Observational cohort study examining apolipoprotein E status and preoperative neuropsychological performance as predictors of post-operative delirium in an older elective arthroplasty population


Published in:
Age and Ageing

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© The Author 2017. Published by Oxford University Press on behalf of the British Geriatrics Society. All rights reserved. This work is made available online in accordance with the publisher’s policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.
Title: Observational cohort study examining apolipoprotein E status and preoperative neuropsychological performance as predictors of postoperative delirium in an older elective arthroplasty population

Authors:

Emma Louise Cunningham

Timothy Mawhinney

David Beverland

Seamus O’Brien

Daniel F McAuley

Rebecca Cairns

Peter Passmore

Bernadette McGuinness

1. Centre for Public Health, Institute of Clinical Sciences, Block B, Queen’s University Belfast, Royal Victoria Hospital, Belfast, BT12 6BA, UK

2. Belfast Health and Social Care Trust, Belfast, UK

3. Centre for Experimental Medicine, Queen’s University Belfast, Belfast, UK

4. Northern Health and Social Care Trust, Ballymena, UK

Corresponding author: Emma Louise Cunningham; email: ecunningham03@qub.ac.uk
Abstract:

Introduction

Delirium following surgery is common and is associated with negative outcomes. Preoperative cognitive impairment has been shown to be a risk factor for postoperative delirium. Often the cognitive tests used are cumbersome. This study tests the hypothesis that the quantification of brain vulnerability, using Apolipoprotein E (ApoE) status and neuropsychological tests, both traditional and more easily administered, can quantify the risk of postoperative delirium following elective primary arthroplasty surgery.

Methods

This observational cohort study recruited participants aged 65 years or older admitted prior to elective primary hip or knee arthroplasty. Baseline data was collected and participants underwent neuropsychological testing and had blood taken for ApoE genotyping preoperatively. Postoperatively participants were assessed daily for delirium using the Confusion Assessment Method (CAM) and charts were reviewed where possible for reports of delirium. Univariate and multivariate analyses of preoperative factors were undertaken to identify independent predictors of delirium.

Results

Between March 2012 and October 2014, 315 participants completed the study with an overall incidence of postoperative delirium of 40/315 (12.7%). Of these 18 fulfilled the CAM criteria for delirium and 22 were deemed delirious by consensus decision based on chart review. ApoE genotype was not associated with postoperative delirium in this cohort. Tim
taken to complete Colour Trails 2, errors in mini mental state examination and level of pain preoperatively were independent predictors of postoperative delirium.

Conclusions

This study challenges the assertion that ApoE4 genotype predicts postoperative delirium. It replicates previous work suggesting cognitive impairment predicts postoperative delirium and shows for the first time that simple cognitive tests can be as effective as more detailed tests.

Introduction

Postoperative delirium is a serious event for patients, families and healthcare workers and is associated with negative outcomes including cognitive impairment, institutionalisation and death (1). Central to delirium pathophysiology, and ultimately prevention and treatment lies in the interaction of patient vulnerability and precipitating insult(s). Cardiac and hip fracture surgery involving, as they do, significant iatrogenic insults and frail populations respectively, have accounted for the majority of postoperative delirium research to date (1). Elective and emergency (most commonly to treat hip fracture) orthopaedic surgery are often considered as a single entity but they are very different populations with elective patients being younger and more robust (2, 3). Elective orthopaedic surgical procedures can also in themselves vary significantly, a revision hip arthroplasty will often take longer, cause more blood loss and require more analgesia than a primary arthroplasty for example.

Identification of predictors of postoperative delirium will facilitate preoperative informed consent; evidence has shown that multidisciplinary modification of perioperative care can reduce delirium incidence (4). It is not yet clear which modifications of anaesthetic
techniques or drug regimens reduce delirium risk and better risk stratification will improve future observational and interventional studies.

Systematic reviews of predictors of postoperative delirium across surgical populations have consistently identified cognitive impairment as a main predictive factor of postoperative delirium (5, 6). Several studies have investigated the utility of detailed cognitive tests in the prediction of postoperative delirium in non-cardiac surgical populations however conclusions have varied (7-11). Participants in these studies underwent a variety of surgical procedures so there is lack of homogeneity. Also detailed cognitive tests require staff training, are time-consuming and so are resource-heavy. After increasing age, carriage of an Apolipoprotein E ε4 (ApoE4) allele is the next most influential risk factor for Alzheimer’s disease (AD), the most common cause of dementia. Whilst it is neither necessary for nor indicative of developing AD dementia, the ApoE4 allele has been associated with amyloid accumulation in those without cognitive impairment. Studies have considered ApoE4 as a risk factor for delirium in ICU, medical, hip fracture and elective populations (12-15) with a meta-analysis in 2009 suggesting an association between ApoE4 carriage and delirium (12).

This observational cohort study tested the hypotheses that:

[1] carriage of an ApoE4 allele is associated with an increased risk of postoperative delirium.

[2] impaired preoperative neuropsychological performance is associated with a higher incidence of postoperative delirium

[3] more easily administered neuropsychological tests are as effective as more detailed assessments.

**Methods**
Study population

Participants aged 65 years or older admitted for primary elective hip or knee arthroplasty to a single surgical centre under the care of three participating consultant orthopaedic surgeons who were expected to undergo spinal anaesthesia with intrathecal diamorphine were eligible for inclusion. Other anaesthetic variations were permitted.

Exclusion criteria: a pre-existing diagnosis of dementia or other neurodegenerative condition; visual or hearing impairment such that the participant was unable to undertake the neuropsychological assessments; illiteracy; colour blindness; stroke with residual deficit.

Study methods

Written informed consent was obtained from all participants and the study was performed in accordance with local ethical committee procedures (REC reference: 10/NIR01/5; protocol number: 09069PP-OPMS). Available medical notes were reviewed and demographic and baseline data, including a comorbidity index (16), American Society of Anesthesiologists (ASA) physical status (17), number of medications, smoking status and alcohol consumption, recorded as outlined in Table 1.

Baseline neuropsychological tests were undertaken as outlined in Tables 1. Estimated intelligence quotient (IQ), pain and depression were assessed as recognised potential confounding factors. Cognitive tests were chosen to capture the breadth of cognitive domains (memory; executive function; language; visuospatial abilities) whilst focussing on executive function and attention testing the hypothesis that these more subtle deficits may be an earlier marker of brain vulnerability (18, 19). The constituent parts of the Mini Mental State Examination (MMSE) (20) were scored individually.
Postoperatively participants were assessed for delirium once daily by a single researcher (TM initially then EC) for the first three days. The MMSE and the Confusion Assessment Method (CAM) (21) were completed on each occasion. Nursing staff were asked about symptoms of delirium as part of each assessment. Error(s) in WORLD backwards and/or, latterly, months of the year backwards (MOTYB) were not requisite or necessarily sufficient to fulfil the inattention feature of the CAM. Where possible, medical and nursing notes for all participants were interrogated post-discharge by the researchers or a research assistant. In all cases where there were reports of delirium but the CAM criteria had not been fulfilled cases were discussed between EC and PP and a consensus reached as to whether they reflected delirium, according to Diagnostic and Statistical Manual of Mental Disorders IV criteria (22).

To reinforce the assessments an additional test of attention – MOTYB – and the modified Richmond Agitation and Sedation Scale (23) to quantify level of consciousness were added during the study, at participants 177 and 198 respectively.

For researcher training and quality control methods see Supplementary Data.

**DNA collection and analysis**

Venous blood was collected preoperatively, at the point of intravenous cannulation, into PAXgene Blood DNA tubes (Qiagen, catalogue number: 761125). Samples were transported on wet ice to the laboratory where they were processed according to the manufacturer’s instructions and stored at -80 °C.

DNA was analysed using the TaqMan Single Nucleotide Polymorphism Genotyping Assay (Life Technologies; catalogue number 4351379) as per the manufacturer’s instructions.
ApoE status was inferred from the genotype at each of the two alleles, rs7412 and rs429358 (24).

Statistical methods

The study was powered to detect a difference in the risk of delirium between the ApoE4 positive and ApoE4 negative groups. This was based on an estimated ApoE4 carrier rate of 25.5% (25) and a rate of delirium in the ApoE4 positive and ApoE4 negative groups of 28% and 11% respectively in an elective non-cardiac surgery population (26). It was estimated that 316 participants would be required to enable detection of significant difference, at the 5% significance level, with 90% power.

Baseline characteristics, ApoE allele status and neuropsychological performances were compared between the delirium and no delirium groups using t-tests, Mann-Whitney U tests and chi-squared test with Yates' correction (or Fisher's exact test) as appropriate. Positively skewed variables were logarithmically (log) transformed to facilitate parametric testing (further description in Supplementary Data). In keeping with conventional practice, MMSE scores were analysed using means and standard deviations, despite the distribution being negatively skewed. Binary logistic regression, with delirium as the dependent variable, was then used to calculate the adjusted odds ratio for each baseline variable. The significance level was set to 10% (P<0.10) for initial entry into the logistic regression model. Variables were then eliminated in a stepwise fashion until all retained variables were significant (P<0.05). Analysis was performed using SPSS for Windows version 22 (IBM Inc., Armonk, NY, USA).
Methods and results are presented in accordance with STROBE guidance (27) where possible.

**Results**

Between 23rd March 2012 and 21st October 2014, 338 participants were recruited with 315 completing the study. The numbers involved at each stage of the recruitment process are shown in Figure 1. The three participating surgeons undertook 2,083 primary hip and knee replacements in patients aged 65 or over during the study period. Mean (SD) age, with range, for the study cohort and entire potentially eligible cohort was 74.4 (5.8), 65 – 92 years and 74.2 (6.1), 65 – 95 years respectively.

**Missing data**

Participants were assessed postoperatively on complete inpatient days. If participants were not assessed during a full inpatient day and had not refused this was considered missing data (numbers shown in Supplementary Data (Table S1)). Each participant was assessed postoperatively on at least one day. Of the 315 participants completing the study, 159 charts were reviewed post-discharge.

**Incidence of delirium**

The cumulative incidence of delirium in the first three postoperative days was 40/315 (12.7%) with 18 CAM positive for delirium and 22 deemed to be delirium by reports. Delirium incidence by day and those symptoms deemed to represent delirium by consensus decision are shown in Tables S2-4.

**Predictors of delirium**
DNA was unavailable for four participants. Of the remaining 311 participants six were homozygous for the E4 allele, 72 heterozygous and 233 non-carriers. No significant association was found between ApoE4 allele carriage and delirium.

The baseline characteristics and performance on cognitive testing of the total cohort, delirium and non-delirium groups, along with the results of the stated univariate analyses are shown in Table 1. Where there is missing data the n=available data is shown in the total cohort column.

Preoperative MMSE scores were considered for each section of the test. The baseline performance in the MMSE sections for the total cohort, delirium and no delirium groups is shown in Table S5. Making more than one error in orientation questions, an error in 3 item recall or in repetition were significantly associated with postoperative delirium in this cohort.

The results from the logistic regression model consisting of all variables significant at the 10% level on univariate analyses are shown in Table 2. These show that for each mm increase in vertical visual analogue pain at rest score the OR of delirium decreased by a factor of 0.98, for every 10-fold increase in time taken to complete Colour Trails 2 the OR of delirium increased by a factor of 15.93, making any error in 3-item recall increased the OR of delirium by a factor of 3.16, and making more than one error in orientation questions in the MMSE increased the OR of delirium by a factor of 7.75.

The incidence of delirium, by CAM and by reports, over the time course of the study was calculated. There was a difference (p=0.053) between delirium rates before (15/167) and after (25/148) the introduction of the additional test of attention. The logistical regression
model was run again with study epoch included as a variable and no independent association was found.

**Discussion**

First this observational cohort study tested the hypothesis that carriage of an ApoE4 allele would be associated with an increased risk of postoperative delirium. The failure to show an association between ApoE4 carriage and postoperative delirium, in conjunction with recent work (28, 29) challenges the conclusions of a 2009 meta-analysis (12) in the postoperative population.

Results from both univariate and multivariate analyses supported the secondary hypothesis, that is, that impaired preoperative neuropsychological performance is associated with a higher incidence of postoperative delirium following elective primary hip and knee arthroplasty. The independent association between performance on Colour Trails 2 and subsequent delirium suggests it may be an adequate measure of cognitive vulnerability to delirium and is in keeping with previous similar work undertaken in elective, non-cardiac surgery populations (9, 11). Furthermore, it may be related to the association between executive dysfunction and vascular risk factors (19) which were common in this study population. Modified from the Trail Making Test by D'Elia et al (30) Colour Trails is a test of executive function and measures the time in seconds taken by participants to sequentially join alternately coloured circles numbered 1 to 25 arranged in an apparently random order. The mean values in this study population (mean age 74.4 years) are similar to those for the 70-75 year old age groups in a large study of older Irish adults (31).

Third, the study tested the hypothesis that more easily administered neuropsychological tests could be as effective as more detailed time and expertise consuming assessments.
This is the first study, to the authors’ knowledge, to consider 3 item recall and more than one error in orientation questions as predictors of postoperative delirium. Pending replication, the inclusion of these simple cognitive tests could aid preoperative risk stratification and counselling and perioperative care planning. The high incidence of errors in 3-item recall in both delirium and no delirium groups would suggest its utility lies in highlighting those at decreased risk, who make no error, rather than identifying those at high risk by making an error. Making more than one error in orientation questions may be a more sensitive test for identifying those patients at higher risk. Consideration of these as binary variables may have increased the power available to demonstrate a significant association. It was felt however that these were the most appropriate methods to analyse the data.

The independent association between decreased levels of pain at rest and increased risk of postoperative delirium is in contrast to the published literature (32-34). Pain scores at rest are most representative of overall pain burden; patients less sensitised to pain at rest may experience the greatest increase in pain over the acute perioperative period, pre-disposing them to delirium. Further research into this area is warranted.

This study had a number of notable strengths and limitations. Delirium was diagnosed on daily assessment by specifically trained researchers using a validated tool. The overall incidence of postoperative delirium found in this study is in keeping with the published literature (35) (further discussion in the Supplementary Data). Introduction of additional outcome measures during a study is not ideal but it was felt that the CAM alone lacked objectivity which is supported by the recent publication of the 3D-CAM with supporting objective tests (36). Symptoms of delirium falling short of a CAM positive screening for
delirium were common in this study. It is possible that some participants with delirium were not identified using the relatively subjective marking of the CAM items. Future work will consider these subsyndromal delirium features further.

The inclusion of delirium by reports will have mitigated this risk to a certain extent. A validated chart review tool was not used however and 156 of the 315 charts were not reviewed. These were spread across the study period but more charts were reviewed in the second epoch and this may have contributed to the increase with incidence over time. Alternatively this increased incidence over time may have been a true finding. Most published studies do not provide the level of detail outlined herein so it is difficult to know if or how these results regarding study epoch differ from previous studies.

No collateral history was taken preoperatively and diagnosis of dementia was not attempted. It is likely that a few patients with undiagnosed dementia (at an early stage) were included in the study. This is arguably more reflective of the elective arthroplasty population where cognitive impairment often goes unrecognised (37). Hearing and visual impairment, which are recognised risk factors for delirium were not adequately recorded in this study. Spinal anaesthetic with intrathecal diamorphine was chosen as a constant at the outset as it was a most commonly used anaesthetic technique. The permission of various other anaesthetic measures introduces the potential for intraoperative confounding factors not captured by these analyses. Future work will need to further consider these confounders.

In conclusion, this study replicates previous work suggesting preoperative cognitive impairment, especially in a test of sustained attention/executive function, predicts postoperative delirium in a specific elective orthopaedic population. It shows for the first
time that simple cognitive tests can be just as effective as more detailed tests. Along with recently published work it challenges the assertion that carriage of an ApoE4 allele predicts delirium, at least in the postoperative setting.

**Key points**

The incidence of postoperative delirium following elective primary hip and knee arthroplasty was 12.7% in this observational cohort study of patients aged 65 years and over.

Preoperative cognitive impairment predicts postoperative delirium.

Despite its established association with Alzheimer’s disease dementia the Apolipoprotein E4 allele does not predict postoperative delirium.

**Conflict of Interest**

None declared

**Funding**

This work was supported by philanthropic funding, specifically the Siew Keok Chin Scholarship, and charitable funding, from the Belfast Arthroplasty Research Trust, both administered via Queen’s University Belfast. Neither played a role in the planning, execution or reporting of this study.

**Acknowledgements**

We gratefully acknowledge the generous support of: the theatre and ward staff of Musgrave Park Hospital; Dr James Murray, for his help and support during the planning of this project; Mr John Conlon, Mrs Hazel Johnston and Mrs Eilish Armstrong, for their laboratory and
cognitive research expertise; Ms Lauren Anderson, for her help with data management and laboratory support; Professor Chris Patterson, for his statistical expertise and patience, Professor Chris Patterson; and Professor Alasdair MacLullich and team, for their generous advice.

References


17. ASA Physical Status Classification System [Internet]. []. Available from: https://www.asahq.org/resources/clinical-information/asa-physical-status-classification-system.


Table 1. Characteristics, including baseline performance in neuropsychological assessments, for delirium and no delirium groups.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Total cohort (n=315)</th>
<th>Delirium (n=40)</th>
<th>No delirium (n=275)</th>
<th>Test statistic</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td>74.4 (5.8)</td>
<td>76.9 (6.0)</td>
<td>74.0 (5.7)</td>
<td>T= -2.945</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Male (%)</td>
<td>136 (43.2)</td>
<td>18/40 (45)</td>
<td>118/275 (42.9)</td>
<td>X² = 0.006</td>
<td>0.94</td>
</tr>
<tr>
<td>Hip arthroplasty (%)</td>
<td>159 (50.5)</td>
<td>14/40 (35)</td>
<td>145/275 (52.7)</td>
<td>X² = 3.709</td>
<td>0.05</td>
</tr>
<tr>
<td>ASA physical status (I/II/III)</td>
<td>15/250/46 n=311</td>
<td>1/30/9</td>
<td>14/220/37</td>
<td>MWU=4840.5</td>
<td>0.11</td>
</tr>
<tr>
<td>Charlson Comorbidity Index (1), 0/1/2/3/4</td>
<td>177/96/26/6/1 n=306</td>
<td>17/13/7/1/0</td>
<td>160/83/19/5/1</td>
<td>MWU=4172.5</td>
<td>0.04</td>
</tr>
<tr>
<td>Number of medications, mean (SD)</td>
<td>5.8 (3.5)</td>
<td>6.3 (4.2)</td>
<td>5.8 (3.4)</td>
<td>T= -0.926</td>
<td>0.36</td>
</tr>
<tr>
<td>Years in education, mean (SD)</td>
<td>11.7 (2.2)</td>
<td>11.4 (2.1)</td>
<td>11.8 (2.3)</td>
<td>T= 1.049</td>
<td>0.30</td>
</tr>
<tr>
<td>Smoking status, current smokers (%)</td>
<td>22 (7.1)</td>
<td>4/39 (10.3)</td>
<td>18/273 (6.6)</td>
<td>FET</td>
<td>0.50</td>
</tr>
<tr>
<td>Units alcohol, ≥ 11 units/week (%)</td>
<td>32 (10.3)</td>
<td>3/40 (7.5)</td>
<td>29/272 (10.7)</td>
<td>FET</td>
<td>0.78</td>
</tr>
<tr>
<td>Bristol Activities of Daily Living score (2), median (IQR)</td>
<td>2 (1 – 3)</td>
<td>3 (1 – 3)</td>
<td>2 (1 – 3)</td>
<td>MWU=5310.5</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>Geriatric Depression Scale score (3), mean (SD)</strong></td>
<td>5.9 (1.7)</td>
<td>5.9 (1.9)</td>
<td>5.9 (1.6)</td>
<td>T= -0.060</td>
<td>0.95</td>
</tr>
<tr>
<td><strong>Range 0 – 11</strong></td>
<td>n=229</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vertical visual analogue pain scale at rest component, median (IQR)</strong></td>
<td>26 (7 – 55)</td>
<td>14 (3 – 29)</td>
<td>28 (8 – 56.25)</td>
<td>MWU=3885</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><strong>Range 0 – 96</strong></td>
<td>n=313</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vertical visual analogue pain scale on movement component, median (IQR)</strong></td>
<td>75 (55 – 88.5)</td>
<td>72 (52 – 85)</td>
<td>76 (55.75 – 89)</td>
<td>MWU=4914</td>
<td>.5</td>
</tr>
<tr>
<td><strong>Range 7 – 100</strong></td>
<td>n=313</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Estimated IQ from National Adult Reading Test errors, mean (SD)</strong></td>
<td>109.5 (9.5)</td>
<td>106.8 (9.1)</td>
<td>109.9 (9.5)</td>
<td>T=1.957</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Range 87.3 – 128.0</strong></td>
<td>n=311</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ 1ApoE4 allele (%)</td>
<td>78 (25.1)</td>
<td>12 (30.1)</td>
<td>66 (24.3)</td>
<td>Χ² = 0.768</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>n=311</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CLOX 1, median (IQR)</strong></td>
<td>13 (11 – 14)</td>
<td>12 (11.25 – 13)</td>
<td>13 (11 – 14)</td>
<td>MWU=4553</td>
<td>.5</td>
</tr>
<tr>
<td></td>
<td>n=303</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CLOX 2, median (IQR)</strong></td>
<td>14 (13 – 15)</td>
<td>14 (13 – 14.75)</td>
<td>14 (13 – 15)</td>
<td>MWU=5197</td>
<td>.90</td>
</tr>
<tr>
<td></td>
<td>n=303</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean letter fluency, mean (SD)</strong></td>
<td>11.8 (4.7)</td>
<td>10.7 (4.4)</td>
<td>11.9 (4.7)</td>
<td>T=1.620</td>
<td>0.11</td>
</tr>
<tr>
<td><strong>Range 0.7 – 27.0</strong></td>
<td>n=315</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mean category fluency, mean (SD)</strong></td>
<td>17.0 (4.3)</td>
<td>15.0 (4.2)</td>
<td>17.3 (4.3)</td>
<td>T=3.099</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><strong>Range 6.5 – 30.0</strong></td>
<td>n=315</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Stroop Colour Word Score, mean (SD)</strong></td>
<td>23.7 (9.0)</td>
<td>20.0 (7.9)</td>
<td>24.3 (9.0)</td>
<td>T=2.789</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td><strong>Range 4 – 52</strong></td>
<td>n=315</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Log transformed Colour Trails I time, geometric mean (IQR)</strong></td>
<td>(50 – 88)</td>
<td>82.3 (54 – 110)</td>
<td>66.6 (49 – 84.5)</td>
<td>Ratio of means (95%CI) 1.24 (1.02 – 1.50)</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>n=313</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Log transformed Colour Trails 2 time, geometric mean (IQR)</strong></td>
<td>(108 – 188)</td>
<td>178.5 (129 – 246)</td>
<td>140.2 (107 – 185)</td>
<td>Ratio of means (95%CI) 1.27 (1.11 – 1.46)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td></td>
<td>n=307</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York Paragraph Recall Test immediate recall, mean (SD)</td>
<td>4.6 (2.1)</td>
<td>4.2 (1.8)</td>
<td>4.6 (2.2)</td>
<td>T=1.296</td>
<td>0.20</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Range 0 – 12 n=315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New York Paragraph Recall Test delayed recall, mean (SD)</td>
<td>5.1 (2.9)</td>
<td>4.1 (2.7)</td>
<td>5.3 (2.9)</td>
<td>T=2.482</td>
<td>0.01</td>
</tr>
<tr>
<td>Range 0 – 14 n=314</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mini Mental State Examination score, mean (SD)</td>
<td>27.3 (2.3)</td>
<td>26.2 (2.7)</td>
<td>27.4 (2.2)</td>
<td>T=3.129</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Range 16 – 30 n=290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Months of the year backwards score, median (IQR)</td>
<td>12 (12 – 12)</td>
<td>12 (12 – 12)</td>
<td>12 (12 – 12)</td>
<td>MWU=1503</td>
<td>0.70</td>
</tr>
<tr>
<td>Range 1 – 12 n=149</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Months of the year backwards any error (%)</td>
<td>21 (14.1)</td>
<td>3/25 (12)</td>
<td>18/124 (14.5)</td>
<td>FET</td>
<td>1.00</td>
</tr>
<tr>
<td>n=149</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(ASA, American Society Anesthesiologists; SD, standard deviation; MWU, Mann-Whitney U test; FET, Fisher’s Exact test; IQR, interquartile range; IQ, intelligence quotient; CI, confidence intervals)

Absolute number of medications was calculated as a crude measure of drug burden with each medication, including as required medications, inhalers and eye drops, assigned a value of one. Level of education was recorded as years of education, assuming age at starting school of 4 years. Alcohol units per week were estimated using the calculator at [www.drinkaware.co.uk](http://www.drinkaware.co.uk). Smoking status was recorded as current, ex or non-smoking.

References shown in Supplementary Data.

P values significant at the 5% level are shown in bold.
Table 2. Independent predictors of postoperative delirium corrected for those variables significant at the 10% level.

<table>
<thead>
<tr>
<th>Variable</th>
<th>P value</th>
<th>Odds Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vertical visual analogue pain scale, pain at rest</td>
<td>0.02</td>
<td>0.98</td>
</tr>
<tr>
<td>Log transformed Colour Trails 2</td>
<td>&lt;0.01</td>
<td>15.93</td>
</tr>
<tr>
<td>Any error in 3 item recall</td>
<td>0.04</td>
<td>3.16</td>
</tr>
<tr>
<td>More than one error in orientation questions</td>
<td>0.04</td>
<td>7.75</td>
</tr>
</tbody>
</table>

Figure 1. Flow chart showing eligible, excluded and withdrawn participants at each stage.