SD for Imaging Variables (mm³) with Group Comparison SMD for Effect-size Assessment

<table>
<thead>
<tr>
<th>Variable</th>
<th>Surgery M</th>
<th>SD</th>
<th>Control M</th>
<th>SD</th>
<th>SMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hcmps left</td>
<td>2,465.31</td>
<td>435.17</td>
<td>2,531.30</td>
<td>540.34</td>
<td>−0.14</td>
</tr>
<tr>
<td>Hcmps right</td>
<td>2,607.87</td>
<td>477.38</td>
<td>2,647.45</td>
<td>390.88</td>
<td>−0.09</td>
</tr>
<tr>
<td>Hcmps total</td>
<td>5,073.18</td>
<td>843.41</td>
<td>5,178.75</td>
<td>856.08</td>
<td>−0.12</td>
</tr>
<tr>
<td>Entorhinal left</td>
<td>1,027.03</td>
<td>556.70</td>
<td>899.65</td>
<td>592.98</td>
<td>0.23</td>
</tr>
<tr>
<td>Entorhinal right</td>
<td>1,056.86</td>
<td>513.26</td>
<td>957.78</td>
<td>543.63</td>
<td>0.19</td>
</tr>
<tr>
<td>Entorhinal total</td>
<td>2,083.89</td>
<td>1,000.27</td>
<td>1,857.43</td>
<td>1,074.13</td>
<td>0.22</td>
</tr>
<tr>
<td>Total LA (raw)</td>
<td>8,182.07</td>
<td>9,570.42</td>
<td>9,005.90</td>
<td>7,372.35</td>
<td>−0.09</td>
</tr>
<tr>
<td>Infarct volume</td>
<td>126.05</td>
<td>126.38</td>
<td>142.33</td>
<td>246.66</td>
<td>−0.10</td>
</tr>
<tr>
<td>TBV</td>
<td>1,264,966.30</td>
<td>125,009.25</td>
<td>1,234,972.1</td>
<td>120,552.73</td>
<td>0.17</td>
</tr>
<tr>
<td>TICV</td>
<td>1,458,260.04</td>
<td>140,704.89</td>
<td>1,427,875.14</td>
<td>126,951.68</td>
<td>0.22</td>
</tr>
<tr>
<td>TBVc</td>
<td>1,292,148.39</td>
<td>261,510.06</td>
<td>1,234,132.01</td>
<td>233,736.17</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Group differences did not reach statistical significance.

Entorhinal = entorhinal cortex (note: surgery \( n = 29 \), control \( n = 9 \)); Hcmps = Hippocampus (note: surgery \( n = 30 \), control \( n = 12 \)); LA = leukoaraiosis; SMD = standardized mean difference; TBV = total brain volume; TBVc = total brain volume (corrected) ((individual intracranial volume/group mean intracranial volume) × individual total brain volume); TICV = total intracranial volume; all volumes from native space.

### Table 4. Surgery and Control Group Percent Reliably Declined (≥1.96) by Test and Time Period

<table>
<thead>
<tr>
<th>Domain</th>
<th>Test</th>
<th>Surgery 3 Weeks</th>
<th>Surgery 3 Months</th>
<th>Surgery 1 Yr</th>
<th>Control 3 Weeks</th>
<th>Control 3 Months</th>
<th>Control 1 Yr</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning and memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HVLT I</td>
<td>0.0%</td>
<td>(0/40)</td>
<td>7.9% (3/38)</td>
<td>8.8% (3/34)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>HVLT D</td>
<td>7.5%</td>
<td>(3/40)</td>
<td>5.3% (2/38)</td>
<td>14.7% (5/34)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>Story I</td>
<td>0.0%</td>
<td>(0/37)</td>
<td>0.0% (0/34)</td>
<td>0.0% (0/33)</td>
<td>9.1% (1/11)</td>
<td>0.0% (0/10)</td>
<td>0.0% (0/13)</td>
</tr>
<tr>
<td>Story D</td>
<td>16.7%</td>
<td>(6/36)</td>
<td>25.0% (8/32)</td>
<td>9.4% (3/32)</td>
<td>9.1% (1/11)</td>
<td>0.0% (0/10)</td>
<td>0.0% (0/12)</td>
</tr>
<tr>
<td>BVMT I</td>
<td>2.5%</td>
<td>(1/40)</td>
<td>0.0% (0/38)</td>
<td>0.0% (0/34)</td>
<td>6.7% (1/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>BVMT D</td>
<td>5.0%</td>
<td>(2/40)</td>
<td>13.2% (5/38)</td>
<td>14.7% (5/34)</td>
<td>13.3% (2/15)</td>
<td>13.3% (2/15)</td>
<td>6.7% (1/15)</td>
</tr>
<tr>
<td>Attention, processing, and Exec. Fx.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DgSpan</td>
<td>0.0%</td>
<td>(0/40)</td>
<td>7.9% (3/38)</td>
<td>8.8% (3/34)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>SpSpan</td>
<td>0.0%</td>
<td>(0/39)</td>
<td>0.0% (0/37)</td>
<td>0.0% (0/33)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>COWA</td>
<td>5.0%</td>
<td>(2/40)</td>
<td>5.3% (2/38)</td>
<td>6.1% (3/33)</td>
<td>6.7% (1/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>DgSymbol</td>
<td>2.6%</td>
<td>(1/38)</td>
<td>11.1% (4/36)</td>
<td>15.2% (5/33)</td>
<td>7.7% (1/13)</td>
<td>0.0% (0/13)</td>
<td>7.7% (1/13)</td>
</tr>
<tr>
<td>Stroop C-W</td>
<td>21.1%</td>
<td>(8/38)</td>
<td>22.2% (8/36)</td>
<td>9.1% (3/33)</td>
<td>7.7% (1/13)</td>
<td>0.0% (0/13)</td>
<td>7.7% (1/13)</td>
</tr>
<tr>
<td>Other language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Animals</td>
<td>5.0%</td>
<td>(2/40)</td>
<td>0.0% (0/38)</td>
<td>2.9% (1/34)</td>
<td>0.0% (0/15)</td>
<td>6.7% (1/15)</td>
<td>6.7% (1/15)</td>
</tr>
<tr>
<td>JLO</td>
<td>2.6%</td>
<td>(1/38)</td>
<td>2.3% (1/36)</td>
<td>3.0% (1/33)</td>
<td>6.7% (1/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>Visuospatial motor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finger D</td>
<td>0.0%</td>
<td>(0/37)</td>
<td>3.1% (1/32)</td>
<td>0.0% (0/32)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
<td>0.0% (0/15)</td>
</tr>
<tr>
<td>Finger NDom</td>
<td>8.1%</td>
<td>(3/37)</td>
<td>3.1 (1/32)</td>
<td>6.3% (2/32)</td>
<td>6.7% (1/15)</td>
<td>0.0% (0/15)</td>
<td>6.7% (1/15)</td>
</tr>
</tbody>
</table>

Denominators vary between time periods due to drop-out or missing data (e.g., corrupt cassette used to record verbatim measurements for the Story Memory Test, broken lever on finger tapper at time of testing).

Animals = animal fluency; BVMT = Brief Visuospatial Memory Test Immediate Total/Delay Total; COWA = Controlled Oral Word Association Test; DgSpan = Wechsler Adult Intelligence Scale Third Edition Digit Span backward; DgSymbol = Wechsler Adult Intelligence Scale Third Edition Digit Symbol subtest; Exec. Fx. = executive functioning Finger Non-Dom = Finger Tapping Non-Dominant; HVLT = Hopkins Verbal Learning Test-Revised Immediate/Delay; JLO = Judgment of Line Orientation Test total; SpSpan = WMS-III Spatial Span backward; Story = Story Memory Test Immediate/Delay; Stroop C-W = Stroop Color-Word Score.

### Frequency of Cognitive Decline

**Learning and Memory.** Postoperative cognitive dysfunction rates were more frequent for TKA patients on the delay relative to immediate memory indices (\( X^2 (1) = 5.98; P = 0.01 \)). The highest POCD rate involved the Story Memory Test delay and the lowest POCD rate involved the Hopkins Verbal Learning Test delay (3 weeks: 17 and 8%; 3 months: 25 and 5%; and 1 yr: 9 and 15%, respectively) with test comparisons at 3 weeks and 3 months, \( P < 0.05 \). The visual memory test was accompanied by equally high rates of impairment in the control group at the 3-week and 3-month time
points (both \( P > 0.05 \)), suggesting high false-positive rates (table 4).

**Executive Functions.** For TKA patients, the highest rates of POCD involved in the inhibitory subtest of the Stroop Color-Word Test, with 21, 22, and 9% at 3 weeks, 3 months, and 1 yr, respectively. The test rate comparison at 3 weeks and 3 months was all nonsignificant.

**Overall Analysis.** The TKA patients had lower z-scores in an overall analysis of time and cognitive domain. The regression coefficient in the doubly repeated model was \(-0.13\) \([t(2,893); P = 0.003]\).

**Neuroanatomical Predictors of Cognitive Decline**

The two tests with the highest rates of POCD and minimal false positives in the control group were examined as outcome markers of memory and executive function (*Memory Test: Story Memory Test Delay Index; Executive Function Test: Stroop Color-Word Test and Color-Word condition*).

**Preoperative Hippocampal/ERC Volumes and Story Memory Test Performance.** Baseline story memory ability was a predictor for postoperative story performance for each time period (\( \beta \)'s: 0.51 to 0.71; all \( P < 0.01 \)). Adding in group status (surgery or control) to the model negatively contributed to memory performance (\( \beta \)'s: \(-0.42\) to \(-0.25\), 3-week and 3-month \( P \) values \( \leq 0.001 \), 1-yr \( P \) value <0.05). The interaction variable of group (surgery or control) by left hippocampus or ERC volume never significantly contributed to the model over that of the other variables. There were no significant findings for the right hippocampus or ERC (tables 5 and 6).

**Preoperative Leukoaraiosis and Lacunae Volumes and Stroop Color-Word Test Performance.** Baseline Stroop performance and total brain volume corrected were significant independent predictors of postoperative Stroop performance for each time period (\( \beta \)'s: 0.76 to 0.82, all \( P \) values <0.001 and \( \beta = -0.23, P = 0.03 \), respectively). Adding group status alone (surgery or control) to the model was not a significant predictor for any time period (all \( P \) values >0.11). Adding in the interaction variable of group (surgery or control) and leukoaraiosis/lacunae volume significantly improved prediction of executive change for 3 weeks (\( \beta = -0.22, P = 0.027 \); adjusted \( R^2 = 0.66, F \text{ change} = 5.31 \) and 1 yr (\( \beta = -0.27, P = 0.01 \), adjusted \( R^2 = 0.68, F \text{ change} = 8.40 \), with a trend for

---

**Table 5.** Hierarchical Regression Analysis Summary for Left Hippocampal Volume and Story Memory Performance at 3-weeks, 3-months, and 1-yr Posttotal Knee Surgery

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>3 Weeks ( R^2 )</th>
<th>3 Months ( R^2 )</th>
<th>1 Yr ( R^2 )</th>
<th>( R^2/R^2\Delta ) Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline SM score</td>
<td>0.51**</td>
<td>0.61**</td>
<td>0.70**</td>
<td></td>
</tr>
<tr>
<td>TBVc</td>
<td>(-0.19)</td>
<td>0.04</td>
<td>(-0.15)</td>
<td>( R^2 = 0.33–0.53 )</td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>(-0.40**)</td>
<td>(-0.42**)</td>
<td>(-0.25**)</td>
<td>( R^2\Delta = 0.06–0.18 )</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L Hcmp × Group</td>
<td>(-0.18)</td>
<td>(-0.17)</td>
<td>0.07</td>
<td>( R^2\Delta = 0.00–0.03 )</td>
</tr>
</tbody>
</table>

**P < 0.001.**

Group = centered group variable (surgery, nonsurgery classification); L Hcmp × Group = interaction variable of Centered Left Hippocampus and Centered Group (TKA surgery/nonsurgery), right sided structures and total volumes (left plus right) showed a similar pattern; SM = story memory; TBVc = total brain volume corrected for intracranial volume; TKA = total knee arthroplasty.

**Table 6.** Hierarchical Regression Analysis Summary for Preoperative Left Entorhinal Cortex Volume on Story Memory Performance at 3-weeks, 3-months, and 1-yr Posttotal Knee Surgery

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>3 Weeks ( R^2 )</th>
<th>3 Months ( R^2 )</th>
<th>1 Yr ( R^2 )</th>
<th>( R^2/R^2\Delta ) Ranges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline SM score</td>
<td>0.51**</td>
<td>0.61**</td>
<td>0.70**</td>
<td></td>
</tr>
<tr>
<td>TBVc</td>
<td>(-0.09)</td>
<td>0.06</td>
<td>(-0.15)</td>
<td>( R^2 = 0.23–0.50 )</td>
</tr>
<tr>
<td>Step 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>(-0.40**)</td>
<td>(-0.42**)</td>
<td>(-0.33**)</td>
<td>( R^2\Delta = 0.10–0.17 )</td>
</tr>
<tr>
<td>Step 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L ERC × Group</td>
<td>(-0.18)</td>
<td>(-0.17)</td>
<td>0.07</td>
<td>( R^2\Delta = 0.00–0.03 )</td>
</tr>
</tbody>
</table>

**P < 0.001.**

Group = centered group variable (surgery, nonsurgery classification); L ERC × Group = interaction variable of left entorhinal cortex and group (TKA surgery/nonsurgery); right sided structures and total ERC showed a similar pattern; TBVc = total brain volume corrected for intracranial volume; TKA = total knee arthroplasty.
contribution at 3 months ($\beta = -0.17, P = 0.08$). Post hoc analyses on preoperative leukoaraiosis and lacunae volumes on postoperative memory changes were not found to be significant (table 7).

**Education and Surgical Considerations**

Education, anesthesia duration (mean ± SD = 225.74 ± 99.39 min), and TKA type (unilateral/bilateral) were examined as covariates but did not contribute significantly to the first step of the regression models (Stroop Color-Word Test: education $\beta$'s = −0.03 to 0.01; anesthesia $\beta$'s = −0.09 to 0.02; TKA type $\beta$'s = −0.11 to −0.26; Story Memory Performance: education $\beta$'s = 0.17 to 0.29; anesthesia $\beta$'s = −0.18 to 0.06; TKA type $\beta$'s = −0.06 to 0.22) and were therefore not retained in the models. In the subsample of surgery participants with emboli measurement, greater emboli number negatively contributed to the acute 3-week postoperative Stroop Color-Word performance ($n = 15$; Stroop Color-Word 3-week $\beta = -0.52, P = 0.03$; all other time point $\beta$'s <−0.16), but not significantly to delay Story Memory Test performance at any time period ($\beta = -0.05$ to −0.27). Although not statistically significant, emboli number was higher in bilateral ($n = 6$; mean ± SD = 24.0 ± 39.06 emboli) than unilateral TKA ($n = 9$; mean ± SD = 8.0 ± 8.97 emboli).

**Discussion**

This is the first prospective pilot study examining the role of presurgical neuroanatomical factors on POCD type after TKA in nondemented older adults. Although we acknowledge the pilot nature of the study, our data suggest that memory and executive declines were the primary forms of cognitive change at 3 weeks post-TKA. Five percent or less of the patients exhibited declines in language, perceptual-spatial, or frontal motor function measures. The pilot study found limited value for using presurgery ERC/hippocampal volumes as neuroanatomical predictors for POCD memory decline at any time point (3 weeks, 3 months, or 1 yr). In contrast, preoperative neuroimaging evidence of microvascular disease (preoperative leukoaraiosis and lacune volume) explained a portion of executive function decline at the 3-week and 1-yr postoperative sessions, with a trend at 3-month postoperation. We encourage researchers to conduct similar but larger-scale prospective studies. Before clinical change can occur, and risk versus surgical benefit can be weighed, we need more definitive evidence regarding the nature of baseline microvascular disease and executive decline, as well as rigorous examinations on neuroanatomical contributors to postoperative memory decline.

**Considerations for Neuropsychological Measures and Potential Perioperative Variables**

**Executive Decline.** Leukoaraiosis and lacunae volume accounted for a significant portion of variance on a well-known measure of executive function and the interference condition of the Stroop Color-Word Test at the 3-week and 1-yr postoperative intervals. Other frontal system tests (i.e., Digit Symbol subtest, which involves processing speed and visual–motor integration) revealed postoperative change, but not to the same extent as the Stroop Color-Word Test. The Stroop Color-Word Test has been sensitive to delirium in similar patient groups. Via functional neuroimaging, this test associated with dorsal and medial frontal lobe, and parietal lobe, thereby implicating the involvement of large frontal-parietal and frontal-subcortical white matter networks. Leukoaraiosis and lacunae disrupt these connections. Thus, the Stroop Color-Word Test warrants consideration as a key neuropsychological measure in future POCD investigations.

Pilot findings extend upon studies showing leukoaraiosis as a risk factor for high-risk cardiac patients. Leukoaraiosis predicts frequency of POCD 3 months after “on-pump” coronary artery bypass surgery, and is associated with cerebral ischemic events after coronary artery bypass grafting. Our pilot study is the first to suggest that leukoaraiosis and lacunae volume should be considered for healthy nondemented older adults without clinically significant atherosclerotic disease who elect orthopedic surgery.
The control group with similar leukoaraiosis and lacunar burden load allowed us to examine the interaction of baseline microvascular disease and group (surgery/control) status. Findings suggest that the preoperative injuries of leukoaraiosis and lacunae did not contribute significantly to executive dysfunction until "insult (perioperative events) was added to injury." A larger study needs to replicate these findings. We also need to identify the perioperative events that negatively interact with presurgery evidence of microvascular disease.

Potential perioperative events that may negatively interact with preoperative leukoaraiosis–lacunae volume include acute anemia,60–63 hypotension,64 oxygen desaturation,65–67 and the production of emboli. In our sample, we acquired valid emboli data on 15 of our surgery participants. The number of emboli in our sample (1 to 100 emboli) is similar to that in the published reports.68 Post hoc findings suggest an association between emboli and acute (3 weeks) Stroop change score, but did not contribute to the predictive model. Although there are more reports that emboli frequency do not relate to cognitive change69–71 than otherwise,72 study sample neurovascular burden factors and the interaction with emboli remain poorly understood. Small emboli may lodge in regions with low blood flow,73 and regions of the brain with leukoaraiosis might represent such regions. Support for this postulate comes from the observation that leukoaraiosis contributes to the development of silent stroke after cardiac surgery.74 We did not identify any significant predictors involving the duration of anesthesia, but the interaction with these and other baseline microvascular diseases warrants future examination.

The transient nature of leukoaraiosis and lacunae volume on cognition across our three postoperative time points suggests that perioperative events around TKA may be associated with ischemia. Ischemia, unlike infarction, is potentially reversible. If infarction was the mechanism of the decline, we would have expected that many participants would not have had a full cognitive reversal. The question of whether transient decline might also be related to changes in other systems (e.g., neurotransmitter systems) needs empirical assessment. Transient decline is a well-established occurrence in cardiac and noncardiac surgery.3,75–78

**Memory Decline.** Memory functions have been classically associated with three neuroanatomic regions in the brain (and the pathways that interconnect them). These include the medial temporal lobe (hippocampus and ERC),33,79 the thalamus (dorsomedial and anterior nuclei),80 and the basal forebrain, which innervates the hippocampus with essential cholinergic neurons.81 In older adults, these regions are most associated with the amnestic form of mild cognitive impairment or clinical presence of Alzheimer disease.11,82,83 Despite the current study’s high incident rate of postoperative recall impairment on sensitive indices of memory function, preoperative ERC and hippocampal volumes did not significantly contribute to postoperative memory abilities.

We do not consider our inability to support the original ERC–hippocampus hypothesis as a consequence of the neuropsychological or imaging measurement approaches. Rather, our reliable change analyses showed a marked decline on a delay recall index that is a characteristic of individuals with amnestic disorder,84,85 and functional changes to the ERC.8 The greatest rate of decline was also observed on a story recall test, which more closely resembles the everyday memory demands for interpersonal discourse, radio and television programs, as well as material that has been read, such as newspapers.3 We used reliable neuroanatomical measurement approaches83 with resultant structure volumes that correspond to published values and ranges.40 We controlled for differences in intracranial and brain volume factors that may contribute to ERC/hippocampal variability between participants.86 Given these methodological strengths, we interpret that preoperative macrostructural measurement of the ERC and hippocampus is not sufficiently sensitive for predicting memory decline for nondemented otherwise “healthy” adults undergoing TKA surgery. In vivo molecular-based examinations of medial temporal integrity and associated regions via diffusion and/or spectroscopy methods appear warranted; changes in microstructure may be earlier markers of undergoing neurodegenerative disease before alterations in macrostructure such as changes in volume become manifest.87

**Study Considerations**

We recognize study limitations. The sample size is small (hence increasing the probability of a Type II error) and unequal within the surgery and control groups. Despite our best attempts at matching the groups on age and education, the control group on average had more years of education. Although we restricted our regression analysis on one neuropsychological index, we recognize that our multiple hypothesis testing increased the experiment-wise type I error rate. An additional limitation was that the neuropsychologist knew of the group (surgery, control) condition. We attempted to rectify this by having all tests rescored and re-entered by trained individuals who were blinded to the groups. Finally, this study was conducted with physically and cognitively well individuals. This limits the applicability of our pilot data to other populations, but does provide some reference for the volume of lacunae and leukoaraiosis in nondemented samples for future comparison purposes. We encourage similar studies on higher-risk patients such as those with metabolic syndrome who have shown high rates of POCD after noncardiac and cardiac surgery.88

Despite the limitations, the pilot study has design strengths. We targeted control group recruitment during the same time period as the surgery recruitment and attempted to identify individuals similar in age, medical comorbidity, and baseline brain variables of interest. Cognitive change was examined using the Reliable Change Index, which expressed change relative to the error estimated from a control group.
matched for age, comorbidity, intelligence, and baseline brain status. The Reliable Change Index has demonstrated adequate specificity to detect POCD in noncardiac samples, with results replicated across studies.\textsuperscript{2,3,48} We further examined our hypotheses regarding POCD types with tests known for their sensitivity in examining memory and executive function. Examining cognitive change over time with a nonsurgery control group allowed us to examine for potential false positives with specific cognitive measures.\textsuperscript{49} This approach was very useful given our small sample size, which prohibited a more powerful confirmatory factor analysis. Now, larger studies are necessary to re-examine our findings not only with the individual cognitive tests but also with specific cognitive composites. The neuroanatomical variables were examined relative to brain and intracranial volumes, a technique that reduces interindividual variability thereby clarifying cognitive–neuroanatomical associations.\textsuperscript{56}

The pilot findings warrant further consideration and larger empirical study. Investigators are encouraged to consider test protocols that include delay memory test conditions and executive inhibitory measures. In addition to standard structural measurements of leukoaraiosis and lacunae volume, future neuroimaging investigations should include molecular diffusion and functional-based sequences. Diffusion-weighted sequences allow quantification of tissue integrity before changes are seen on traditional clinical sequences.\textsuperscript{59} Researchers have shown that tensor-based algorithms have relevancy for understanding 1-yr survival after major cardiac and brain trauma,\textsuperscript{59,61} as well as delirium.\textsuperscript{62} Resting and functional task-based sequences can be incorporated for understanding neuronal network risk profiles. Overall, it is our expectation that similar hypothesis-driven investigations using sensitive neuropsychological tools combined and molecular- and functional-based tools in addition to standard clinical scans (i.e., fluid attenuated inversion recovery scans used to calculate leukoaraiosis) will improve our appreciation for presurgery neuroanatomical risk profiles and POCD types.

Acknowledgments

The authors are most grateful to the participants for their time. The authors appreciate and thank the insightful comments from the Anesthesiology reviewers and associated editors. The authors are also grateful for the research assistance by Illicita Shugarman, M.D., (College of Medicine, University of Florida, Gainesville, Florida), who helped with the procedural aspects of the investigation, and thank Ria van der Aa, M.S., Katherine Moyses, B.A., and Jacqueline Nicosia, B.S. (Departments of Clinical and Health Psychology or Anesthesiology at the University of Florida), who assisted with double and triple data entry and management. The authors give special thanks to Jenny Mace, M.D. (Department of Anesthesiology, University of Florida), a valued mentor and advocate for research on postoperative cognitive dysfunction.

Funding was provided by the National Institutes of Health F32 AG021363 from the National Institute on Aging, Bethesda, Maryland (to Dr. Price); The University of Florida Opportunity Fund (to Dr. Monk); NIH General Clinical Research Center Grant No. ULI TR000064; I. Heermann Anesthesia Foundation, Gainesville, Florida (to Drs. Monk and Price).

Competing Interests

The authors declare no competing interests.

Correspondence

Address correspondence to Dr. Price: Department of Clinical and Health Psychology, PO Box 100165, Gainesville, Florida 32610. cep23@phhp.ufl.edu. This article may be accessed for personal use at no charge through the Journal Web site, www.anesthesiology.org.

References


Anesthesiology 2014; 120:601-13

Price et al.

13. Sapolsky RM: Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. Arch Gen Psychiatry 2000; 57:925–55.


Anesthesiology 2014; 120:601-13

Price et al.


