

Association of Delirium With Long-term Cognitive Decline A Meta-analysis

Terry E. Goldberg, PhD; Chen Chen, MS; Yuanjia Wang, PhD; Eunice Jung, BS; Antoinette Swanson, BS; Caleb Ing, MD; Paul S. Garcia, MD, PhD; Robert A. Whittington, MD; Vivek Moitra, MD

 Supplemental content

IMPORTANCE Delirium is associated with increased hospital costs, health care complications, and increased mortality. Long-term consequences of delirium on cognition have not been synthesized and quantified via meta-analysis.

OBJECTIVE To determine if an episode of delirium was an independent risk factor for long-term cognitive decline, and if it was, whether it was causative or an epiphenomenon in already compromised individuals.

DATA SOURCES A systematic search in PubMed, Cochrane, and Embase was conducted from January 1, 1965, to December 31, 2018. A systematic review guided by Preferred Reporting Items for Systematic Reviews and Meta-analyses was conducted. Search terms included delirium AND postoperative cognitive dysfunction; delirium and cognitive decline; delirium AND dementia; and delirium AND memory.

STUDY SELECTION Inclusion criteria for studies included contrast between groups with delirium and without delirium; an objective continuous or binary measure of cognitive outcome; a final time point of 3 or more months after the delirium episode. The electronic search was conducted according to established methodologies and was executed on October 17, 2018.

DATA EXTRACTION AND SYNTHESIS Three authors extracted data on individual characteristics, study design, and outcome, followed by a second independent check on outcome measures. Effect sizes were calculated as Hedges *g*. If necessary, binary outcomes were also converted to *g*. Only a single effect size was calculated for each study.

MAIN OUTCOMES AND MEASURES The planned main outcome was magnitude of cognitive decline in Hedges *g* effect size in delirium groups when contrasted with groups that did not experience delirium.

RESULTS Of 1583 articles, data subjected from the 24 studies (including 3562 patients who experienced delirium and 6987 controls who did not) were included in a random-effects meta-analysis for pooled effect estimates and random-effects meta-regressions to identify sources of study variance. One study was excluded as an outlier. There was a significant association between delirium and long-term cognitive decline, as the estimated effect size (Hedges *g*) for 23 studies was 0.45 (95% CI, 0.34-0.57; *P* < .001). In all studies, the group that experienced delirium had worse cognition at the final time point. The *I*² measure of between-study variability in *g* was 0.81. A multivariable meta-regression suggested that duration of follow-up (longer with larger *gs*), number of covariates controlled (greater numbers were associated with smaller *gs*), and baseline cognitive matching (matching was associated with larger *gs*) were significant sources of variance. More specialized subgroup and meta-regressions were consistent with predictions that suggested that delirium may be a causative factor in cognitive decline.

CONCLUSIONS AND RELEVANCE In this meta-analysis, delirium was significantly associated with long-term cognitive decline in both surgical and nonsurgical patients.

Author Affiliations: Department of Psychiatry, Columbia University Irving Medical Center, New York, New York (Goldberg, Chen, Wang, Jung, Swanson); Department of Anesthesiology, Columbia University Irving Medical Center, New York, New York (Goldberg, Ing, Garcia, Whittington, Moitra); Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, New York (Wang).

Corresponding Author: Terry E. Goldberg, PhD, New York State Psychiatric Institute (NYSPI), 1051 Riverside Dr, #2409, New York, NY 10032 (teg2117@cumc.columbia.edu).

JAMA Neurol. 2020;77(11):1373-1381. doi:10.1001/jamaneurol.2020.2273
Published online July 13, 2020. Corrected on August 31, 2020.

Delirium is a common feature of postoperative recovery and is also observed in patients with critical illnesses, such as sepsis, respiratory failure, and cardiogenic shock.¹⁻³ It is characterized by an acute onset or fluctuating course, inattention, and either disorganized thought (manifesting as memory, language, and orientation difficulties) or altered level of consciousness.^{4,5} It generally arises 1 to 3 days after surgery.⁶ In the United States, delirium in postsurgical populations (cardiac, noncardiac, orthopedic) may range from 11% to 51%.^{1,2} In a recent large group of patients undergoing diverse surgeries, the rate of delirium was nearly 25%.⁷ In the intensive care unit, delirium rates have been reported to be as high as 82%.¹ Variability may be due to some extent to multiple risk factors (eg, age, hydration and nutritional status, sensory loss, preexisting cognitive impairment, alcohol or substance use, polypharmacy, multiple comorbidities, duration of surgery, and extent of surgical trauma).⁸ Delirium is associated with patient and family stress, increased hospital costs, increased duration of hospital stay, escalation of care, and increased mortality and morbidity including institutionalization.⁹ It is the most common surgical complication in adults older than 65 years.¹

In 2010, Witlox et al¹⁰ demonstrated that delirium was associated with increased mortality (hazard ratio, 1.96) in a meta-analysis based on 7 studies and 2957 patients. Based on 2 studies, they also suggested that delirium may be associated with increases in dementia. While observational studies have examined delirium-cognition associations since then, the literature on long-term cognitive decline following delirium has not been quantitatively synthesized. Furthermore, this literature comprised studies that used different cognitive outcomes, considered using continuous or binary outcomes (ie, dementia present/absent), examined delirium in the context of surgery or outside the surgical context (eg, intensive care unit), assessed delirium with different instruments, and investigated sample sizes that ranged from small (under 100) to large (over 100), thus making qualitative conclusions about delirium effects difficult.

Methods

General

We sought to determine if delirium was associated with increases in cognitive impairment or dementia incidence at least 3 months after such an episode by conducting a systematic review and meta-analysis of observational studies that met our inclusion criteria. We quantitatively controlled for factors listed above as well as others we deemed relevant. We further designed our analytic approach so that we might empirically address the question of whether delirium unmasks cognitive decline in those individuals who were already compromised and on a downward trajectory or whether delirium may potentially be causative (ie, precipitating).

Search

Our search strategy is illustrated in the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA)

Key Points

Question Is delirium associated with long-term cognitive decline?

Findings In this meta-analysis of 23 studies (after 1 outlier study was excluded), delirium was associated with long-term cognitive decline with a Hedges *g* effect size of 0.45. Effect sizes were similar between surgical and nonsurgical groups; meta-regressions were consistent with the hypothesis that delirium played a causative role.

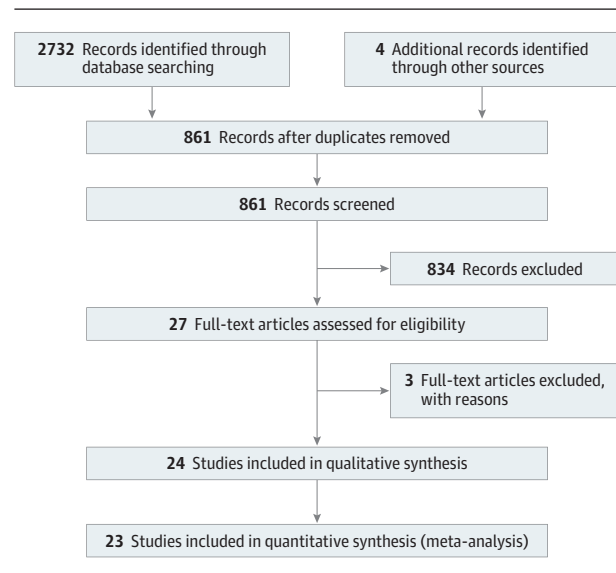
Meaning Delirium may be an independent risk factor for long-term cognitive decline in surgical and nonsurgical patient groups.

reporting guideline flowchart in **Figure 1**. It is described in detail in the eAppendix in the **Supplement**. PubMed, Cochrane, and Embase were searched from January 1, 1965, to December 31, 2018. Search terms included delirium AND postoperative cognitive dysfunction; delirium and cognitive decline; delirium AND dementia; and delirium AND memory. The electronic search was conducted according to established methodologies and was executed on October 17, 2018.

Inclusion Criteria

Diagnosis of delirium could be determined by validated delirium scales, usually the Confusion Assessment Method (CAM) (a screening method), *Diagnostic and Statistical Manual of Mental Disorders*, *International Classification of Diseases* diagnostic criteria, or validated case extraction methods. Measurement of cognition was required to be at least 3 months after the episode of delirium. It was required that cognition was measured by objective tests. The Informant Questionnaire on Cognitive Decline in the Elderly for baseline matching in studies, which were nonsurgical (necessitating an estimate of premorbid cognitive level), was accepted. The clinical diagnosis of dementia was also accepted as an outcome. Dementia diagnoses were based on medical record review, consensus conference using cognitive measures and function, or

Figure 1. PRISMA Flowchart



diagnostic examination. Most critically, a contrast between delirium present vs absent groups at end point was necessary for inclusion. All studies were cohort based.

Exclusion Criteria

Studies that used only self-report or informant report as an outcome were excluded. We excluded studies or subgroups within a study that measured delirium in the postanesthesia care unit because it could be confounded with emergence agitation¹¹ or cognition while patients experienced a delirium episode. We excluded studies that investigated delirium in the context of neurosurgery.

Extraction of Data

Study data were extracted by 2 of the authors (A.S. and E.J.).¹²⁻³⁵ The lead author (T.E.G.) reviewed all variable values and extracted all values needed to calculate Hedges g , which is used to measure the effect size of means between 2 groups in terms of pooled SD or odds ratio (OR). For continuous cognitive outcome variables, a single cognitive screening measure or a composite of multiple cognitive measures was used as the single outcome and then converted to an effect size. For dementia outcomes (ie, a binary outcome in which individuals were categorized as having dementia or not having dementia, an OR was extracted or calculated. Each study contributed a single effect size or OR. Only data from the last follow-up point were used. If multiple cognitive measures were included, only the composite was used. The extracted values were then subjected to meta-analytic procedures by 2 authors (Y.W. and C.C.).

Study Quality

We rated study quality using the Nottingham Ottawa Scale for observational studies (rated by T.G.).³⁶ The summary score was used in the analysis.

Statistical Approach

Effect sizes were calculated based on outcome means and SDs. The effect sizes were then converted to between-group effect size, Hedges g .³⁷ The combined effect size was measured using Hedges g weighted based on the inverse variances of each g . Analyses were conducted to test the primary hypothesis within subgroups.

Meta-regression was then used to investigate potential factors that affect how delirium and cognitive decline are associated. All analyses were conducted using random-effects models with the R metafor package (R Foundation).³⁸ All tests were 2-sided. Statistical significance was set at α less than .05.

Publication Bias

Funnel plots and Egger test of asymmetry were used to test potential publication bias and identify outliers.³⁹ They were also used to identify outliers.

Sensitivity Analyses

We conducted sensitivity analyses. The first was of 3 studies with overlapping samples (planned). The second analysis was done after removing a study based on funnel plot outlier inspection (unplanned).

Study Quality

We gave each article a single semiquantitative grade of quality, using Nottingham-Ottawa Scale criteria for observational designs. These values were also subjected to a meta-regression.

Heterogeneity Using I^2 and Meta-Regressions

We first measured I^2 in our final panel of 23 studies to assess between-study heterogeneity. We then analyzed the role of potential contributors to I^2 through meta-regression.

We elected to extract multiple variables that were quantitative in nature and that potentially might modulate or confound results and so account for heterogeneity. These included sample size, attrition rate, type of cognitive outcome, etc. Each of these variables was then used to assess the association with the meta-analytic outcome using meta-regressions in univariate analyses. This approach was then followed by a full multivariable analysis of those variables found to have $R^2 > 0$. All individual-level ($n = 3$) and study-level variables ($n = 8$) are specified below.

We also conducted a meta-regression in which we separated studies with continuous vs binary end points to determine if the magnitude of delirium associations to these measures was similar, and inferentially, to determine if the outcomes were thus measuring the same construct of cognitive decline over time.

Last, we sought to examine the possibility that cognitive decline after delirium was or might be an independent and causative agent in decline, rather than an epiphenomenon associated with preexisting cognitive impairments. Therefore, we conducted the following analyses: (1) performed a random-effects model analysis to determine g in studies that included only healthy individuals with no cognitive impairment at baseline (19 studies) (if delirium was an epiphenomenon, then g should be nonsignificant as neither group would decline); (2) conducted a similar analysis to the first step but in studies that comprised a high proportion of patients with cognitive impairment (3 studies) (if delirium was an epiphenomenon, then both delirium and nondelirium groups should decline equivalently with the contrast being nonsignificant, given that both groups were compromised and hence would have downward trajectories); and (3) examined the association of adjustment for baseline cognition with g by meta-regression. If g were found to be larger in those studies that adjusted for baseline cognition, then this would suggest that outcome differences were less likely to be the result of preexisting differences in cognition. However, if unmatched group studies were associated with larger g s, this would be in keeping with the hypothesis that preexisting baseline differences drove outcomes. These 3 analyses are based on the view that the more similar the groups of delirium present and delirium absent are at baseline (eg, both healthy, both with dementia, or baseline adjusted), the less likely that preexisting conditions can account for decline, if observed.

Results

Study Characteristics

There were differences in study-related methods in the 24 studies in Table 1.¹²⁻³⁵ Five articles used a cognitive

Table 1. Hedges *g* and SEs for Studies and Selected Individual- and Study-Level Variables

| Source | No. of patients at baseline | Age at baseline, y | Sex (male vs female), % | Delirium at baseline, % | Duration of follow-up, y | No. of covariates | Matching/adjustment at baseline | Attrition rate, % | <i>g</i> (SE) |
|---|-----------------------------|--------------------|-------------------------|-------------------------|--------------------------|-------------------|---------------------------------|-------------------|---------------|
| Bickel et al, ¹² 2008 | 180 | 73.8 | 31 | 23 | 3.2 | 4 | Yes | 9 | 2.04 (0.63) |
| Brown et al, ¹³ 2018 | 143 | 70 | 75 | 54 | 1.0 | 5 | Yes | 24 | 0.18 (0.17) |
| Davis et al, ¹⁴ 2012 | 558 | 88.3 | 20 | 13 | 10.0 | 3 | Yes | 1 | 1.05 (0.13) |
| Davis et al, ¹⁵ 2017 | 987 | 90 ^a | 31 | 28 | 5.2 | 4 | Yes | 0 | 0.22 (0.07) |
| Franck et al, ¹⁶ 2016 | 720 | 70 | 54 | 21 | 0.3 | 8 | Yes | 23 | 0.52 (0.25) |
| Fong et al, ¹⁷ 2009 | 408 | 73.9 | 43 | 18 | 1.4 | 8 | Yes | 0 | 0.47 (0.13) |
| Fong et al, ¹⁸ 2012 | 771 | 77.2 | 43 | 53 | 2.0 | 10 | Yes | 23 | 0.38 (0.13) |
| Girard et al, ¹⁹ 2018 | 964 | 62 | 60 | 69 | 1.0 | 8 | Yes | 35 | 0.15 (0.07) |
| Gross et al, ²⁰ 2012 | 263 | 78.3 | 43 | 56 | 3.2 | 8 | Yes | 69 | 0.45 (0.13) |
| Inouye et al, ²¹ 2016 | 560 | 76.7 | 42 | 24 | 3.0 | 10 | Yes | 44 | 0.23 (0.1) |
| Kat et al, ²² 2008 | 112 | 82.7 | NA | 63 | 2.5 | 4 | Yes | 56 | 0.63 (0.2) |
| Krogseth et al, ²³ 2016 | 287 | 85.7 | 26 | 70 | 0.4 | 1 | Yes | 30 | 0.33 (0.13) |
| Lingehall et al, ²⁴ 2017 | 114 | 76.5 | 69 | 53 | 5.0 | 13 | Yes | 6 | 0.75 (0.19) |
| Lundström et al, ²⁵ 2003 | 67 | 79.1 | 23 | 28 | 5.0 | 3 | Yes | 40 | 0.94 (0.4) |
| Mitchell et al, ²⁶ 2018 | 148 | 57 | 69 | 19 | 0.5 | 9 | No | 47 | 0.12 (0.21) |
| Neerland et al, ²⁷ 2017 | 387 | 84.5 | 24 | 32 | 0.8 | 4 | Yes | 31 | 1.05 (0.3) |
| Pandharipande et al, ²⁸ 2013 | 821 | 61 | 51 | 75 | 1.0 | 9 | Yes | 43 | 0.5 (0.03) |
| Rockwood et al, ²⁹ 1999 | 203 | 79 | 43 | 19 | 3.0 | 3 | Yes | 50 | 0.98 (0.33) |
| Rudolph et al, ³⁰ 2008 | 1161 | 68.7 | 53 | 9 | 0.3 | 3 | No | 19 | 0.14 (0.19) |
| Saczynski et al, ³¹ 2012 | 225 | 73 | 76 | 46 | 1.0 | 8 | Yes | 0 | 0.26 (0.13) |
| Sauër et al, ³² 2017 | 184 | 67.1 | 72 | 13 | 1.0 | 3 | Yes | 21 | 0.67 (0.08) |
| Sprung et al, ³³ 2017 | 1667 | 80 | 53 | 2 | 0.8 | 5 | Yes | 31 | 0.61 (0.28) |
| Vasunilashorn et al, ³⁴ 2018 | 560 | 76.7 | 42 | 56 | 3.0 | 9 | Yes | 11 | 0.25 (0.09) |
| Wacker et al, ³⁵ 2006 | 29 | 77.1 | 34 | 48 | NA | 0 | No | 0 | 1.03 (0.39) |

Abbreviation: NA, not applicable.

^a Mean age at baseline.

composite,^{13,18,19,21,26,28,32,34} 8 used a cognitive screening measure (ie, the Mini-Mental State Examination), and 8 used a categorical outcome.^{12,16,18,25,27,29,30,33} Sixteen studies used the CAM or CAM-intensive care unit as a delirium measure, 6 used other prospective measures,^{17,24-26,29} and 2 studies used a retrospective measure of delirium.^{14,15} Three studies included nonsurgical patients.^{15,19,28} Six studies did not adjust for baseline cognition.^{13,26,30,32,35} These differences were examined in meta-regressions.

One set of articles from a single research group had potentially near fully overlapping samples.^{17,18,20} We elected to include all articles after conducting sensitivity analyses. A second set of articles had a limited degree of overlap. Thus, Davis et al¹⁴ included more than 550 individuals from the Vantaa study. Their 2017 study included only 250 individuals from this study in addition to approximately 600 new individuals from 2 other large community samples.¹⁵ Therefore, we elected to include both articles.

For 2 studies, we elected to examine only subgroups. For the study by Franck et al,¹⁶ we elected to use the subgroup in which outcome was measured after postanesthesia care unit convalescence in keeping with other surgery-based studies and to reduce potential confounding of delirium with emergence agitation syndromes. For the study by Girard et al,¹⁹ we se-

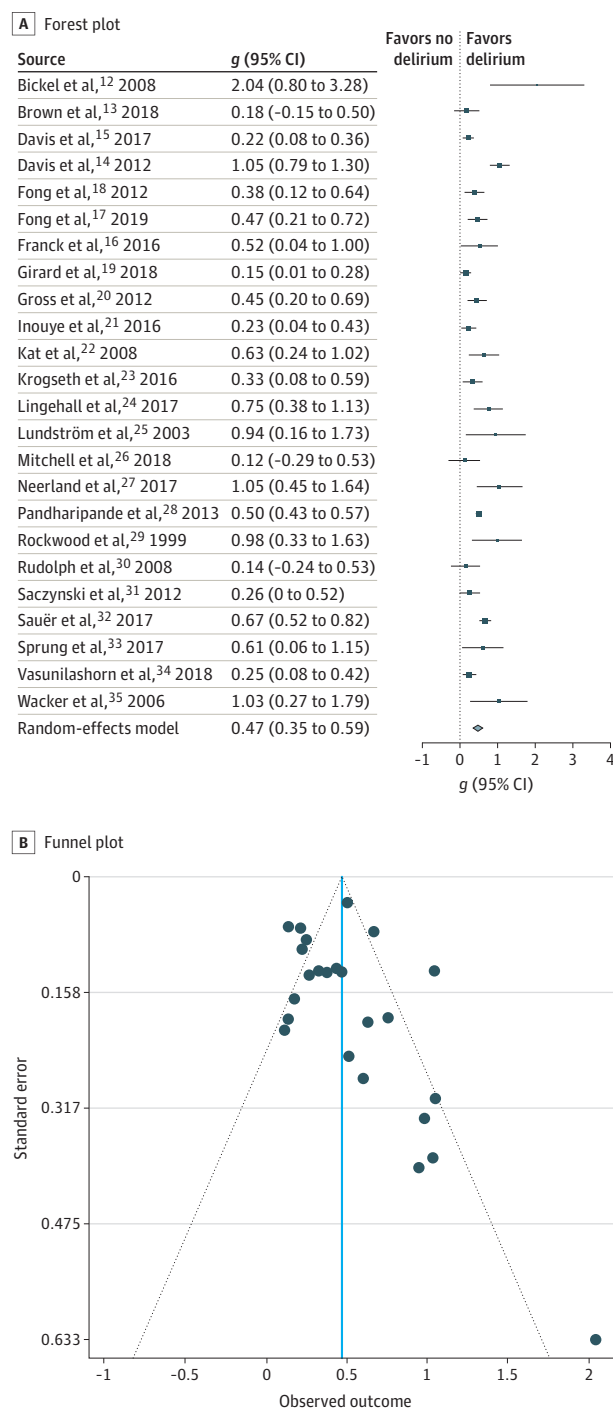
lected the largest delirium subgroup for which a cause was identified.

The mean (SD) study age in the panel was 75.4 (7.6) years. The mean (SD) number of individuals was 441.7 (352.7) for the sample. The mean (SD) percentage of individuals with delirium was 37.2 (10) in the 24 studies. The mean (SD) percentage of male individuals was 46.9 (10.0). The mean (SD) duration of follow-up after a delirium episode was 2.4 (2.3) years. The modal delirium rating instrument was CAM/CAM-intensive care unit. The modal cognitive instrument was the Mini-Mental State Examination.

g and ORs

We identified 3562 patients with delirium (delirium present) and 6987 patients without delirium (delirium absent) in 24 studies. We first examined the association of cognitive outcome *g* to delirium in our complete panel of studies in a random-effects model. In all instances, we used adjusted ORs and adjusted effect sizes that we derived from individual studies. As shown in Table 1,¹²⁻³⁵ every study demonstrated that the group that experienced delirium had worse neurocognitive outcomes at 3 or more months after the episode, with effect sizes ranging from large (>0.80) to small (0.15). The summary *g* of this meta-analysis of 24 studies was 0.47 as shown in Figure 2A.

Figure 2. Forest and Funnel Plots for All Studies



Thus, the effect size was medium and the result was highly significant ($g = 0.47$; 95% CI, 0.35-0.59; $P < .001$).

The effect size in g units was converted to an OR for further interpretative clarity. The conversion was based on assuming an underlying continuous trait for the binary outcome³⁷ and implemented using the Campbell collaborative program.⁴⁰ The resulting value was 2.30 (95% CI, 1.85-2.86). Thus, patients who experienced delirium had 2.30 times the odds to

demonstrate a given cognitive decline when contrasted with the odds of patients who did not experience delirium.

Funnel Plot Publication Bias and Outlier Analysis

Examination of the funnel plot in Figure 2B indicated some reduction of studies in the lower left quadrant. Additionally, an evident outlier was present.¹² The Egger asymmetry index was 3.15, and the P value was highly significant ($P = .002$) as shown in eFigure 3 in the Supplement. When the outlier was removed, asymmetry lessened and the index became marginally significant ($z = 2.37$; $P = .02$). This pattern suggests that smaller studies showing weaker associations may have been less likely to have been submitted or published, resulting in a skewed body of evidence.

Study Quality

Two studies were deemed of low quality by the Nottingham Ottawa Scale owing to lack of control over baseline cognition. Nevertheless, in meta-regressions, study quality was not a significant determinant of g ($R^2 = 0$; $P = .56$).

Sensitivity

One study¹² had an OR greater than 41, as noted above. This was an order of magnitude greater than any other study. It also was the most significant factor in determining funnel plot asymmetry. When converted to g , its value was more than 2. When we reconducted our meta-analysis without this study, the overall g declined trivially to 0.45 and remained significant (95% CI, 0.34-0.57; $P < .001$). This study was excluded from all further meta-regressions and analysis of delirium as a causative factor (eFigures 1-3 in the Supplement).

Three studies from a single investigator group had potentially various combinations of overlapping samples. (We contacted the lead author but did not receive a response.) We reconducted our meta-analysis after using only the largest sample (and thereby excluding the 2 smaller samples). The summary g did not change.

Continuous vs Binary Outcomes

We divided the sample into those using a categorical outcome (dementia or cognitive impairment present/absent) and those using continuous cognitive measures, as it could be viewed that they are reflecting qualitatively different types of outcome that might have differing associations with delirium. Thus, we contrasted studies with continuous outcomes ($n = 15$) and binary outcomes ($n = 8$) by meta-regression. The result was nonsignificant (difference in $g = 0.14$; 95% CI, -0.14 to 0.43). We also determined the association between delirium and these outcomes separately (ie, in 2 separate meta-analyses). In both types of outcome measures, the overall g was highly significant. Point estimates for these meta-regressions are in Table 2. Studies with binary outcomes ($g = 0.57$) had a larger effect size than did studies using continuous measures ($g = 0.42$).

I^2 Heterogeneity

We next examined the heterogeneity of our results. I^2 was high (0.81), which reflected high between-study variability in ef-

Table 2. Univariable Meta-Regressions for Individual- and Study-Level Variables

| Variable | β | z | P value | 95% CI | R^2 , % |
|---|---------|-------|---------|-------------------|-----------|
| Individual-level variables | | | | | |
| Age at baseline, y | 0.02 | 1.98 | .05 | 0.0002 to 0.03 | 3.28 |
| Education | 0.009 | 0.42 | .67 | -0.05 to 0.03 | 0 |
| Sex (male vs female) | -0.58 | -1.63 | .10 | -1.29 to 0.12 | 0 |
| Study-level variables | | | | | |
| No. | -0.0003 | -1.62 | .10 | -0.0006 to 0.0001 | 10.18 |
| Measure of cognitive (composite cognitive battery vs screening) | 0.16 | 1.07 | .28 | -0.13 to 0.46 | 0 |
| Measure of delirium (retrospective/other/Confusion Assessment Method) | 0.19 | 1.02 | .31 | -0.18 to 0.56 | 0 |
| Surgery (yes/no) | 0.04 | 0.03 | .77 | -0.23 to 0.31 | 0 |
| Duration of follow-up, y | 0.05 | 2.39 | .01 | 0.009 to 0.01 | 22.6 |
| Attrition rate, % | -0.10 | -0.32 | .75 | -0.69 to 0.49 | 0 |
| Outcome (binary vs continuous) | 0.14 | 0.99 | .32 | -0.14 to 0.43 | 0 |
| No. of covariates | -0.03 | 1.56 | .12 | -0.06 to 0.007 | 4.5 |

Table 3. Multivariable Meta-Regression^a

| Variable | β | P value | z |
|---------------------------------|---------|---------|-------|
| Age | -0.01 | .02 | -2.39 |
| Duration of follow-up | 0.07 | <.001 | 3.75 |
| No. of covariates | 0.04 | <.001 | -3.37 |
| Matching/adjustment at baseline | 0.40 | .003 | -3.02 |

^a All variables used had $R^2 > 0$ in univariable analyses. Sample size did not enter.

fect sizes. To identify sources of variance in g , we conducted a series of univariate meta-regressions that examined the proportion of variance that could be explained by various individual- or study-level factors in meta-regressions (Table 2).

Four of these study variables accounted for R^2 : duration of follow-up (longer durations were associated with larger g s), number of covariates (more covariates were associated with smaller g s), number of individuals (larger sample sizes were associated with smaller g s), and baseline matching of cognition (if present, it was associated with larger g s), and age (older samples were associated with larger g s). We then conducted a multivariable meta-regression using these variables as predictors of between-study variance. Four entered the regression (duration, age, covariate number, baseline matching/adjustment) with a total $R^2 = 0.86$ ($P < .001$) (Table 3). Thus, these were able to account for 70% of the I^2 .

The percentage of nonsurgical cases at baseline (due to sepsis, other types of infection, cardiac shock, respiratory failure, etc) was not a significant modulator of outcome (Table 2). This suggested that surgical cases, presumably associated with inflammation and anesthesia, did not yield outcomes that were significantly different from nonsurgical cases.

Delirium as an Epiphenomenon or as Causative

We first derived a summary g in only those studies examining patients without cognitive impairment. If delirium was an epiphenomenon, then neither the absence or presence of delirium should demonstrate decline. However, results for the summary g were positive ($g = 0.42$; 95% CI, 0.29-0.56; $P < .001$). Second, we examined studies with high proportions of patients with cognitive impairment. It should be the

case that if delirium-cognitive decline associations were an epiphenomenon, studies in Alzheimer disease-only groups should be nonsignificant because both delirium-present and delirium-absent groups would decline equivalently. This was not the case as the mean g of these studies was 0.44 and consistent with the overall summary g . Last, we examined baseline cognitive matching. Results indicated that studies that adjusted for baseline cognition had larger g s than those that did not (univariable estimate = 0.23; $P = .08$; 95% CI, -0.55 to 0.10; $R^2 = 5.11$; multivariable estimate $P = .003$ in Table 3).

These results, while consistent with the hypothesis that delirium may play a causative role in long-term cognitive decline after such an episode, are not conclusive, given difficulties with establishing causality in observational studies.

Discussion

In this meta-analysis of 24 trials enrolling 3562 patients who experienced delirium and 6987 controls who did not, delirium was significantly associated with cognitive impairment 3 or months or longer after the delirium episode. Remarkably, in every study included, the group that experienced delirium demonstrated worse cognitive performance 3 or months or longer after the episode. The summary g statistic was 0.46 with an equivalent OR = 2.30 for 23 studies (after 1 study was excluded as an outlier). The effect size can be considered medium and was not the result of study-related confounders or duplication. Such a medium effect size has been associated with clinically significant differences between groups (here delirium present and absent).^{41,42} We did not find

differences in cognitive outcomes between surgical and non-surgical studies by meta-regression, suggesting that the underlying pathophysiological events associated with delirium may be similar and speculatively and may be associated with inflammatory processes common to both contexts.⁴³ We also did not find significant differences between cognition treated as a continuous variable based on neurocognitive test scores or as a binary variable based on the presence or absence of dementia, suggesting that these 2 measures were monitoring the same underlying construct, namely cognitive decline and associated functional compromises.

Interstudy heterogeneity was present ($I^2 = 0.81$) and significant. Several factors accounted for I^2 in this quantitative synthesis. Studies of longer duration yielded greater differences; studies with more covariates tended to yield smaller differences. Studies without baseline cognitive matching yielded smaller differences. With multivariable meta-regression, we were able to account for 0.70 of the I^2 variance. Other potential confounders, including study quality, size, delirium measures, and outcome measure (binary or continuous, cognitive screening or cognitive composite), did not significantly account for variance. Moreover, the meta-regressions that were conducted to determine sources of I^2 suggested that variance was associated with primarily study design features, rather than study participant composition. By implication (longer) duration of follow-up, control over (multiple) covariates, and baseline matching for neurocognitive level might yield increasingly precise and valid results in future studies. Sensitivity analyses confirmed that neither studies that contained overlapping samples nor a study that was an outlier drove the results. Funnel plot asymmetry may reflect publication bias against studies with small g s and large SEs, although significance was marginal.

The view that postoperative delirium is not a risk factor for cognitive decline but rather an epiphenomenon of presurgical cognitive compromise is widespread. In this view, delirium is a biomarker for already compromised cerebral function that is then unmasked by surgery, anesthesia, shock, or infection. It implies the delirium group in each study would have declined more steeply even if delirium was not present. However, this perspective is not supported by our data when we devised analyses to adjudicate between the views. First, in cognitively intact well-characterized matched groups, the delirium group declined more. Second, in studies that explicitly examined groups with cognitive impairment (ie, dementia), the delirium group experienced greater decline. Third, those studies that adjusted for baseline cognition were associated with higher g s than those that did not, again suggesting that preexisting cognitive compromises were not a major driver of g . It might also be argued that unmeasured neuropathology might be worse in the delirium group and hence influence outcome. However, in the 1 study that directly addressed this, Davis et al¹⁵ did not find significant differences in Alzheimer disease histopathology on postmortem examinations. In general, the persistence of an unfavorable cognitive outcome years after the index episode also suggests that delirium is not marker of preexisting condition. Moreover, the associa-

tion of delirium with g could not be explained by such factors as attrition, sex, sample size, and to a degree, comorbidities. However, it remains a possibility that delirium may interact with preclinical disease to accelerate decline.

While our analyses were consistent with a causal hypothesis, causality cannot be confirmed because these studies were designed as observational in demonstrating associations. Findings based on prospective randomized clinical trials, albeit difficult to implement, might help to resolve this issue. Such trials would involve manipulation of delirium; the outcome would be long-term cognition.

Limitations

We recognize that the observational studies used in the meta-analysis here were not designed to adjudicate between delirium as a causative factor in consequent cognitive decline or an epiphenomenon. Another possible limitation relates to the large I^2 measure of heterogeneity among studies. Three variables accounted for 70% of the variance in plausible directions. However, the relatively small number of studies in some of the meta-regressions may have affected power to identify other sources of heterogeneity. Similarly, the negative findings in our meta-regressions relating to causation may also be limited by relatively small study panel sizes. While differences in the way dementia was diagnosed and the type of cognitive variable used as an outcome measure may have added undue variance to the analysis, meta-regressions did not find significant differences among them in terms of their effect on g . Last, we were unable to examine the interaction of delirium with such potential accelerators of biological aging as frailty.⁴⁴ This will be an important area of investigation in the future.

Conclusions

From a public health standpoint, delirium represents a clear target to improve population health. Delirium is robustly associated with increases in mortality and, as shown here, long-term cognitive decline. Assuming that delirium complicates stay in about 20% of those 11.8 million individuals older than 65 years who are hospitalized per year, costs attributable to delirium may be between \$143 billion and \$152 billion owing to longer hospital stays, outpatient visits, nursing home care, and rehabilitation.⁹ Reduction in delirium incidence would have salutary effects in this demographic. Indeed, several approaches to reducing delirium incidence have been implemented. These range from systematic environmental approaches (eg, orientation, sleep-wake cycle regularity, access to glasses and hearing aids, promotion of mobility) to those involving medication, including reduction in the use of antipsychotics.^{45,46} Other approaches that are prophylactic (eg, use of anti-inflammatory drugs, cognitive-enhancing drugs, or cognitive training) have been less used.⁴⁵ We suggest that additional research in this area may rather rapidly yield reductions in delirium and *pari passu*, postdelirium cognitive decline.

ARTICLE INFORMATION

Accepted for Publication: April 8, 2020.

Published Online: July 13, 2020.
doi:10.1001/jamaneurol.2020.2273

Correction: This article was corrected on August 31, 2020, to fix errors in the byline and reference list.

Author Contributions: Drs Goldberg and Wang had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Goldberg, Chen, Ing, Garcia, Whittington, Moitra.

Acquisition, analysis, or interpretation of data: Goldberg, Chen, Wang, Jung, Swanson, Ing, Garcia, Moitra.

Drafting of the manuscript: Goldberg, Chen, Wang, Jung, Swanson, Garcia, Whittington, Moitra.

Critical revision of the manuscript for important intellectual content: Goldberg, Chen, Wang, Jung, Ing, Garcia, Whittington, Moitra.

Statistical analysis: Goldberg, Chen, Wang, Jung.

Obtained funding: Chen, Whittington.

Administrative, technical, or material support: Chen, Jung, Swanson, Ing, Garcia, Moitra.

Supervision: Chen, Moitra.

Conflict of Interest Disclosures: Dr Garcia reports grants from James S. McDonnell Foundation during the conduct of the study. Dr Whittington reports grants from National Institute of General Medical Sciences during the conduct of the study and personal fees from Anesthesia and Analgesia outside the submitted work. No other disclosures were reported.

Funding/Support: This study was supported by the Departments of Anesthesiology and Biostatistics, Columbia University Irving Medical Center.

Role of the Funder/Sponsor: The funder had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

REFERENCES

- Inouye SK, Robinson T, Blaum C, et al; American Geriatrics Society Expert Panel on Postoperative Delirium in Older Adults. Postoperative delirium in older adults: best practice statement from the American Geriatrics Society. *J Am Coll Surg*. 2015;220(2):136-48. doi:10.1016/j.jamcollsurg.2014.10.019
- Hsieh TT, Inouye SK, Oh ES. Delirium in the elderly. *Psychiatr Clin North Am*. 2018;41(1):1-17. doi:10.1016/j.psc.2017.10.001
- Slooter AJ, Van De Leur RR, Zaal IJ. Delirium in critically ill patients. *Handb Clin Neurol*. 2017;141:449-466. doi:10.1016/B978-0-444-63599-0.00025-9
- Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). *Crit Care Med*. 2001;29(7):1370-1379. doi:10.1097/00003246-200107000-00012
- Inouye SK, van Dyck CH, Alessi CA, Balkin S, Siegel AP, Horwitz RI. Clarifying confusion: the confusion assessment method a new method for detection of delirium. *Ann Intern Med*. 1990;113(12):941-948. doi:10.7326/0003-4819-113-12-941
- Lloyd DG, Ma D, Vizcaychipi MP. Cognitive decline after anaesthesia and critical care. *BJA Educ*. 2012;12(3):105-109. doi:10.1093/bjaeaccp/mks004
- Wildes TS, Mickel AM, Ben Abdallah A, et al; ENGAGES Research Group. Effect of electroencephalography-guided anesthetic administration on postoperative delirium among older adults undergoing major surgery: the ENGAGES Randomized Clinical Trial. *JAMA*. 2019;321(5):473-483. doi:10.1001/jama.2018.22005
- Zaal IJ, Devlin JW, Peelen LM, Slooter AJ. A systematic review of risk factors for delirium in the ICU. *Crit Care Med*. 2015;43(1):40-47. doi:10.1097/CCM.0000000000000625
- Leslie DL, Inouye SK. The importance of delirium: economic and societal costs. *J Am Geriatr Soc*. 2011;59(suppl 2):S241-S243. doi:10.1111/j.1532-5415.2011.03671.x
- Witlox J, Eurelings LS, de Jonghe JF, Kalisvaart KJ, Eikelenboom P, van Gool WA. Delirium in elderly patients and the risk of postdischarge mortality, institutionalization, and dementia: a meta-analysis. *JAMA*. 2010;304(4):443-451. doi:10.1001/jama.2010.1013
- Safavynia SA, Arora S, Pryor KO, Garcia PS. An update on postoperative delirium: clinical features, neuropathogenesis, and perioperative management. *Curr Anesthesiol Rep*. 2018;8(3):252-262. doi:10.1007/s40140-018-0282-5
- Bickel H, Grading R, Kochs E, Förstl H. High risk of cognitive and functional decline after postoperative delirium: a three-year prospective study. *Dement Geriatr Cogn Disord*. 2008;26(1):26-31. doi:10.1159/000140804
- Brown CH IV, Probert J, Healy R, et al. Cognitive decline after delirium in patients undergoing cardiac surgery. *Anesthesiology*. 2018;129(3):406-416. doi:10.1097/ALN.0000000000002253
- Davis DH, Muniz Terrera G, Keage H, et al. Delirium is a strong risk factor for dementia in the oldest-old: a population-based cohort study. *Brain*. 2012;135(pt 9):2809-2816. doi:10.1093/brain/awb190
- Davis DH, Muniz-Terrera G, Keage HA, et al; Epidemiological Clinicopathological Studies in Europe (ECLIPSE) Collaborative Members. Association of delirium with cognitive decline in late life: a neuropathologic study of 3 population-based cohort studies. *JAMA Psychiatry*. 2017;74(3):244-251. doi:10.1001/jamapsychiatry.2016.3423
- Franck M, Nerlich K, Neuner B, et al. No convincing association between post-operative delirium and post-operative cognitive dysfunction: a secondary analysis. *Acta Anaesthesiol Scand*. 2016;60(10):1404-1414. doi:10.1111/aas.12779
- Fong TG, Jones RN, Shi P, et al. Delirium accelerates cognitive decline in Alzheimer disease. *Neurology*. 2009;72(18):1570-1575. doi:10.1212/WNL.0b013e3181a4129a
- Fong TG, Jones RN, Marcantonio ER, et al. Adverse outcomes after hospitalization and delirium in persons with Alzheimer disease. *Ann Intern Med*. 2012;156(12):848-856. doi:10.7326/0003-4819-156-12-201206190-00005
- Girard TD, Thompson JL, Pandharipande PP, et al. Clinical phenotypes of delirium during critical illness and severity of subsequent long-term cognitive impairment: a prospective cohort study. *Lancet Respir Med*. 2018;6(3):213-222. doi:10.1016/S2213-2600(18)30062-6
- 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. doi:10.1001/archinternmed.2012.3203
- Inouye SK, Marcantonio ER, Kosar CM, et al. The short-term and long-term relationship between delirium and cognitive trajectory in older surgical patients. *Alzheimers Dement*. 2016;12(7):766-775. doi:10.1016/j.jalz.2016.03.005
- Kat MG, Vreeswijk R, de Jonghe JF, et al. Long-term cognitive outcome of delirium in elderly hip surgery patients: a prospective matched controlled study over two and a half years. *Dement Geriatr Cogn Disord*. 2008;26(1):1-8. doi:10.1159/000140611
- Krogseth M, Watne LO, Juliebø V, et al. Delirium is a risk factor for further cognitive decline in cognitively impaired hip fracture patients. *Arch Gerontol Geriatr*. 2016;64:38-44. doi:10.1016/j.archger.2015.12.004
- Lingehall HC, Smulter NS, Lindahl E, et al. Preoperative cognitive performance and postoperative delirium are independently associated with future dementia in older people who have undergone cardiac surgery: a longitudinal cohort study. *Crit Care Med*. 2017;45(8):1295-1303. doi:10.1097/CCM.0000000000002483
- Lundström M, Edlund A, Bucht G, Karlsson S, Gustafson Y. Dementia after delirium in patients with femoral neck fractures. *J Am Geriatr Soc*. 2003;51(7):1002-1006. doi:10.1046/j.1365-2389.2003.51315.x
- Mitchell ML, Shum DHK, Mihala G, Murfield JE, Aitken LM. Long-term cognitive impairment and delirium in intensive care: a prospective cohort study. *Aust Crit Care*. 2018;31(4):204-211. doi:10.1016/j.aucc.2017.07.002
- Neerland BE, Krogseth M, Juliebø V, et al. Perioperative hemodynamics and risk for delirium and new onset dementia in hip fracture patients: a prospective follow-up study. *PLoS One*. 2017;12(7):e0180641. doi:10.1371/journal.pone.0180641
- Pandharipande PP, Girard TD, Jackson JC, et al; BRAIN-ICU Study Investigators. Long-term cognitive impairment after critical illness. *N Engl J Med*. 2013;369(14):1306-1316. doi:10.1056/NEJMoa1301372
- Rockwood K, Cosway S, Carver D, Jarrett P, Stadnyk K, Fisk J. The risk of dementia and death after delirium. *Age Ageing*. 1999;28(6):551-556. doi:10.1093/ageing/28.6.551
- Rudolph JL, Marcantonio ER, Culley DJ, et al. Delirium is associated with early postoperative cognitive dysfunction. *Anaesthesia*. 2008;63(9):941-947. doi:10.1111/j.1365-2044.2008.05523.x
- Saczynski JS, Marcantonio ER, Quach L, et al. Cognitive trajectories after postoperative delirium. *N Engl J Med*. 2012;367(1):30-39. doi:10.1056/NEJMoa1112923
- Sauër AC, Veldhuijzen DS, Ottens TH, Slooter AJC, Kalkman CJ, van Dijk D. Association between delirium and cognitive change after cardiac surgery.

Br J Anaesth. 2017;119(2):308-315. doi:10.1093/bja/aex053

33. Sprung J, Roberts RO, Weingarten TN, et al. Postoperative delirium in elderly patients is associated with subsequent cognitive impairment. *Br J Anaesth*. 2017;119(2):316-323. doi:10.1093/bja/aex130

34. Vasunilashorn SM, Fong TG, Albuquerque A, et al. Delirium severity post-surgery and its relationship with long-term cognitive decline in a cohort of patients without dementia. *J Alzheimers Dis*. 2018;61(1):347-358. doi:10.3233/JAD-170288

35. Wacker P, Nunes PV, Cabrita H, Forlenza OV. Post-operative delirium is associated with poor cognitive outcome and dementia. *Dement Geriatr Cogn Disord*. 2006;21(4):221-227. doi:10.1159/000091022

36. Wells GA, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. The Ottawa Hospital Research Institute. Accessed June 5, 2020. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp

37. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. *Introduction to Meta-analysis*. John Wiley & Sons; 2011.

38. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Software*. 2010;36(3):48. doi:10.18637/jss.v036.i03

39. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634. doi:10.1136/bmj.315.7109.629

40. Wilson DB. Practical meta-analysis effect size calculator. Accessed June 5, 2020. <https://campbellcollaboration.org/escalc/html/EffectSizeCalculator-Home.php>

41. Norman GR, Sloan JA, Wyrwich KW. The truly remarkable universality of half a standard deviation: confirmation through another look. *Expert Rev Pharmacoecon Outcomes Res*. 2004;4(5):581-585. doi:10.1586/14737167.4.5.581

42. Farivar SS, Liu H, Hays RD. Half standard deviation estimate of the minimally important difference in HRQOL scores? *Expert Rev*

Pharmacoecon Outcomes Res. 2004;4(5):515-523. doi:10.1586/14737167.4.5.515

43. Whittington RA, Planel E, Terrando N. Impaired resolution of inflammation in Alzheimer's disease: a review. *Front Immunol*. 2017;8:1464. doi:10.3389/fimmu.2017.01464

44. Wallace LMK, Theou O, Godin J, Andrew MK, Bennett DA, Rockwood K. Investigation of frailty as a moderator of the relationship between neuropathology and dementia in Alzheimer's disease: a cross-sectional analysis of data from the Rush Memory and Aging Project. *Lancet Neurol*. 2019;18(2):177-184. doi:10.1016/S1474-4422(18)30371-5

45. Inouye SK, Westendorp RG, Saczynski JS. Delirium in elderly people. *Lancet*. 2014;383(9920):911-922. doi:10.1016/S0140-6736(13)60688-1

46. Wu YC, Tseng PT, Tu YK, et al. Association of delirium response and safety of pharmacological interventions for the management and prevention of delirium: a network meta-analysis. *JAMA Psychiatry*. 2019;76(5):526-535. doi:10.1001/jamapsychiatry.2018.4365