The Short- and Long-Term Relationship between Delirium and Cognitive Trajectory in Older Surgical Patients

Sharon K. Inouye, M.D., MPH[1,2], **Edward R. Marcantonio, M.D., SM**[1,2], **Cyrus M. Kosar, MA**[2], **Douglas Tommet, MS**[2,3], **Eva M. Schmitt, PhD**[2], **Thomas G. Travison, PhD**[1,2], **Jane S. Saczynski, PhD**[2,4], **Long H. Ngo, PhD**[1], **David C. Alsop, PhD**[5], and **Richard N. Jones, ScD**[2,3]

[1]Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

[2]Aging Brain Center, Institute for Aging Research, Hebrew SeniorLife

[3]Department of Psychiatry and Human Behavior, Warren Alpert Medical School, Brown University, Providence, RI

[4]University of Massachusetts Medical Center and Meyers Primary Care Institute, Worcester, MA

[5]Department of Radiology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA

Abstract

INTRODUCTION—Since the relationship between delirium and long-term cognitive decline has not been well-explored, we evaluated this association in a prospective study.

METHODS—SAGES is an on-going study involving 560 adults age 70+ without dementia scheduled for major surgery. Delirium was assessed daily in the postoperative period using the Confusion Assessment Method. General Cognitive Performance (GCP) and the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE) were assessed preoperatively then repeatedly out to 36 months.

RESULTS—On average, patients with post-operative delirium had significantly lower preoperative cognitive performance, greater immediate (1 month) impairment, equivalent recovery at 2 months, and significantly greater long-term cognitive decline relative to the non-delirium group. Proxy reports corroborated the clinical significance of the long-term cognitive decline in delirious patients.

Corresponding Author: Richard N. Jones, Sc.D., Director, Quantitative Science Program, Associate Professor, Department of Psychiatry & Human Behavior, and Neurology, Warren Alpert Medical School, Brown University, 345 Blackstone Boulevard, Box G-BH, Providence, Rhode Island 02906-4800, Richard_Jones@Brown.edu, Phone: 401-444-1943, Fax: 401-444-1911.

DISCUSSION—Cognitive decline following surgery is biphasic and accelerated among persons with delirium. The pace of long-term decline is similar to that seen with Mild Cognitive Impairment.

Keywords

Delirium; acute confusional state; cognitive decline; dementia; mild cognitive impairment; geriatrics; surgical outcomes; surgical complications

Background

Delirium is a common, serious, often fatal disorder affecting as many as 50% of older people during the course of surgery or hospitalization, and costing more than $164 billion (2011 US dollars) per year.[2] While previously considered a transient condition, increasing evidence suggests that delirium may be associated with longer term effects, including cognitive decline.[2,3] However, the contribution of delirium to long-term outcomes has been difficult to ascertain since prior studies have been limited by the lack of baseline cognitive assessment and short follow-up.[4]

Cognitive decline has been documented following surgery,[5–9] hospitalization,[10,11] and intensive care.[12,13] Each of these inciting events has been associated with long-term cognitive impairment in previous studies, focusing on 3–12 months follow-up. However, methodological challenges have hampered separation of the relative contributions of pre-existing cognitive impairment, acute illness, or surgery.[5,6,9,10] Moreover, the independent contribution of delirium beyond one year remains unclear.

Our goal was to examine the changes in short- and long-term cognitive trajectory associated with delirium up to 36 months following scheduled major surgery in a well-characterized cohort of patients without dementia at baseline. Our hypotheses were that delirium would be independently associated with an acute cognitive decline and with a higher long-term rate of cognitive decline.

Methods

Study Population

Participants were enrolled in an ongoing longitudinal cohort study of older adults undergoing scheduled major non-cardiac surgery, the Successful Aging after Elective Surgery (SAGES) study.[14] Patients were identified from the operating room booking schedules, and were enrolled in their homes. Eligible participants were age 70 years and older, English speaking, and scheduled to undergo surgery at two academic medical centers with an anticipated length of stay of ≥ 3 days. Eligible surgeries were: total hip or knee replacement, lumbar, cervical, or sacral laminectomy, lower extremity arterial bypass surgery, open abdominal aortic aneurysm repair, and open or laparoscopic colectomy. Exclusion criteria[14] included evidence of dementia (self or clinician report, medical record diagnosis, or baseline Modified Mini-Mental State Examination (3MS[15]) score <69 or education-adjusted equivalent), delirium or hospitalization within 3 months, terminal condition, severe blindness or deafness, history of schizophrenia, and alcohol abuse. A total

of 566 eligible patients were enrolled between June 18, 2010 and August 8, 2013. Six patients were subsequently excluded for possible dementia after neuropsychological testing and clinical adjudication, yielding a final sample of 560 patients. Participants were excluded for possible dementia by NIA-AA (National Institutes of Aging Alzheimer's Association) criteria.[16] A 7-member interdisciplinary expert panel consisting of one neurologist, two neuropsychologists, two geriatricians, and two geriatric psychiatrists used a two-step, modified Delphi strategy, with consensus agreement requiring agreement of a minimum of 5 of 7 panel members. The NIA-AA criteria for dementia were evaluated by each panel member by review of the following data: impairment in memory or other cognitive patient-reported basic and instrumental activities of daily living (ADLs, IADLs), proxy Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE[17]), Confusion Assessment Method (CAM[18]) to rule out delirium, and Geriatric Depression Scale (GDS[19]) to rule out severe depression as a cause of poor cognitive performance. To support a classification of dementia, cognitive performance had to be more than 2 standard deviations (sd) below mean (age-adjusted norms) on two or more of the selected tests. A Flow diagram summarizing the screening and enrollment stages is presented in the Appendix. Participants provided written informed consent according to procedures approved by the institutional review boards of all study hospitals.

Characteristics of the non-surgical comparison group (n = 119), used to control for practice or retest effects, are described in detail in the Appendix. Briefly, the surgical and nonsurgical groups had nearly identical mean (±sd) age (77±5 years in both groups), baseline 3MS score (both 94±5), IQCODE score (3.1±0.2, both groups) and GCP score (58±7 vs 58±10 surgical and non-surgical, respectively). The surgical sample had relatively fewer men (42%) than the comparison group (66%) and a lower mean GDS score (2.5±2.5 vs 1.3±1.8).

Data Collection

Participants underwent baseline assessment and medical record review within 30 days before surgery (median 9 days, interquartile range, IQR, 5–17). From postoperative day 1 (i.e., the day following the operative date) through discharge, participants received a daily delirium assessment. After discharge, participants were followed at 1, 2, 6, 12 months, and every 6 months up to 18–36 months. Interviews were conducted by bachelor's and master's level research associates who received 4–6 weeks of intensive training and semiannual standardization. A separate team conducted the in-hospital delirium assessments. Thus, the neuropsychological assessors were blinded to the patient's delirium status. An experienced research physician conducted all chart reviews. Data collection staff members were blinded to the study hypotheses.

Assessment of Delirium

The delirium assessment included brief cognitive testing,[14,20] Delirium Symptom Interview (DSI),[21] and interviews with family and nurses conducted daily. Delirium was rated using the Confusion Assessment Method (CAM),[18] a standardized approach with high sensitivity (94–100%) and specificity (90–95%).[22,23] The CAM ratings demonstrated high reliability (kappa statistic= 0.92 in 71 paired ratings). Combined with a validated chart review

method,[24,25] patients were classified as delirious if either CAM or chart criteria were met on a given day.

Assessment of Cognitive Function

We assessed cognitive performance using neuropsychological testing and proxy ratings using the Informant Questionnaire for Cognitive Decline in the Elderly (IQCODE).[17]

The neuropsychological battery of 11 tests of attention, memory, language, and executive functioning was administered at baseline and each follow-up (Appendix).[14,26] We created a weighted composite summary measure, the General Cognitive Performance (GCP) score following standard procedures,[27] and calibrated this measure to a nationally representative sample of adults age 70 years[28] to yield a mean score = 50 and standard deviation = 10.[26] The GCP, which is sensitive to change with minimal floor and ceiling effects,[26,27] has been applied in prior studies.[27,29–31]

To control for learning effects (improvements over time with repeated testing), we applied an accepted approach (Appendix),[32–35] which involves subtracting a correction derived from repeated administrations in a comparison sample. We used 119 age-matched primary care patients who received the same testing protocol. The corrections derived from the comparison group were subtracted from the scores of each surgical patient at follow-up. Our approach involves subtracting the same value from both groups, and therefore does not impact on the mean difference between groups.

To assess for preoperative cognitive decline, we used the proxy IQCODE, which asks family members to rate current abilities for daily cognitive tasks relative to 10 years ago. We used accepted thresholds[36] to define three categories of IQCODE ratings: IQCODE < 3.2, 3.2 IQCODE < 3.7, and IQCODE 3.7.

Additional Study Variables

The baseline interview assessed sex, race, ethnicity, education, marital status, 15-item Geriatric Depression Scale (GDS),[37] Modified Mini-Mental State (3MS),[15] basic (ADL) and instrumental activities of daily living (IADL).[38,39] Age, surgical type, and Charlson comorbidity score[40] were determined from chart review. The GDS score was missing in 2 participants, which were imputed using Bayesian estimation. Other covariables had no missing data.

At follow-up, participants reported any hospitalizations since the prior assessment. Reliability of these reports, compared with chart review (N=208), revealed overall agreement on number of rehospitalizations of 90% (kappa=0.79, 95% confidence interval, CI, 0.71–0.87) and agreement on admission dates of 99% (kappa=0.78, 95% CI, 0.72–0.85). Participants also reported any new intercurrent illnesses since the index surgery.[40,41]

Statistical Analysis

Mixed effects regression models were used to characterize the adjusted mean GCP scores and IQCODE scores and their trajectory over time. For the GCP analyses, based on examination of the raw data, the model included fixed effects at the 1 and 2 month

assessments and random effects for baseline and change following two months to capture acute decline, recovery, and long-term trajectory (Appendix). Delirium was regressed on baseline cognitive functioning using logistic regression nested within the linear mixed effect model for cognitive change, allowing the pre-operative level to predict the risk of delirium. In turn, the acute decline, recovery, and long-term trajectory were regressed on delirium to capture differential effects by delirium status. Finally, the Wald test was used to examine whether estimated scores at 36 months were significantly different from baseline.

The baseline, acute decline, recovery and long-term trajectory effects models were adjusted for baseline covariables likely to be related to delirium and cognitive decline, including age, gender, non-white race, education, Charlson score, GDS score, impairment in any instrumental activity of daily living, surgery type, IQCODE, and baseline GCP score. Since we wanted to adjust for baseline covariables yet avoid over-controlling for variables that might be intermediaries between delirium and its effect on cognitive trajectory,[42,43] we did not adjust for hospital-related factors. While follow-up was complete to 18 months, only 56% of participants completed 36 month follow-up because of rolling enrollment. Our analyses included all participants using maximum likelihood methods for parameter estimation under the assumption that data are missing at random. Potential bias due to missing data was probed in multiple ways (Appendix). Analyses were conducted with Mplus software (Version 7.3, Muthén & Muthén, Los Angeles CA).

For the IQCODE analysis, we conducted a mixed effect generalized linear model with IQCODE as a repeatedly observed outcome (at months 0, 6, 12, 18, 24, 30 and 36). A random slope for linear time captured changes in IQCODE score, and an interaction of time and delirium status captured differences in linear change over time attributable to delirium. Covariables were identical to those used in the GCP analysis except for the exclusion of IADL and IQCODE as control variables, both of which were considered endogenous with the outcome. Analyses were conducted with Stata software (Version 14.0, Stata Corporation, College Station TX).

Sensitivity Analyses

To assess whether incident delirium is independently associated with long-term cognitive decline or whether the association is confounded or mediated by rehospitalization, intercurrent illnesses or major postoperative complications, we conducted three additional analyses. The first two included the subsample of 288 participants (51% of the study sample) with complete data up to 36 months, and the sensitivity analysis adjusting for major complications used all 560 participants.

Results

Baseline characteristics overall and by delirium group are shown in Table 1. The mean age (standard deviation, SD) of the sample was 76.7 (5.2) years and 58% were women. Delirium occurred in 134/560 (24%). At baseline, the delirium group had higher levels of comorbidity with 43% having a Charlson index score of 2; lower 3MS, and GCP scores 0.38 sample SD lower than the non-delirium group at baseline (all P<0.05). Despite lower values in the delirious group, the baseline GCP values in both groups were above the expected mean (50)

for a representative older U.S. population. The delirium group also had a higher mean IQCODE score [3.2 (0.3)] relative to the non-delirium group [3.1 (0.2)], amounting to 0.29 sample SD units.

Median duration of follow-up for this ongoing cohort was 36 months (IQR 24–37). Deaths, ascertained from family interviews, charts, or obituary reviews, occurred in 37 (7%) patients after a median follow-up of 19 months (IQR 12–26). The cumulative rate of death was similar with (12/134, 9%) and without (25/426, 6%) delirium (P=0.21). An additional 27 (5%) participants withdrew from follow-up after a median of 5 months (IQR 3–12). Rates of drop-out were similar between the groups with (8/134, 6%) and without (19/426, 4%) delirium (P=0.48). A total of 496 (89%) eligible participants completed all planned study visits, a range of 1–9 visits per participant. In multiple analyses, we did not detect evidence of bias in the results due to missing observations (Appendix) which were predominantly due to the rolling enrollment into the study.

Cognitive function scores over time

We examined the change in raw GCP scores from baseline up to 36 months in the overall sample, and stratified by delirium status (Table 2). The mean raw baseline GCP score was 57.6 points (SD 7.3), indicating that the mean cognitive performance was about 3/4 of the population standard deviation (population SD = 10)[26] above the national average. While the degree of GCP change varied, overall and in the two subgroups (delirium and non-delirium), the GCP followed a similar pattern of significant decline at 1 month, followed by recovery above baseline at 2 months, then gradual decline to below baseline from 12–36 months.

Cognitive trajectory by delirium status

Table 3 shows the differences in mean GCP scores stratified by delirium status and adjusted for covariables – including baseline GCP – in the multivariable mixed effects model. At 1 month, the delirium group has a significantly greater decline (−1.03 points) compared to those without delirium (P<0.003). At 2 months, cognitive function has recovered with no significant difference between groups (P=0.99). Beyond 2 months both groups decline on average, but the delirium group declines significantly more with a net −1.07 -point greater decline in adjusted mean scores at 36 months (P=0.02). Comparing changes from baseline to 36 months reveals a significant change for the delirium group (difference=−1.30, 95% CI −2.06 to −0.54, P < 0.01) and no significant change for the group without delirium (difference= −0.23, 95% CI −0.65 to 0.20, P = 0.30).

Summary results of the mixed effects model, presented in Figure 1, Panels A and B. The figure demonstrates the biphasic relationship of delirium status with general cognitive performance over time. In the short-term, delirium is associated with an accentuated decline at 1 month and recovery back to baseline by 2 months. In the long-term, delirium is associated with a steeper slope, indicating a faster pace of decline over time compared with the non-delirium group. In Figure 1, Panel A, we start at the overall cohort baseline mean since delirium occurred subsequent to baseline. In Figure 1, Panel B, we start at adjusted baseline GCP values: mean ± standard error 55.6 ± 0.5 points for the delirium group and 58.2 ± 0.3 points for the non-delirium group.

Delirium was associated with a significant mean adjusted acute cognitive decline of −1.04 point on GCP at 1 month (Appendix). This effect is resolved by 2 months. In the non-delirium group, the estimated slope for long-term cognitive change (from month 2–36) was −0.21 points/year (95% CI −0.36, −0.06) (Appendix). Over 3 years, the magnitude of change is equivalent to 8.6% of the total sample baseline standard deviation (SD = 7.3). The delirium group demonstrates a significantly greater rate of cognitive decline relative to those without delirium, with an *additional* decrease of −0.38 points in GCP per year (95% CI −0.70, −0.07, P = 0.02). Thus, expected change in the delirious patients is −0.59 GCP points/year, or 24.2% of the total sample baseline standard deviation over 3 years. The difference in the rate of cognitive decline observed between the delirium and non-delirium groups can be interpreted as the acceleration of cognitive decline attributable to delirium (−0.59/−0.21 GCP points), which is a 2.8-fold increase (Appendix).

Family-rated cognitive decline by delirium status

The analysis of IQCODE change revealed that the greater cognitive decline experienced by delirious patients was noticed by their families. The IQCODE ratings increased (worsened) at a pace of 0.03 IQCODE points per year in patients who did not develop delirium, and at a pace of 0.06 points per year in the delirium group, amounting to 0.10 IQCODE units at 3 years (95% CI 0.04, 0.17) (Figure 2). With a baseline pooled SD of 0.24, this yields a moderate effect size of 0.44 SD units over 3 years.

In Table 4, we summarize the IQCODE results using clinical thresholds applied to the IQCODE score for the baseline and 36 month follow-up ratings. The overall odds ratio (OR) for being in a higher category for delirious vs. non-delirious, derived from an adjusted ordinal logistic regression model, is 2.5 (95% CI 1.4, 4.5). If we consider only those patients who had low IQCODE ratings (IQCODE < 3.2) at baseline, 21/60 (35%) of those who developed delirium were above the 3.2 threshold at 36 months, versus 34/181 (19%) of the non-delirium group.

Sensitivity analyses

To examine the effects of delirium independent of subsequent rehospitalization and intercurrent illnesses on trajectories of GCP score, we restricted the sample to 288 participants who completed 36 month follow-up (Appendix), which ensured that participants had passed through all possible risk periods for rehospitalization and intercurrent illnesses. In adjusted analyses, patients with delirium experienced a significant acceleration of cognitive decline of −0.32 GCP points (95% CI −0.70 to +0.05, P = 0.09) per year, which approximates the effect from the main analysis. The higher slope persisted after controlling for any new intercurrent illness or rehospitalization (adjusted difference in slope = −0.32 points, 95% CI −0.69, +0.05 P = 0.09). For the sensitivity analysis that included adjustment for major complications we found essentially no impact on the association of delirium and long-term cognitive decline (−0.36 points, 95% CI −0.69, −0.04 P = 0.02). More details on the sensitivity analyses are presented in the Appendix.

Discussion

In a prospective cohort study of initially high functioning older persons without dementia undergoing major scheduled non-cardiac surgery, we demonstrated a cognitive trajectory characterized by acute decline and recovery, followed by gradual decline out to 36 months. Patients who developed delirium experienced significantly greater short-term decline and accelerated long-term cognitive decline. At baseline, delirium and non-delirium groups differed by about 2.9 GCP units (Table 2). At two months, the delirium and non-delirium groups differed by 3.2 units, about the same difference. Accounting for baseline differences and the effects of potentially confounding variables, there was no significant difference between delirious and non-delirious patients at 2 months (0.01 GCP units, Table 3). This was observed despite the raw score difference of 3.9 GCP units (Table 2) and adjusted difference of 1.03 GCP units at 1 month (Table 3). Therefore, despite large and significant differences in cognitive performance at 1 month after surgery, those with an intervening delirium were, on average, no worse off cognitively than those who did not develop delirium at 2 months. At 36 months, the estimated level of cognitive performance was significantly lower than baseline in the delirium group, while the non-delirium group remained stable. The effect of delirium remains undiminished after controlling for subsequent rehospitalizations, intercurrent illnesses and major postoperative complications. The estimated longitudinal change in GCP over 3 years is about 0.24 (95% CI 0.13 to 0.36) sample SD units relative to baseline among those with delirium and 0.09 (95% CI 0.02 to 0.15) sample SD units among those without delirium. Patients with delirium experienced an approximately 3-fold greater rate of decline, and this decline was noticed by family members as rated by IQCODE. The rate of cognitive decline in the group without delirium is low (0.03 sample SD units/year), attributed to the effects of cognitive aging. This value falls within the range of published studies examining initially high functioning groups, from 0.01[44] to 0.04[45] SD units/year. On the other hand, the observed decline associated with delirium (0.08 SD units/year) approaches that for Mild Cognitive Impairment (about 0.1 SD units/year).[44] By contrast, the effect of dementia would be about 0.33–0.66 SD units/year.[46,47] Perhaps the most significant finding in this study is that even in this high-functioning group, who started substantially above the population norm, delirium exerted potentially permanent effects.

This study allowed us to substantively extend our previous work,[7] which examined cognitive trajectories for one year following cardiac surgery. We previously demonstrated a significant decline in mean cognitive scores at one month, but no significant difference at one year using the Mini-Mental State Examination. The present study utilized a larger sample size, better characterization of baseline and follow-up status with formal neuropsychological testing, more complete data collection, and extended follow-up to 36 months. Such prolonged follow-up is essential to detect changes in cognitive trajectory beyond the effect of cognitive aging alone.

Our demonstration of a biphasic pattern of cognitive decline following delirium, separated by a *return to baseline* was not anticipated. We believe that this distinct biphasic pattern has not been demonstrated before because prior studies have included at least some patients with dementia. The return to baseline following the initial acute decline among patients both with

and without delirium provides evidence of resilience, and thus, diminishes the likelihood of substantial pre-existing dementia in either group. The acute decline (most prominent in the delirium group) likely represents the impact of transient precipitating events, such as anesthesia, surgery, hospitalization, which resolve relatively early. The later effects, however, with a higher rate of cognitive decline following delirium suggest a more long-lasting process. The biological basis of these findings requires further investigation. It is possible that delirium sets off a cascade of events, such as compromised blood-brain barrier, neuro-inflammation, stress response, or alterations in neurotransmission,[2,48,49] which–after an initial rebound that cannot be maintained–lead to progressive, long-lasting effects. Alternatively, delirium may be associated with a pre-existing higher rate of cognitive decline that is not detectable at baseline even with our careful multi-step screening for dementia; however, without repeated neuropsychological testing, we are unable to fully establish the preoperative trajectory. In either case, delirium may serve as a robust marker for persons with poor cognitive reserve at heightened risk for accelerated long-term cognitive decline.[50] If pre-existing subclinical dementia or early cognitive decline did exist, our study documents that the cognitive decline trajectory was substantially worsened following delirium. Even if not causative, delirium identifies a high risk group for subsequent decline who remain important to target for close clinical follow-up and preventive interventions.

Strengths of the study include the large prospective cohort with rigorous data collection; careful characterization of cognition at baseline; neuropsychological testing over time; and standardized delirium assessments. Another important strength was the exclusion of patients with even mild dementia, allowing for estimation of delirium effects free of this potentially confounding influence. The success of the study in excluding dementia is demonstrated by the high baseline GCP scores and the low rate of APOE-ε4 demonstrated previously.[51] Adjustment for learning effects and extended follow-up after delirium represent other advances. In this paper, we do not control for post-operative cognitive decline (POCD), typically defined in terms of cognitive decline, since it is an intermediate (endogenous) variable and would necessarily bias the results.[52]

Several caveats about this study should be noted. While we controlled for learning effects, patients recovered above baseline levels at 2 months suggesting either that this control was incomplete, or that patients had depressed cognitive levels at baseline likely due to pain, pre-admission narcotics, other psychoactive medications, or immobility. It is unlikely that delirium or subsyndromal delirium at baseline contributed to this finding, since these were carefully ruled out. While delirium severity and duration may have influenced the subsequent cognitive trajectory, analyses of these complex influences were considered beyond the scope of this study. Despite our rigorous epidemiologic methods, we cannot rule out pre-clinical dementia at baseline. Patients with delirium had lower GCP scores at baseline than those without delirium, although both groups were above the population mean. In addition, the study population represents a highly educated sample with relatively low racial diversity from a single city. It is important to note, however, that the diversity characteristics of our sample (92% white) are representative of the greater Boston metropolitan area in the over 65 age group (2008–2012 census data).[53] While the internal validity of our results is not threatened, generalizability may be limited and the results require replication in more diverse samples. While we controlled for confounding due to

rehospitalization and intercurrent illnesses, these may be incomplete. Many factors, including perioperative factors and surgical complications, may mediate the association between delirium and long-term cognitive outcomes. While we wanted to avoid over-controlling for variables that might be intermediaries between delirium and cognitive decline, such control will be important in future work to clarify mechanisms and targets for intervention. Another caveat is the incomplete follow-up. Whereas 560 surgical patients are included in the baseline sample, only 312 are followed through the 36 month follow-up time point. However, most of this incomplete follow-up can be viewed as missing completely at random and due to the rolling accrual of the study sample. We only failed to collect follow-up data on 4 of 316 surgical patients eligible for participating in the month 36 follow-up (Table 2), putting our 36 month response rate at 98.7%. Finally, our analysis focuses on a single composite cognitive outcome. Exploration of domain-specific changes in cognition was beyond the scope of the present analysis, and represents an important area for future investigation.

This study provides a novel presentation of the biphasic relationship of delirium and cognitive trajectory, both its well-recognized acute effects but also long-term effects. Our results suggest that following a period of initial recovery, patients with delirium experience a substantially accelerated trajectory of cognitive aging. Our data provide substantive support for previous work suggesting a heightened risk of dementia following delirium.[3] Whether delirium actually causes this long-term decline, or serves as a marker of individuals with reduced reserve at risk for this decline, remains uncertain. Given that delirium is preventable in 40% of cases,[2] this new conceptualization of delirium will warrant renewed attention to the importance of prevention to diminish its acute and potential long-term adverse effects. Moreover, our results hold implications for clinical trials, since treatment during the acute phase where most patients will recover may not necessarily lead to long-term benefit; thus, new interventions may be required to forestall long-term cognitive decline following delirium.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Popp J. Delirium and cognitive decline: more than a coincidence. Curr Opin Neurol. 2013; 26(6): 634–639. [PubMed: 24152819]

2. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. The Lancet. 2014; 383(9920):911–922.

3. Witlox J, Eurelings LSM, de Jonghe JFM, Kalisvaart KJ, Eikelenboom P, Van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia. JAMA. 2010; 304(4):443–451. [PubMed: 20664045]

4. MacLullich AMJ, Beaglehole A, Hall RJ, Meagher DJ. Delirium and long-term cognitive impairment. International Review of Psychiatry. 2009; 21(1):30–42. [PubMed: 19219711]

5. Avidan MS, Evers AS. Review of clinical evidence for persistent cognitive decline or incident dementia attributable to surgery or general anesthesia. J Alzheimer's Dis. 2011; 24(2):201–216. [PubMed: 21447878]

6. Newman S, Stygall J, Hirani S, Shaefi S, Maze M. Postoperative cognitive dysfunction after noncardiac surgery: a systematic review. Anesthesiology. 2007; 106(3):572–590. [PubMed: 17325517]

7. Saczynski J, Marcantonio ER, Inouye SK, et al. Cognitive trajectories after postoperative delirium. N Engl J Med. 2012; 367(1):30–39. [PubMed: 22762316]

8. Selnes OA, Grega MA, Bailey MM, et al. Do Management Strategies for Coronary Artery Disease Influence 6-Year Cognitive Outcomes? The Annals of Thoracic Surgery. 2009; 88(2):445–454.e443. [PubMed: 19632391]

9. Selnes OA, Gottesman RF, Grega MA, Baumgartner WA, Zeger SL, McKhann GM. Cognitive and neurologic outcomes after coronary-artery bypass surgery. N Engl J Med. 2012; 366(3):250–257. [PubMed: 22256807]

10. Mathews SB, Arnold SE, Epperson CN. Hospitalization and cognitive decline: Can the nature of the relationship be deciphered? The American Journal of Geriatric Psychiatry. 2014; 22(5):465–480. [PubMed: 23567430]

11. Wilson R, Hebert L, Scherr P, Dong X, Leurgens S, Evans D. Cognitive decline after hospitalization in a community population of older persons. Neurology. 2012; 78(13):950–956. [PubMed: 22442434]

12. Ehlenbach WJ, Hough CL, Crane PK, et al. Association between acute care and critical illness hospitalization and cognitive function in older adults. JAMA: the journal of the American Medical Association. 2010; 303(8):763–770. [PubMed: 20179286]

13. Pandharipande PP, Girard TD, Jackson JC, et al. Long-term cognitive impairment after critical illness. N Engl J Med. 2013; 369(14):1306–1316. [PubMed: 24088092]

14. Schmitt EM, Marcantonio ER, Alsop DC, et al. Novel risk markers and long-term outcomes of delirium: the successful aging after elective surgery (SAGES) study design and methods. J Am Med Dir Assoc. 2012; 13(9):818 e811–810. [PubMed: 22999782]

15. Teng EL, Chui HC. The Modified Mini-Mental State (3MS) examination. J Clin Psychol. 1987; 48(8):314–318.

16. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimer's & Dementia: The Journal of the Alzheimer's Association. 7(3):263–269.

17. Jorm A, Scott R, Cullen J, MacKinnon A. Performance of the informant questionnaire on cognitive decline in the elderly (IQCODE) as a screening test for dementia. Psychol Med. 1991; 21:785–790. [PubMed: 1946866]

18. Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegal AP, Horwitz RI. Clarifying confusion: The Confusion Assessment Method. A new method for detection of delirium. Ann Intern Med. 1990; 113(12):941–948. [PubMed: 2240918]

19. Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a geriatric depression screening scale: a preliminary report. J Psychiatr Res. 1983; 17(1):37–49. [PubMed: 7183759]

20. Simon S, Bergmann M, Jones RN, Murphy K, Orav E, Marcantonio E. Reliability of a structured assessment for non-clinicians to detect delirium among new admissions to post-acute care. Journal of the American Medical Directors' Association. 2006; 7(7):412–415. [PubMed: 16979083]

21. Albert MS, Levkoff SE, Reilly C, et al. The Delirium Symptom Interview: An interview for the detection of delirium symptoms in hospitalized patients. J Geriatr Psychiatry Neurol. 1992; 5(1):14–21. [PubMed: 1571069]

22. Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does This Patient Have Delirium? JAMA: The Journal of the American Medical Association. 2010; 304(7):779–786. [PubMed: 20716741]

23. Wei LA, Fearing MA, Sternberg EJ, Inouye SK. The Confusion Assessment Method: a systematic review of current usage. J Am Geriatr Soc. 2008; 56(5):823–830. [PubMed: 18384586]

24. Inouye SK, Leo-Summers L, Zhang Y, Bogardus ST Jr, Leslie DL, Agostini JV. A chart-based method for identification of delirium: validation compared with interviewer ratings using the confusion assessment method. J Am Geriatr Soc. 2005; 53(2):312–318. [PubMed: 15673358]

25. Saczynski JS, Kosar CM, Xu G, et al. A tale of two methods: chart and interview methods for identifying delirium. J Am Geriatr Soc. 2014; 62(3):518–524. [PubMed: 24512042]

26. Gross AL, Jones RN, Fong TG, Tommet D, Inouye SK. Calibration and validation of an innovative approach for estimating general cognitive performance. Neuroepidemiology. 2014; 42(3):144–153. [PubMed: 24481241]

27. Jones RN, Rudolph J, Inouye S, et al. Development of a unidimensional composite measure of neuropsychological functioning in older cardiac surgery patients with good measurement precision. J Clin Exp Neuropsychol. 2010; 32(10):1049–1051.

28. Langa KM, Plassman BL, Wallace RB, et al. The Aging, Demographics, and Memory Study: study design and methods. Neuroepidemiology. 2005; 25(4):181–191. [PubMed: 16103729]

29. Saczynski JS, Inouye SK, Kosar CM, et al. Cognitive and brain reserve and the risk of postoperative delirium in older patients: analysis of data from a prospective observational study. The Lancet Psychiatry. 2014; 1(6):437–443.

30. Cavallari M, Hshieh TT, Guttmann CR, et al. Brain atrophy and white matter hyperintensities are not significantly associated with incidence and severity of postoperative delirium in older persons without dementia. Neurobiol Aging. 2015

31. Gross A, Sherva R, Mukherjee S, et al. Calibrating longitudinal cognition in Alzheimer's disease across diverse test batteries and datasets. Neuroepidemiology. 2014; 43(3–4):194–205. [PubMed: 25402421]

32. Lewis M, Maruff P, Silbert B. Statistical and conceptual issues in defining post-operative cognitive dysfunction. Neurosci Biobehav Rev. 2004; 28(4):433–440. [PubMed: 15289007]

33. Evered L, Scott DA, Silbert B, Maruff P. Postoperative cognitive dysfunction is independent of type of surgery and anesthetic. Anesth Analg. 2011; 112(5):1179–1185. [PubMed: 21474666]

34. Devapalasundarum A, Silbert B, Evered L, Scott D, MacIsaac A, Maruff P. Cognitive function in patients undergoing coronary angiography. Heart Asia. 2010; 2(1):75–79. [PubMed: 27325949]

35. Soinne L, Helenius J, Tikkala I, et al. The effect of severe carotid occlusive disease and its surgical treatment on cognitive functions of the brain. Brain Cogn. 2009; 69(2):353–359. [PubMed: 18823689]

36. Harrison JK, Fearon P, Noel-Storr AH, McShane R, Stott DJ, Quinn TJ. Informant Questionnaire on Cognitive Decline in the Elderly (IQCODE) for the diagnosis of dementia within a general practice (primary care) setting. Cochrane Database of Systematic Reviews. 2014; 7

37. Sheikh JI, Yesavage J. Geriatric Depression Scale (GDS): Recent evidence and development of a shorter version. Clin Gerontol. 1986; 5:1–2.

38. Katz S, Ford A, Moskowitz R, Jackson B, Jaffe M. Studies of illness in the aged. The index of ADL: A standardized measure of biological and psychosocial function. Journal of the American Medical Association. 1963; 185(12):914–919. [PubMed: 14044222]

39. Lawton M, Brody E. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. 1969; 9:179–186. [PubMed: 5349366]

40. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. Journal of Chronic Disease. 1987; 40(5):373–383.

41. Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW. Can comorbidity be measured by questionnaire rather than medical record review? Med Care. 1996; 34(1):73–84. [PubMed: 8551813]

42. Glymour, MM.; Greenland, S. Causal diagrams. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. Modern Epidemiology. Third. New York: Wolters Kluwer, Lippincott Williams & Wilkins; 2008. p. 183-209.

43. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology (Cambridge, Mass). 2009; 20(4):488.

44. Johnson JK, Gross AL, Pa J, et al. Longitudinal change in neuropsychological performance using latent growth models: a study of mild cognitive impairment. Brain imaging and behavior. 2012; 6(4):540–550. [PubMed: 22562439]

45. Hayden KM, Reed BR, Manly JJ, et al. Cognitive decline in the elderly: an analysis of population heterogeneity. Age Ageing. 2011; 40(6):684–689. [PubMed: 21890481]

46. Fong T, Jones RN, Shi P, et al. Delirium accelerates cognitive decline in Alzheimer disease. Neurology. 2009; 72(18):1570–1575. [PubMed: 19414723]

47. Gross AL, Jones RN, Habtemariam DA, et al. Delirium and long-term cognitive trajectory among persons with dementia. Arch Intern Med. 2012; 172(17):1324–1331. [PubMed: 23403619]

48. Montagne A, Barnes SR, Sweeney MD, et al. Blood-Brain Barrier Breakdown in the Aging Human Hippocampus. Neuron. 2015; 85(2):296–302. [PubMed: 25611508]

49. MacLullich AMJ, Ferguson KJ, Miller T, de Rooij SEJA, Cunningham C. Unravelling the pathophysiology of delirium: a focus on the role of aberrant stress responses. J Psychosom Res. 2008; 65(3):229–238. [PubMed: 18707945]

50. Jones RN, Fong TG, Metzger E, et al. Aging, brain disease, and reserve: implications for delirium. Am J Geriatr Psychiatry. 2010; 18(2):117–127. [PubMed: 20104068]

51. Vasunilashorn SM, Ngo L, Kosar CM, et al. Does Apolipoprotein E Genotype Increase Risk of Postoperative Delirium? Am J Geriatr Psychiatry. 2015; 23(10):1029–1037. [PubMed: 26238230]

52. Weinberg C. Toward a clearer definition of confounding. Am J Epidemiol. 1993; 137(1):1–8. [PubMed: 8434568]

53. US Census Bureau. American FactFinder 2010. 2013. http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. Accessed 8 August, 2015

Research in Context

1. Systematic Review

The authors reviewed the literature using PubMed and hand searches of bibliographies. Recent publications describing the relationship of delirium to long-term cognitive decline are appropriately cited. A recent systematic review concluded that this area remains controversial and underexplored.[1]

2. Interpretation

This study represents the longest-running prospective observational study of the long-term consequences of delirium ascertained by direct assessment. We found, on average, that in persons without delirium, post-operative effects on cognition resolve to preoperative baseline levels; by contrast, in persons with delirium cognition declines significantly below baseline levels over 36 months at a pace similar to that of Mild Cognitive Impairment.

3. Future Directions

This study justifies further research on delirium's cognitive sequelae and relationship to dementia and Alzheimer's disease. Our results call for additional studies to elucidate whether delirium is simply a risk marker for brain vulnerability or if delirium itself constitutes a potential contributor to dementia.
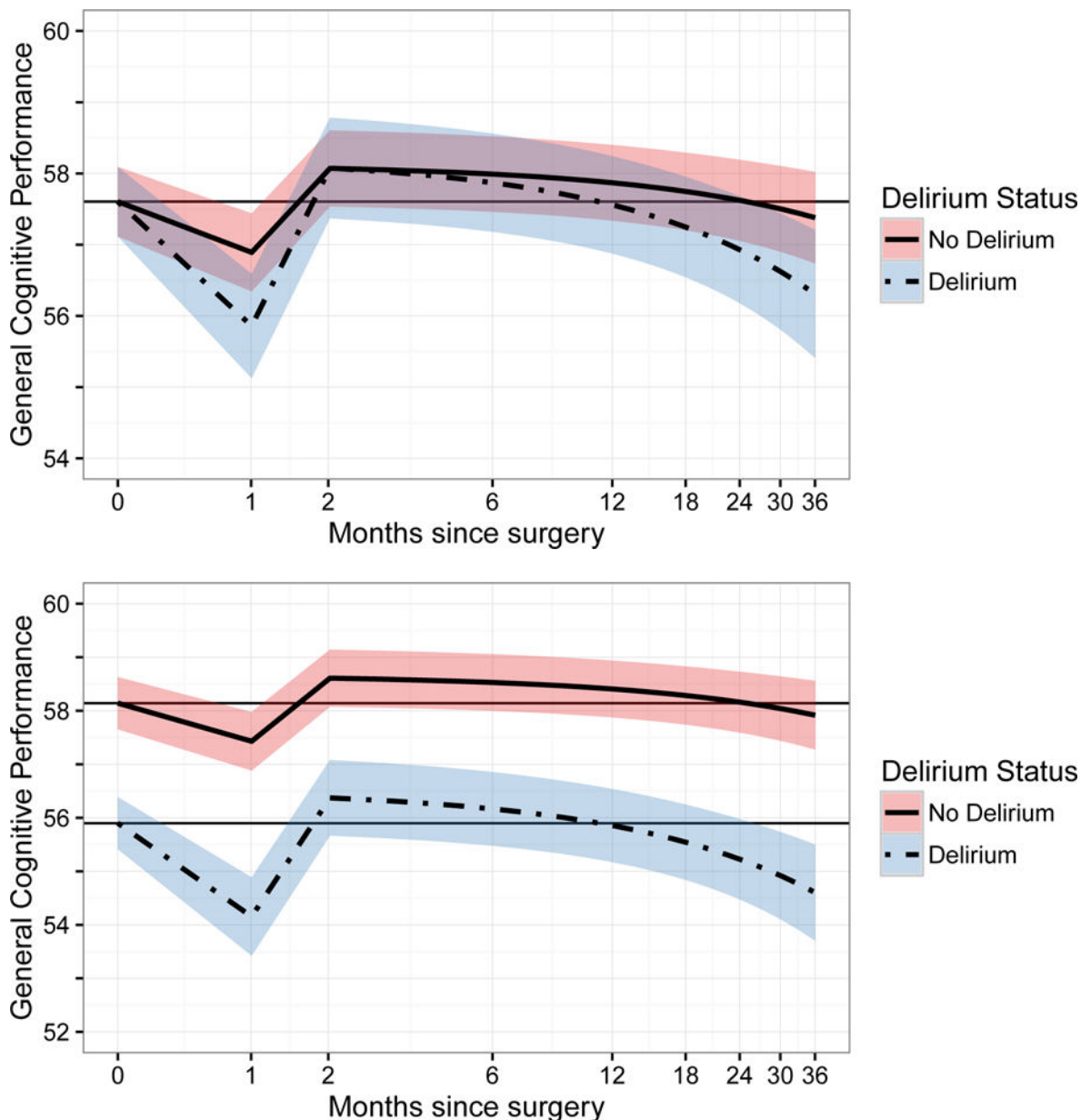
Figure 1. GCP Trajectory by Delirium Status

Panel A demonstrates the relationship between the estimated General Cognitive Performance (GCP) scores derived from random effects models (y-axis) and time since surgery on a natural log scale (x-axis). In this figure, the models are adjusted for the effect of baseline GCP score in delirium and cognitive change, since preoperative cognition predicts delirium. Therefore, at baseline delirium status is unknown and both groups start at the overall sample mean value of GCP. The zero-time values represent the preoperative baseline, and follow-up time points are relative to the index surgery and placed on a log scale. The delirium group is indicated by the solid black line surrounded by its associated 95% confidence interval in dark gray shading. The no delirium group is indicated by a gray line, surrounded by its associated 95% confidence interval in light gray shading. Solid gray

reference lines indicate the baseline level of GCP. The biphasic relationship of delirium with GCP over time is demonstrated. In the acute phase, patients with delirium experienced a significantly accentuated acute decline in cognitive impairment at 1 month and recovery back to baseline by 2 months. In the long-term phase, patients with delirium experienced a significantly higher rate of cognitive decline over time (see text for details).

Panel B demonstrates the same models and relationships as Panel A, however, for these models, the starting point of the curves is offset by the mean difference in baseline GCP according to delirium status during hospitalization.
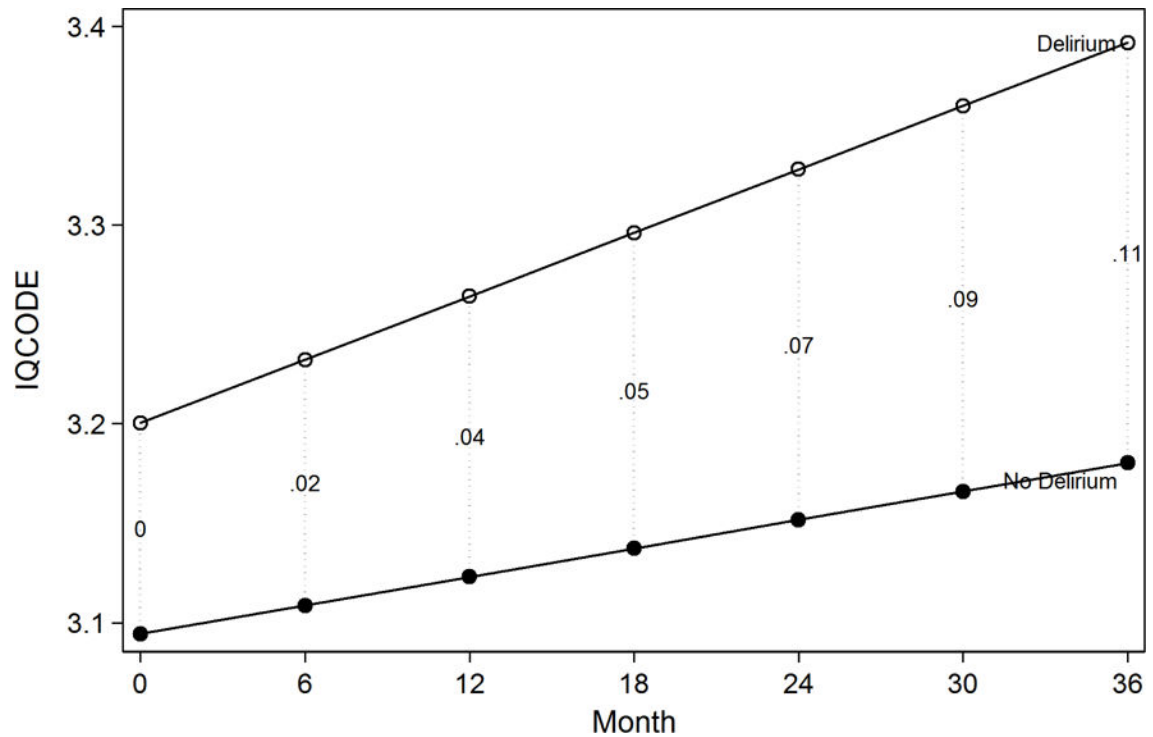
Figure 2. IQCODE Trajectory by Delirium Status
The figure demonstrates the relationship between the estimated IQCODE score derived from a random effects model (y-axis) and time since surgery (x-axis). In this figure, the models are adjusted for baseline GCP score, since this score predicts delirium and therefore both groups start at the same mean value of GCP. Estimates are also adjusted for age, sex, race, English as a second language, Charlson comorbidity index and surgery type. The vertical dotted lines connecting the cotemporaneous estimates for delirious and non-delirious groups are labeled with the net effect differences (differences at that time point net of baseline differences). Proxy IQCODE ratings were available on 548 patients (98% of 560) at baseline and 290 patients (52% of 560) at 36 month follow-up.

Table 1

Baseline characteristics of the study cohort (N=560)[*]

Characteristic	Full Sample N= 560	Delirium N=134	No Delirium N=426
Age- mean years (SD)	76.7 (5.2)	77.5 (5.0)	76.4 (5.2)
Female- n (%)	326 (58)	81 (60)	245 (58)
Nonwhite- n (%)	42 (8)	13 (10)	29 (7)
Education- mean years (SD)	15.0 (2.9)	14.7 (3.0)	15.1 (2.9)
Married- n (%)	332 (59)	79 (59)	253 (59)
Lives Alone- n (%)	167 (30)	39 (29)	128 (30)
Charlson Score- n (%)			
0	257 (46)	54 (40)	203 (48)
1	139 (25)	23 (17)	116 (27)
2+	164 (29)	57 (43)	107 (25)
GDS 15 score- mean (SD)	2.5 (2.5)	3.0 (2.8)	2.3 (2.4)
GCP score- mean (SD)	57.6 (7.3)	54.7 (6.5)	58.5 (7.3)
3MS score- mean (SD)	93.5 (5.4)	91.6 (5.8)	94.1 (5.1)
IQCODE – mean (SD)	3.12 (0.24)	3.19 (0.29)	3.10 (0.21)
Impaired in ADL-n (%)	42 (8)	12 (9)	30 (7)
Impaired in IADL-n (%)	152 (27)	48 (36)	104 (24)
Surgery type- n (%)			
Orthopedic	454 (81)	105 (79)	349 (82)
Vascular	35 (6)	11 (8)	24 (6)
General	71 (13)	18 (13)	53 (12)

[*] ADL = Activities of Daily Living, impairment indicated by human assistance to complete any activity; GCP= General Cognitive Performance; GDS 15= Geriatric Depression Scale 15 point version, range (0–15), higher is worse; IADL = Instrumental Activities of Daily Living, impairment indicated by human assistance to complete any activity; 3MS = Modified Mini-Mental State Exam, range (0–100), lower indicates impairment; SD= standard deviation. IQCODE = Informant Questionnaire for Cognitive Decline in the Elderly, high scores imply greater proxy-reported cognitive change relative to 10 years previously. The Charlson comorbidity score ranged from 0–35, with scores of 2 or more indicating higher comorbidity. GDS scores were missing in 2 participants; no missing data for other variables.

Table 2

Descriptive statistics of corrected raw GCP scores over time (N=560)[*]

Month visit	N	Mean (SD)	Mean change from baseline
Overall			
0	560	57.6 (7.3)	–
1	548	56.8 (7.9)	−0.94 (3.3)
2	536	58.0 (7.9)	0.26 (3.4)
6	528	58.2 (7.5)	0.37 (3.4)
12	511	58.4 (7.6)	0.53 (3.3)
18	499	58.3 (8.0)	0.40 (3.4)
24	474	58.2 (8.0)	0.11 (3.7)
30	325	57.5 (8.2)	−0.16 (4.2)
36	312	57.1 (8.4)	−0.47 (4.5)
Delirium			
0	134	54.7 (6.5)	–
1	129	52.9 (7.0)	−1.95 (3.5)
2	126	54.8 (7.4)	−0.05 (4.4)
6	124	55.3 (7.2)	0.29 (4.0)
12	120	55.1 (6.4)	−0.00 (3.9)
18	117	54.7 (7.6)	−0.13 (3.9)
24	106	55.4 (7.3)	−0.54 (4.1)
30	83	54.5 (7.8)	−1.07 (4.1)
36	86	53.7 (7.7)	−1.90 (4.9)
No delirium			
0	426	58.5 (7.3)	–
1	419	58.0 (7.7)	−0.64 (3.2)
2	410	59.0 (7.8)	0.35 (3.1)
6	404	59.1 (7.4)	0.39 (3.2)
12	391	59.4 (7.6)	0.69 (3.1)
18	382	59.4 (7.8)	0.56 (3.3)
24	368	59.2 (7.7)	0.29 (3.6)
30	242	58.4 (8.0)	0.15 (4.1)
36	226	58.4 (8.0)	0.08 (4.2)

GCP = General Cognitive Performance

[*] Note: All postoperative GCP values corrected for learning effects (see text for details). The number of participants completing each the interview/the number of participants eligible for the interview for each time point follows with amount of attrition from the prior time point in parentheses. Baseline: 560/560 (0); Month 1: 548/552 (8); Month 2: 536/546 (5); Month 6: 528/539 (7); Month 12: 511/527 (8); Month 18: 499/516 (8); Month 24: 474/489 (13); Month 30: 325/342 (6); Month 36: 312/316 (1).

Table 3

Adjusted mean GCP scores over time by delirium status (N=560)[*]

Assessment Time	Adjusted mean GCP score (95% CI)		Difference in adjusted scores	
	Delirium (N=135)	No Delirium (N=431)	Net difference[†] (95% CI)	P value
Estimated mean performance				
Baseline (preoperative)	57.6 (57.1, 58.1)		–	–
After surgery				
1 month	55.9 (55.1, 56.6)	56.9 (56.4, 57.5)		
2 months	58.0 (57.4, 58.7)	58.0 (57.5, 58.5)		
6 months	57.9 (57.2, 58.5)	58.0 (57.4, 58.5)		
12 months	57.6 (56.9, 58.3)	57.9 (57.4, 58.5)		
18 months	57.3 (56.6, 58.0)	57.8 (57.3, 58.4)		
24 months	57.1 (56.3, 57.8)	57.8 (57.2, 58.3)		
30 months	56.8 (55.9, 57.6)	57.7 (57.1, 58.3)		
36 months	56.5 (55.6, 57.5)	57.6 (57.0, 58.3)		
Change from baseline				
1 month	−1.75 (−2.32, −1.17)	−0.71 (1.02, −0.400	−1.03 (−1.71, −0.360)	<0.01
2 months	0.47 (−0.06, 1.02)	0.47 (0.17, 0.77)	0.01 (−0.64, 0.65)	0.99
6 months	0.26 (−0.26, 0.79)	0.39 (0.10, 0.67)	−0.12 (−0.73, 0.49)	0.70
12 months	−0.05 (−0.56, 0.46)	0.26 (−0.01, 0.54)	−0.31 (−0.91, 0.28)	0.31
18 months	−0.36 (−0.89, 0.17)	0.14 (−0.15, 0.43)	−0.50 (−1.13, 0.12)	0.11
24 months	−0.67 (−1.26, −0.09)	0.02 (−0.30, 0.34)	−0.69 (−1.38, −0.01)	0.048
30 months	−0.99 (−1.65, −0.32)	−0.10 (−0.47, 0.26)	−0.88 (−1.65, −0.10)	0.03
36 months	−1.30 (−2.06, −0.54)	−0.23 (−0.65, 0.20)	−1.07 (−1.96, −0.18)	0.02

GCP = General Cognitive Performance

[*] Note: Values for GCP are corrected for learning effects and adjust for covariables, including age, female sex, nonwhite race, education, IQCODE score, Charlson, depression, and impairment in instrumental activities of daily living, surgery type, and baseline GCP score (see text for details). The number of participants completing each the interview/the number of participants eligible for the interview for each time point follows with amount of attrition from the prior time point in parentheses. Baseline: 560/560 (0); Month 1: 548/552 (8); Month 2: 536/546 (5); Month 6: 528/539 (7); Month 12: 511/527 (8); Month 18: 499/516 (8); Month 24: 474/489 (13); Month 30: 325/342 (6); Month 36: 312/316 (1).

[†] Net difference refers to the difference in score between those with (Delirium – No delirium)

Table 4

IQCODE Categories at Baseline and 36 Month Follow-up by Delirium Status[*]

IQCODE Category at baseline	IQCODE Score Category at 36 Month Follow-up								
	No post-operative delirium				Post-operative delirium				
	< 3.2	3.2, <3.7	3.7	Total	< 3.2	3.2, <3.7	3.7	Total	
< 3.2	147	29	5	181	39	16	5	60	
3.2, <3.7	7	11	6	24	1	4	9	14	
3.7	0	0	3	3	0	1	3	4	
Total	154	40	14	208	40	21	17	78	

[*] Proxy IQCODE ratings were available on 548 patients (98% of 560) at baseline and 290 patients (52% of 560) at 36 month follow-up.