The Association of Preoperative Frailty and Postoperative Delirium: A Meta-Analysis

Thomas J Gracie, B.S.[1] **[Medical Student]**, **Christine Caufield-Noll, M.L.I.S.**[2] **[Manager]**, **Nae-Yuh Wang, PhD**[3] **[Associate Professor]**, **Frederick Sieber, M.D.**[4] **[Professor]**

[1]Johns Hopkins University School of Medicine, Baltimore, Maryland [2]Library Services, Harrison Library, Johns Hopkins Bayview Medical Center, Baltimore, Maryland [3]Departments of Medicine (General Internal Medicine), Biostatistics, and Epidemiology, Johns Hopkins University, Baltimore, Maryland [4]Department of Anesthesiology and Critical Care Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland

Abstract

Background: Both frailty and postoperative delirium (POD) are common in elective surgical patients 65 years of age and older. However, the association between preoperative frailty and POD remains difficult to characterize owing to the large number of frailty and POD assessment tools employed in the literature, only a few of which are validated. Furthermore, some validated frailty tools fail to provide clear score cutoffs for distinguishing frail and nonfrail patients. We performed a meta-analysis to estimate the relationship between preoperative frailty and POD.

Methods: We searched several major databases for articles that investigated the relationship between preoperative frailty and POD in patients with mean age ≥ 65 who were undergoing elective, non-emergent inpatient surgery. Inclusion criteria included articles published in English no earlier than 1999. Both preoperative frailty and POD must have been measured with validated tools using clear cutoff scores for frailty and delirium. Articles were selected and data extracted independently by two researchers. Risk of bias (ROBINS-I) and presence of confounders were summarized. Odds ratios (OR) for POD associated with frailty relative to nonfrailty were computed with adjusted ORs when available. Original estimates were pooled by random-effects analysis. Statistical significance was set at 2-sided $P<0.05$.

Results: Nine studies qualified for meta-analysis. The Fried score or a modified version of it was used in five studies. Frailty prevalence ranged from 18.6% to 56%. Delirium was assessed with the

Corresponding Author: Frederick Sieber, M.D., Professor, Department of Anesthesiology and Critical Care Medicine, Johns Hopkins Bayview Medical Center, 4940 Eastern Avenue, Baltimore, MD 21224, 410-550-0942, fsieber1@jhmi.edu.

Confusion Assessment Method (CAM) or CAM-ICU in seven studies, Delirium Observation Scale in one study, and Intensive Care Delirium Screening Checklist in one study. The incidence of POD ranged from 7% to 56%. ROBINS-I risk of bias was low in one study, moderate in four studies, serious in three studies, and critical in one study. Random effects analysis (n = 794) of the OR for POD in frail versus nonfrail patients based on adjusted OR estimates was significant with an OR of 2.14 and a 95% confidence interval of 1.43–3.19. The I^2 value was low range at 5.5, suggesting small variability from random effects. Funnel-plot analysis did not definitively support either the presence or absence of publication bias.

Conclusions: This meta-analysis provides evidence for a significant association between preoperative frailty and POD in elective surgical patients aged 65 or older.

Introduction

Frailty—defined as a state of vulnerability to physiologic insults—has been characterized as both a phenotype and as a state of accumulated deficits.[1] It is common in community dwelling older persons, with a prevalence of 4.0–17.0%.[2] Its prevalence has been estimated to be even higher in older adult surgical patients,[3] and is associated with poorer postoperative outcomes, including increased mortality, length of stay, discharge to skilled care, readmission, and complications.[4] Postoperative delirium (POD) is a particularly common complication suspected to be associated with preoperative frailty.[5,6] It is important to assess for POD risk because of its independent association with increased length of stay, complications, institutionalization and mortality.[7–9] Quantification of the relationship between preoperative frailty and POD could enhance perioperative decision making, potentially mitigating unnecessary morbidity and mortality associated with POD.

One challenge in studying this relationship is the wide variety of frailty assessment instruments, only a few of which have been validated.[10] Furthermore, frailty tools are divided into categorical scales and continuous indices.[1] Few continuous indices offer a consistent, validated cutoff point for categorizing patients as either nonfrail or frail.[11] In contrast, categorical tools for frailty phenotype assessment—such as the Fried index[12]— usually consist of relatively fewer items and offer clear, well-validated score cutoffs, allowing for consistent and efficient categorization of patients as nonfrail or frail in clinical settings. These phenotypic tools are suitable from a clinician's standpoint as they facilitate quick, definitive judgements of frailty based on reproducible data. Similarly, although many POD assessment tools exist, only a few have been well-validated in the literature.[13]

This meta-analysis builds upon previously published work in this field[6] and is of relevance to anesthesiologists because of its narrow focus on elective surgeries. In addition, this meta-analysis used several specific strategies to better quantify the relationship between preoperative frailty and POD. First, we included only studies that used validated tools for both preoperative frailty and POD. We focused on frailty and delirium instruments that specify frailty status or presence of POD using consistent and validated cutoff points— metrics practical for the anesthesiologist's or surgeon's clinical practice. Second, given that nonrandomized studies of interventions constitute the literature on this topic, we carefully

assessed each eligible study for confounding variables and methodological limitations that could bias the analysis results and their conclusions.

METHODS

Literature Search

Our meta-analysis was conducted in accordance with PRISMA recommendations[14] using the Covidence web platform. The databases searched included PubMed, EMBASE, Scopus, and Cochrane Library. The search algorithm retrieved articles that included either a MESH term or text-word from each of the following three tiers. The first tier of keywords included confusion, delirium, acute confusion, organic brain syndrome, acute encephalopathy, cognitive dysfunction, cognitive impair*, cogniti* disorder*, postoperative complication*, postoperative medical complication*, and altered mental status. The second tier of keywords included postoperative care, postoperative complications, postoperative, "surgical procedures, operative," postsurgical, after surgery, and following surgery. The third tier of keywords included frail elderly, frailty, frail*, geriatric assessment, gait speed, gait analysis, Fried index, Edmonton Frail Scale, hand strength, grip strength, sarcopenia, timed up and go, and cumulative index rating scale. An asterisk placed after the keyword captured every possible variation of the preceding word or portion of a word.

The search algorithm was supplemented with a manual search of relevant literature. Specifically, all studies in the full-text screening category were entered into Scopus, and articles citing and cited by these studies were manually reviewed and the relevant studies uploaded for title and abstract review. Furthermore, review articles generated by the database searches as well as by citation search were manually examined; those with similar inclusion/ exclusion criteria were mined for studies that could be included in title and abstract review.

Article Selection

Inclusion criteria consisted of studies that investigated the relationship between preoperative frailty and POD in patients with a mean age of 65 years and older undergoing elective, non-emergent surgery in an inpatient setting. Articles must have been available in an English full-text version and should have been published no earlier than 1999, prior to which frailty had not been characterized in the literature with consistent metrics.[15] The search included articles until December 31, 2019. Both preoperative frailty and postoperative delirium must have been measured with well-validated tools as defined by the literature.[10,13] Because false-negative rates of delirium are high when assessed by chart review,[16] assessments of frailty and delirium must have been conducted in person. Validated frailty tools that did not offer a previously validated cutoff point for categorization of patients as either nonfrail or frail were excluded. Furthermore, studies that used validated frailty tools must have utilized all components of those tools, rather than just a subset. Other exclusion criteria comprised study populations with any reported history of alcohol use disorder or other substance use disorders, psychiatric illness, head trauma, neurological surgery, or stroke. Case reports, letters, oral presentations, review articles, and abstract-only publications were also excluded.

Two researchers (FS, TG) independently conducted the title and abstract review, full-text review, bias assessment, and data extraction, and conflicts were resolved by consensus to ensure consistency and inclusiveness for each step of the PRISMA algorithm.[17] Extracted data included the following: study authors; location; funding source; type of surgical procedure performed; number of patients; central measure of age; proportion female; measurement tools for frailty; number and percentage per frailty category; preoperative cognitive status; cognitive status assessment tool; burden of comorbidities; comorbidity measurement tool;[18] the odds ratio (OR), relative risk, or hazard ratio for delirium; and variables used for adjustment of the frailty-delirium relationship analysis. Bias was assessed with the ROBINS-I tool for non-randomized studies, per Cochrane recommendations.[19] Importantly, ROBINS-I bias assessments are made based on the comparison between a given study and a theoretical randomized-controlled trial with ideal design for the study question —the latter of which represents the standard for a "low-risk" study. Given that phenotypic frailty is often visually apparent, a blinded study may be impossible, even in theory. For this reason, the "low-risk standard" for bias assessment was defined as an ideal observational study instead.

Statistical Analysis

Studies that distinguished nonfrailty and frailty with clear and validated cutoff scoring were assessed through meta-analysis. Frailty instruments that described a "prefrail" group were also accepted. The primary measure of interest was the summary OR for POD in frail versus nonfrail patients. Furthermore, several exploratory analyses were carried out for studies that described prefrail groups. First, we estimated the OR for POD in prefrail versus nonfrail patients. Additionally, based on the assumption that frailty instruments giving a binary outcome (frail/nonfrail) would classify prefrail patients as nonfrail, we combined prefrail and nonfrail patients together, and then compared this pool to frail patients with regards to the OR for POD.

Adjusted ORs were used in the meta-analyses when available. We used raw data to compute the unadjusted OR in studies that did not report OR of postoperative delirium associated with frailty. We used Firth's penalized likelihood approach to estimate unadjusted OR while correcting for bias due to small sample size and sparse outcome when applicable. Random effects models were used to pool the estimates obtained from the included studies. Forest plots were used to display variation in OR estimates that expressed the association between frailty and delirium. Heterogeneity was tested by using Q and I^2 statistics. We also calculated E-value to evaluate the degree of unmeasured confounding necessary to attenuate the observed POD and frailty association to null (i.e. OR = 1).[20] We converted OR to relative risk (RR) using the optimal minimax approximation[21] for E-value calculation and then converted the resulting E-value back to OR for results reporting and discussion. Furthermore, studies in the meta-analysis were assessed for publication or other reporting biases on a funnel plot. For all tests, significance was set at $P<0.05$, and 95% confidence intervals (CIs) are presented. Corresponding calculations and graphical visualizations of data were carried out in SAS (Cary, NC).

Power Evaluation

For the analysis of frail versus nonfrail patients, we assumed a POD incidence of 15.0% in the nonfrail group with 9 studies having an average sample size of 45 patients per group. The meta-analysis would have 80.2% power to detect an OR of 2.38 (POD incidence of 29.6% in the frail group) with low heterogeneity, but only 62.9% power for the same OR with moderate heterogeneity, and 36.7% power with high heterogeneity among the included studies. For prefrail versus nonfrail patients, we assumed the same POD incidence and sample size per group, but with 4 studies only, the analysis would have 80.4% power to detect an OR of 3.31 (or POD incidence of 36.9% in the prefrail group) with low heterogeneity, 63.1% power with moderate heterogeneity, and 36.8% power for high heterogeneity. For comparing frail versus prefrail and nonfrail patients combined, we assumed a POD incidence of 18.0% in the combined group, 9 studies with average sample size of 55 patients per group, this analysis would have 80.2% power to detect an OR of 2.08 (POD incidence of 31.3% in the frail group) with low heterogeneity, 63.0% power with moderate heterogeneity, and 36.7% power for high heterogeneity.

RESULTS

Figure 1 shows the PRISMA diagram. The initial search produced 1272 articles, from which 730 duplicates were removed. Of the remaining 542 articles, 509 were excluded for the following reasons: 230 were deemed irrelevant because their study designs failed to investigate either frailty or delirium, or failed to include an operative intervention;131 failed to measure frailty with a validated tool; 75 were review articles; 30 were conducted in an emergency or nonelective setting; 27 failed to measure delirium with a validated tool; eight were case reports; five had a mean age below 65; and three were notes or letters.

After the screening stage, 33 studies underwent full-text review. Of these, 11 failed to measure delirium with a validated tool, four were abstracts or reviews, three failed to measure frailty with a validated tool, two were excluded for the context of emergency surgery, two included a high prevalence of alcohol use disorder in the study population, one contained overlap of the study participants with another study included in our synthesis, and one was excluded because we were unable to access the full-text article. Thus, nine studies were eligible for meta-analysis.

Characteristics of included studies are reported in Table 1. In total, our systematic review comprised a population of 1008 participants with a mean age of 74 years, of which 42% were female. Three studies were European,[22–24] four were North American,[25–28] and two were Asian.[29,30] Of the nine studies in the meta-analysis, two did not adjust for confounders[24,30] and one adjusted for EUROSCORE only.[25]

Six studies included American College of Cardiology/American Heart Association-defined high-risk surgeries such as open or transcatheter aortic valve repair, coronary artery surgery, or vascular procedures.[22–25,27,29] Three studies featured intermediate-risk procedures such as general surgery/laparotomy, thoracic surgery, spine surgery, or arthroplasty and total knee replacement (See Supplemental Table 1).[26,28,30] The Fried score or one of its variants—such as the Modified Fried Score or Japanese version of the Cardiovascular Health Study (J-CHS)

—was used in five of the studies.[25–27,29,30] Each of the remaining four studies used a different frailty tool. Overall, the prevalence of frailty across studies ranged between 18.6% and 56%. Four studies described a subset of patients with prefrailty, the prevalence of which ranged from 27% to 57%.[22, 26–28] Compared with the array of tools used to measure preoperative frailty, the instruments used for POD assessment were much less variable. Seven studies used the Confusion Assessment Method (CAM) or CAM for the intensive care unit,[22,24–28,30] one used the Delirium Observation Scale (DOS),[23] and one used the Intensive Care Delirium Screening Checklist (ICDSC).[29] The incidence of postoperative delirium in the studies ranged from 7% to 56%.[22–30]

Quality of Methodology

Consensus ROBINS-I judgments of bias are summarized in Table 2, and reasons for bias are documented in Supplemental Table 2. With regard to risk for overall bias, one study was judged to have low risk ,[28] four studies were judged to have moderate risk,[22,26,29,30] three studies were judged to have serious risk,[23,24,27] and one study was judged to have critical risk.[25] In the domain of confounding, three studies had low risk of bias.[26,28,29] as they adjusted for the important confounders of age and cognition. One study[25] was classified as having critical risk for bias in the confounding domain as several delirium confounders differed between frail and non-frail groups but were not adjusted for in the final analysis. Two studies were at moderate risk for bias because either age or cognition, but not both, were adjusted for in analysis.[22,27] Serious risk was evident in one study not providing sufficient information for bias assessment of confounders,[23] and another unadjusted study with baseline cognitive differences.[24] Table 3 compares confounding variables between frail and nonfrail groups in studies for which data were available.

The domain of participant selection was the most common source of risk for bias; three studies were judged as having moderate risk of bias,[25,26,29] often due to measures that excluded patients who might be disproportionately frail. A fourth study was judged to have a serious risk of bias in this domain[27] because it selectively included patients from two parent studies with different exclusion criteria. All studies were at low risk in their classification of interventions. Missing data were a source of moderate risk for bias in two studies.[22,24] The risk was due to either loss of data from patient deaths or failure to assess for delirium on weekends. A third study was found to have a serious risk for bias because the delirium follow-up period was unclear.[23] In the domain of measurement outcome bias, one study[27] was judged to have moderate bias owing to the occasional use of chart review for delirium screening, whereas another study[23] was judged to have a serious risk of bias because a "filtering" step was used in the screening process, potentially underestimating delirium incidence. With regard to risk of bias in selection of the reported result, one study was classified as being at moderate risk of bias because it failed to address patient deaths in analysis.[22]

Meta-analysis of Frailty

Random effects analysis (n = 794) of the OR for POD in frail versus nonfrail patients based on adjusted OR estimates was significant, with an OR of 2.14 and a 95% CI of 1.43–3.19 (Figure 2). The I^2 value was in the low range at 5.5, suggesting small variability as a result

of random effects. The E-value to attenuate the observed OR of 2.14 between POD and frailty to 1, and to render the observed association statistically non-significant (i.e. attenuate the lower 95% CI bound of 1.43 to 1) was 4.08 and 2.36, respectively. Funnel-plot analysis of these eight studies did not definitively support either the presence or absence of publication bias (Supplemental Figure 1).

Meta-analysis of Prefrailty

Random effects analysis of the OR for POD in prefrail versus nonfrail patients was not significant (n = 349; OR = 2.30; 95% CI, 0.67–7.88; Supplemental Figure 2). When prefrail patients were combined with nonfrail patients and this pool was compared to frail patients, frail patients showed significantly greater odds of POD (n = 1008; OR = 2.05; 95% CI, 1.47–2.86) than the combined nonfrail-prefrail pool (Supplemental Figure 3). Although only three studies[22,27,28] delineated a prefrail category of patients for comparison, a fourth study[26] provided sufficient data for abstraction, such that four studies were available for analysis.

DISCUSSION

In a meta-analysis of adjusted data, older adult surgical patients with preoperative frailty had significantly greater risk of developing POD than their nonfrail counterparts. This finding suggests that assessment of preoperative frailty can be helpful for POD risk stratification. Although the frailty relationship was strong, the data regarding an association between prefrailty and POD were inconclusive.

Regardless of whether frailty assessment tools were based on phenotype or an index of accumulated deficits, the strong relationship between frailty and POD stands.

However, this conclusion must be tempered by the range of bias observed among studies as well as the small study size. Nevertheless, variability as a result of random effects was small. It remains unclear which components of frailty scales are the primary drivers of this relationship, or which frailty instrument is most predictive of POD. Should future research confirm the observation that frailty scales precisely stratify surgical and POD risk, frailty assessment could become a necessary part of preoperative care.

Despite the fact that the current meta-analysis, the two meta-analyses cited[5,6] and a recently published review[31] included studies with very little overlap, their results were comparable, with the relative risk of delirium to be just above 2. Despite the similarity of findings, our study is uniquely robust for several reasons. Studies using retrospective chart-based delirium determinations[5,6] or unspecified means of delirium assessment[5] were employed in previous meta-analysis. Our data derive solely from studies that assessed delirium at the bedside. Thus, we avoided bias from the high false-negative rate of chart review tools for delirium assessment.[16] Indeed, chart review tools are not recommended for delirium diagnosis.[16] Studies incorporating clinical judgement[6] or comprehensive geriatric assessment[5,31] were included in past and recent meta-analyses. Our aim was to make this meta-analysis as clinically relevant as possible for perioperative care. For this reason, we included only

studies whose frailty assessment tools were validated with specific cutoff points and could provide definitive categorization without the need for geriatric consultation.

Although this analysis suggests an effect of preoperative frailty on risk for POD, the exploratory analysis for prefrailty failed to show a similar effect. In the relatively small subset of prefrail patients, we found no significant difference in POD compared to nonfrail patients, although the OR estimate was similar in magnitude to that of the frail patients. Furthermore, when prefrail patients were grouped with nonfrail patients, this combined pool did not obliterate the overall effect of frailty on POD. As such, this exploratory analysis cannot definitively reconcile the relationship between prefrailty and POD. Several factors complicate the data. First, it is possible that prefrail patients do have increased odds of POD relative to nonfrail patients, but that our analysis was insufficiently powered to detect this difference. It's worth noting that analysis of this scale (number of studies and patients per study) only has sufficient power to detect a much greater effect size than observed, so the observed analysis result was subject to a high probability of type 2 error. Alternatively, some frail—or nonfrail—patients may have been misclassified as prefrail. In several of the frailty tools, prefrailty is distinguished from both the nonfrail and frail states by only a single point or domain of scoring. As such, any human or instrumental error in the scoring process could easily lead to misclassification, introducing greater variability in the exposure assessment. Nondifferential exposure measurement error or misclassification would bias the outcome association toward the null and mask the weaker association in studies without sufficient sample sizes. Of note, in studies that did not describe a prefrail group, similar misclassification could have occurred between the frail and nonfrail groups, underestimating the true degree of frailty-delirium association through these studies. As a consequence, the prefrailty category may have limited clinical utility for predicting postoperative delirium until more studies are available.

Strengths and Limitations

The strengths of this review include the restriction of frailty assessment tools to those that are validated and offer clear cutoff scores for frailty versus nonfrailty. Such frailty tools simplify assessment relative to continuous indices and could be easily implemented in the perioperative setting. This consideration is especially important given that the relationship between frailty and delirium appears robust. Our study also characterizes study bias through the use of ROBINS-I assessment—a necessity given that most studies of frailty and delirium appear to be nonrandomized studies of interventions—allowing for a more nuanced understanding of the study findings and similar publications in the field. Lastly, this study is strengthened by its exclusion of study populations—such as stroke or neurosurgery patients—whose comorbidities or surgical procedures independently associate with delirium. Exclusion of these patients allows better generalization of study findings to the surgical setting.

Limitations of this study include the lack of randomized-controlled trials; none were available based on the study question and eligibility criteria. Given that research personnel in each study were not blinded to the exposure status during outcome assessments, it is possible that assumptions about the effect of frailty or other confounding variables may have

introduced bias into the delirium screening process. Studies with different frailty instruments were included in the meta-analysis. Different frailty scales capture different but overlapping groups of patients resulting in different estimates for frailty prevalence.[32] Seven studies measured frailty using instruments based on the frailty phenotype while 2 used deficit accumulation approaches. The range of frailty across studies (18.6%–56%) is higher than previously reported systematic reviews examining community dwelling adults and may reflect the older surgical population.[2] POD ranged from 7%–56% among the studies used in this meta-analysis. All three delirium instruments utilized have been assessed for sensitivity/ specificity in comparison to the gold standard of psychiatric examination using DSM criteria. Specificity of these instruments for delirium is high; sensitivity varies but is still very high.[33] The key risk factors for postoperative delirium across studies including age, cognition, and co-morbidity were examined, however no discernable pattern explaining the variation in delirium incidence was observed. Variation in delirium incidence may be related to surgical procedure as the cardiac surgery studies in this meta-analysis as well as others[5] report consistently higher POD incidence. The scope of our study is limited by the high mean age; few younger patients were included because frailty and postoperative delirium are predominantly present in adults older than age 65. Although frailty is more common in women, only 42 percent of the subjects in these studies were women. In accordance with these limitations, our results may not be generalizable to all age groups, or female populations. The work elucidating the independent risk posed by surgery and anesthesia is evolving and knowing whether the surgeries/procedures themselves carry a high risk for postoperative complications, such as postoperative delirium, is important. As demonstrated in supplemental Table 1, our study did not include many patients undergoing low-cardiac-risk surgical procedures. This observation may be relevant in regard to post-operative delirium, as frailty may exert greater influence with higher risk procedures. Therefore, our results may not be applicable to all surgical procedures. Restricting to the search criteria to English can introduce bias. To determine risk of bias the search was rerun, both with and without the English limiter. Using google translate, the 3 non-English manuscripts would not have been included. So, for this topic, non-English language research was not a significant component of the results. There is also concern about bias due to unadjusted confounding. The E-value[20] for the observed OR of 2.14 between POD and frailty, the magnitude of association for unadjusted confounding with POD as well as with frailty that is needed to attenuate the OR from 2.14 to 1 above and beyond the measured confounding accounted for in the analysis, was 4.08. In a subgroup analysis that included the 3 studies which at least adjusted for age and baseline cognitive function, two of the strongest POD predictors established in the literature, the observed OR estimate for POD-frailty association increased slightly to 2.44. Putting these together, while residual confounding is likely, the observed association is not likely to be explained by unadjusted confounding.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Disclosure of funding:

Glossary of Terms

AHA	American Heart Association
ASA	American Society of Anesthesiologists
CAM	Confusion Assessment Method
DOS	Delirium Observation Scale
EFS	Edmonton Frail Scale
GFI	Groningen Frailty Indicator
ICDSC	Intensive Care Delirium Screening Checklist
LESCORE	Logistic European System for Cardiac Operative Risk Evaluation
MeSH	Medical Subject Headings
MMSE	Mini-Mental-State Examination
MoCA	Montreal Cognitive Assessment
POD	Postoperative delirium
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SOF	Study of Osteoporotic Fractures
TICS	Telephone Interview for Cognitive Status

References

1. Walston JD, Bandeen-Roche K. Frailty: a tale of two concepts. BMC Med 2015;13:185. [PubMed: 26265077]

2. Collard RM, Boter H, Schoevers RA, Oude Voshaar RC. Prevalence of frailty in community-dwelling older persons: a systematic review. J Am Geriatr Soc 2012;60:1487–1492. [PubMed: 22881367]

3. Cooper Z, Rogers SO Jr., Ngo L, et al. Comparison of frailty measures as predictors of outcomes after orthopedic surgery. J Am Geriatr Soc 2016;64:2464–2471. [PubMed: 27801939]

4. Makary MA, Segev DL, Pronovost PJ, et al. Frailty as a predictor of surgical outcomes in older patients. J Am Coll Surg 2010;210:901–908. [PubMed: 20510798]

5. Watt J, Tricco AC, Talbot-Hamon C, et al. Identifying older adults at risk of delirium following elective surgery: a systematic review and meta-analysis. J Gen Intern Med 2018;33:500–509. [PubMed: 29374358]

6. Persico I, Cesari M, Morandi A, et al. Frailty and delirium in older adults: a systematic review and meta-analysis of the literature. J Am Geriatr Soc 2018;66:2022–2030. [PubMed: 30238970]

7. Marcantonio ER, Goldman L, Mangione CM, et al. A clinical prediction rule for delirium after elective noncardiac surgery. JAMA 1994;271:134–139. [PubMed: 8264068]

8. Robinson TN, Raeburn CD, Tran ZV, Angles EM, Brenner LA, Moss M. Postoperative delirium in the elderly: risk factors and outcomes. Ann Surg 2009;249:173–178. [PubMed: 19106695]

9. Buurman BM, Hoogerduijn JG, de Haan RJ, et al. Geriatric conditions in acutely hospitalized older patients: prevalence and one-year survival and functional decline. PLoS One 2011;6:e26951. [PubMed: 22110598]

10. Nidadavolu LS, Ehrlich AL, Sieber FE, Oh ES. Preoperative evaluation of the frail patient. Anesth Analg 2020;130:1493–1503. [PubMed: 32384339]

11. Romero-Ortuno R An alternative method for Frailty Index cut-off points to define frailty categories. Eur Geriatr Med 2013;4:10.1016.

12. Fried LP, Tangen CM, Walston J, et al. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001;56:M146–156. [PubMed: 11253156]

13. Oh ES, Fong TG, Hshieh TT, Inouye SK. Delirium in older persons: advances in diagnosis and treatment. JAMA 2017;318:1161–1174. [PubMed: 28973626]

14. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009;62:e1–34. [PubMed: 19631507]

15. Rockwood K, Howlett SE. Fifteen years of progress in understanding frailty and health in aging. BMC Med 2018;16:220. [PubMed: 30477486]

16. 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;53:312–318. [PubMed: 15673358]

17. Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4:1. [PubMed: 25554246]

18. Ali MJ, Davison P, Pickett W, Ali NS. ACC/AHA guidelines as predictors of postoperative cardiac outcomes. Can J Anaesth 2000;47:10–19. [PubMed: 10626712]

19. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. Bmj 2016;355:i4919. [PubMed: 27733354]

20. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med. 2017;167(4):268–274. [PubMed: 28693043]

21. VanderWeele TJ. Optimal approximate conversions of odds ratios and hazard ratios to risk ratios. Biometrics 2020;76(3):746–752. [PubMed: 31808145]

22. Eide LS, Ranhoff AH, Fridlund B, et al. Comparison of frequency, risk factors, and time course of postoperative delirium in octogenarians after transcatheter aortic valve implantation versus surgical aortic valve replacement. Am J Cardiol 2015;115:802–809. [PubMed: 25644851]

23. Pol RA, van Leeuwen BL, Visser L, et al. Standardised frailty indicator as predictor for postoperative delirium after vascular surgery: a prospective cohort study. Eur J Vasc Endovasc Surg 2011;42:824–830. [PubMed: 21810543]

24. Partridge JS, Fuller M, Harari D, Taylor PR, Martin FC, Dhesi JK. Frailty and poor functional status are common in arterial vascular surgical patients and affect postoperative outcomes. Int J Surg 2015;18:57–63. [PubMed: 25907322]

25. Jung P, Pereira MA, Hiebert B, et al. The impact of frailty on postoperative delirium in cardiac surgery patients. J Thorac Cardiovasc Surg 2015;149:869–875.e861–862. [PubMed: 25486976]

26. Leung JM, Tsai TL, Sands LP. Brief report: preoperative frailty in older surgical patients is associated with early postoperative delirium. Anesth Analg 2011;112:1199–1201. [PubMed: 21372278]

27. Nomura Y, Nakano M, Bush B, et al. Observational study examining the association of baseline frailty and postcardiac surgery delirium and cognitive change. Anesth Analg 2019;129:507–514. [PubMed: 30540612]

28. Mahanna-Gabrielli E, Zhang K, Sieber FE, et al. Frailty is associated with postoperative delirium but not with postoperative cognitive decline in older noncardiac surgery patients. Anesth Analg 2020;130:1516–1523. [PubMed: 32384341]

29. Itagaki A, Sakurada K, Matsuhama M, Yajima J, Yamashita T, Kohzuki M. Impact of frailty and mild cognitive impairment on delirium after cardiac surgery in older patients. J Cardiol 2020;76:147–153. [PubMed: 32156513]

30. Khan SA, Chua HW, Hirubalan P, Karthekeyan RB, Kothandan H. Association between frailty, cerebral oxygenation and adverse post-operative outcomes in elderly patients undergoing non-cardiac surgery: An observational pilot study. Indian J Anaesth 2016;60:102–107. [PubMed: 27013748]

31. Tjeertes EK, van Fessem JM, Mattace-Raso FU, et al. Influence of Frailty on Outcome in Older Patients Undergoing Non-Cardiac Surgery - A Systematic Review and Meta-Analysis. Aging Dis 2020; 11:1276–1290. [PubMed: 33014537]

32. Theou O, Brothers TD, Mitnitski A, RockwoodK. Operationalization of frailty using eight commonly used scales and comparison of their ability to predict all-cause mortality. J Am Geriatr Soc 2013;61:1537–1551. [PubMed: 24028357]

33. van Velthuijsen EL, Zwakhalen SM, Warnier RM, et al. Psychometric properties and feasibility of instruments for the detection of delirium in older hospitalized patients: a systematic review. Int J Geriatr Psychiatry. 2016; 31:974–89. [PubMed: 26898375]

Key Points

Question:

To what extent is preoperative frailty associated with increased odds of developing postoperative delirium in elective surgical patients over age 64?

Findings:

Frail patients over 64 years of age undergoing elective surgical procedures have greater odds for developing postoperative delirium than do nonfrail patients of the same age.

Meaning:

This meta-analysis provides evidence for a significant association between preoperative frailty and POD in elective surgical patients aged 65 or older.

Clinical Implications

This meta-analysis provides evidence for an association between preoperative frailty and POD in elective surgical patients aged 65 or older. Use of simple, validated frailty scales in the preoperative setting can facilitate better estimation of odds for POD in this population. These findings suggest that brain vulnerability to health status changes is reflected in frailty status. As such, it is possible that interventions focused on delirium prevention, such as nonpharmacologic interventions or the ASA brain-health initiative recommendations, may become important cornerstones in perioperative management of the frail patient. Should future research confirm that frailty assessment can help to stratify POD risk, frailty assessment will be a necessary component of the older patient's preoperative evaluation.
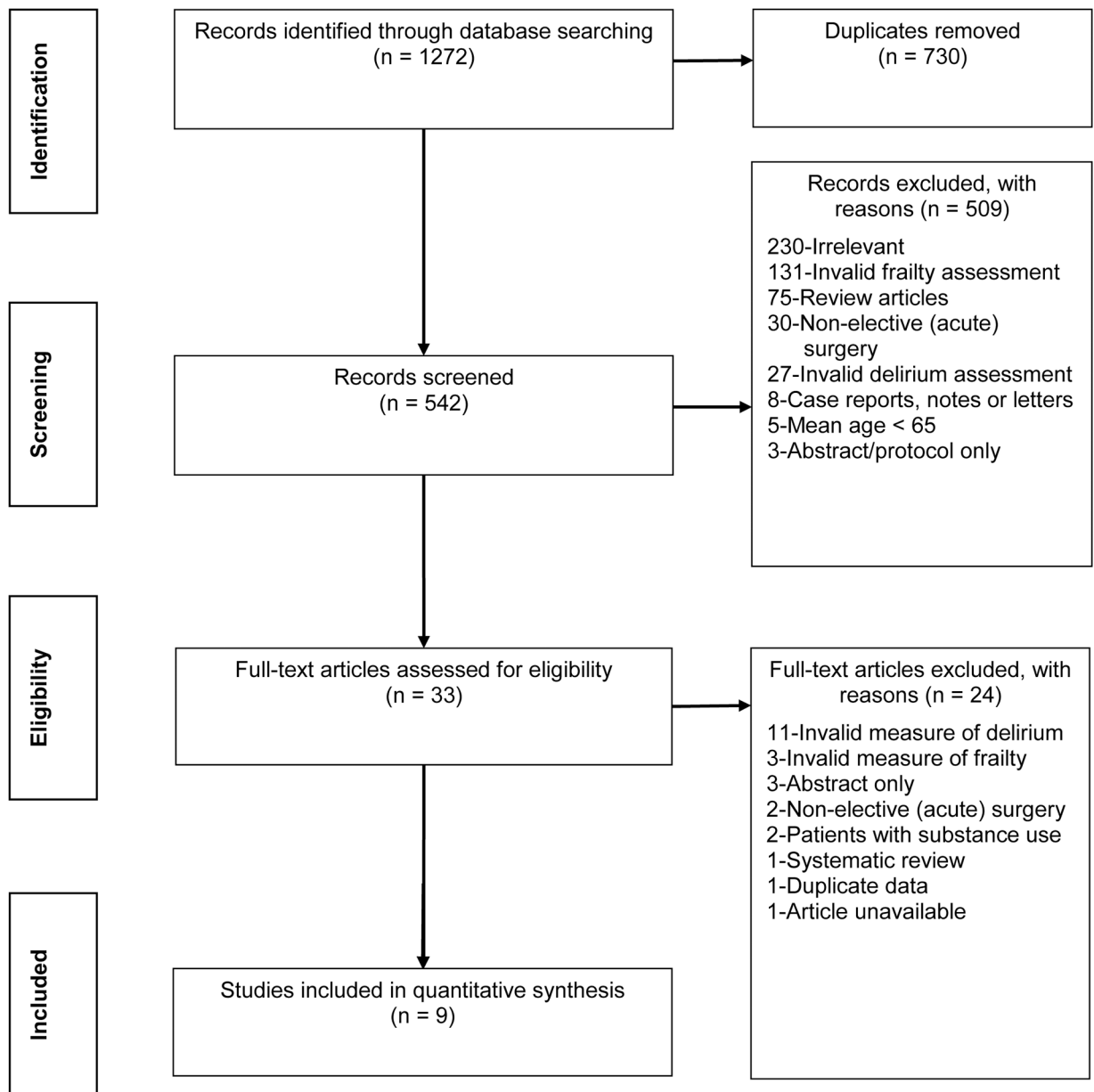
Figure 1.
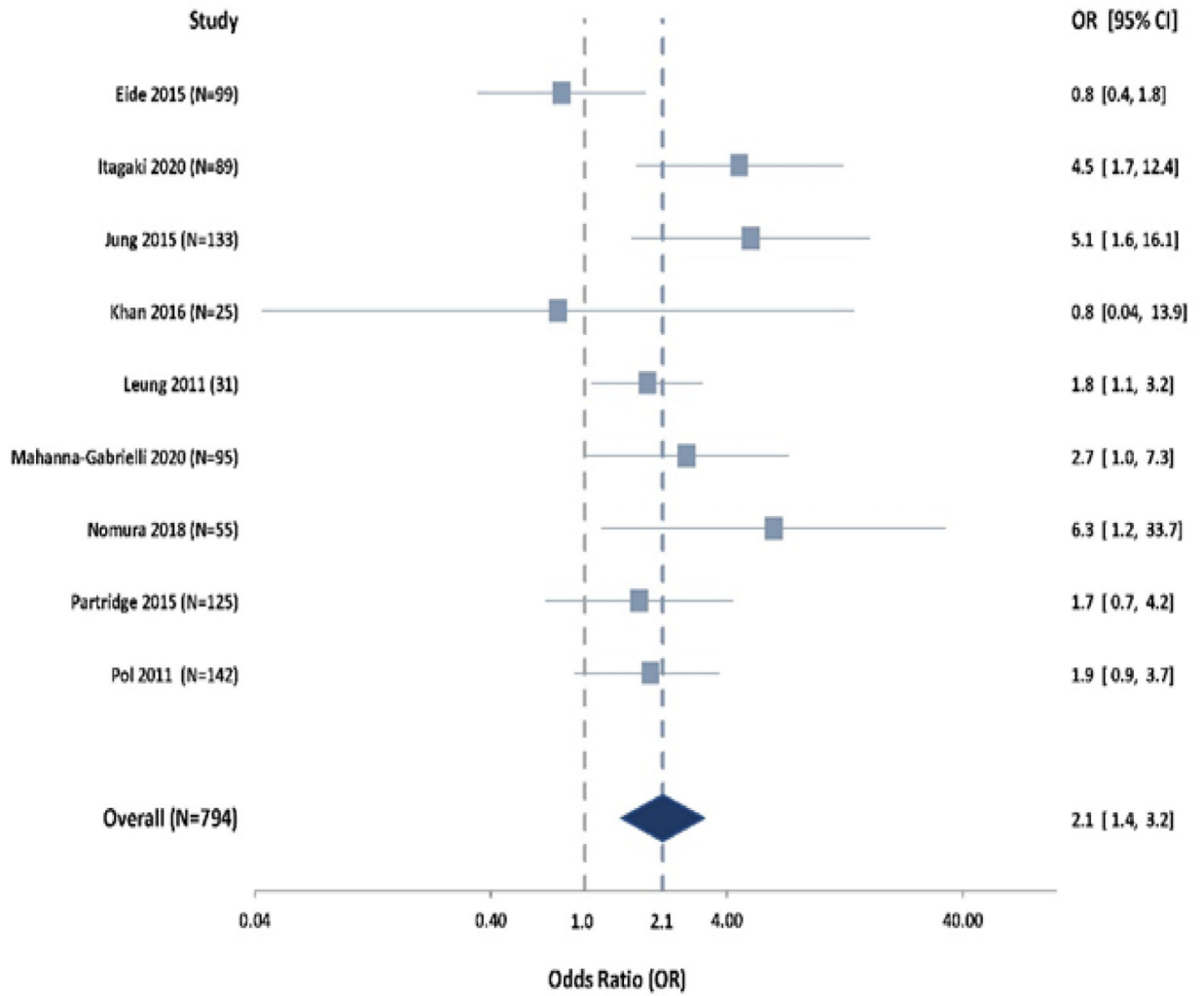Preferred reporting items for systematic reviews and meta-analyses flowchart.

Figure 2.
Forest plot of random-effect analysis comparing postoperative delirium in frail and nonfrail patients. Prefrail patients were not included in this analysis.

Table 1.

Characteristics of Studies Selected for Qualitative Review and Meta-Analysis

Study (Year)	Patients, N	Frailty Metric	Score Indicating Frailty	Percent Frail	Percent Prefrail	Postoperative Delirium Metric	Percent Delirium	Adjusted Odds Ratio (95% CI)	Variables Adjusted For
Eide et al[22] (2015)	136	SOF	0 = nonfrail, 1 = prefrail, 2–3 = frail	39	27.3	CAM	55.9	0.8 (0.4–1.8) P = 0.059	Sex, ADL, A-fib, CCI, MMSE, Post-op opioids, post-op anxiolytics, TAVI
Itagaki et al[29] (2020)	89	J-CHS (Fried)	0 = nonfrail, 1–2 = prefrail, 3–5 = frail	38.2	-	ICDSC	28.1	4.5 (1.7–12.4) P = 0.003	MoCA, sex, age, albumin, operation
Jung et al[25] (2015)	133	Modified Fried	>2 = frail	54.1	-	CAM, CAM-ICU	18.0	5.1 (1.6–16.1) P = .0015	EUROSCORE II
Khan et al[30] (2016)	25	Modified Fried	>2 = frail	56.0	-	CAM-ICU	8.0	0.77 (0.04–13.87)[a]	Unadjusted*
Leung et al[26] (2011)	63	Fried	0 = nonfrail, 1–2 = prefrail, 3–5 = frail	33.3	50.8	CAM	25.4	1.8 (1.1–3.2)[b] P = 0.028	Age, ADL, IADL, TICS, GDS
Mahanna-Gabrielli et al[28] (2020)	167	FRAIL scale	0 = nonfrail, 1–2 = prefrail, 3–5 = frail	18.6	43.1	CAM-ICU	22.8	2.7 97.5% (1.0–7.3)[c]	Age, sex, education, surgical duration, surgical type, ASA status, baseline cognition
Nomura et al[27] (2019)	128	Fried	0 = nonfrail, 1–2 = prefrail, 3–5 = frail	31.3	57	CAM-ICU	43.8	6.3 (1.2–33.7) P = 0.31	Age, sex, education, LESCORE
Partridge et al[24] (2015)	125	EFS	>6.5 = frail	52	-	CAM	19.2	1.7 (0.7–4.2)[d]	Univariate analysis only
Pol et al[23] (2011)	142	GFI	>3 = frail	35.2	-	DOS	7.0	1.9 (0.9–3.7) P = 0.05[e]	CRP, CCI, GFI, impaired renal function

Abbreviations: ADL, activities of daily living; A-fib, atrial fibrillation; ASA, American Society of Anesthesiologists; CAM, Confusion Assessment Method; CAM-ICU, Confusion Assessment Method for the intensive care unit; CCI, Charlson comorbidity index; CRP, C-reactive protein; DOS, Delirium Observation Scale; EFS, Edmonton Frail Scale; GDS, geriatric depression scale; GFI, Groningen Frailty Indicator; IADL, instrumental activities of daily living; ICDSC, Intensive Care Delirium Screening Checklist; J-CHS, Japanese version of the cardiovascular health study; MMSE, Mini-Mental-State Examination; LESCORE, Logistic European System for Cardiac Operative Risk Evaluation; SOF, Study of Osteoporotic Fractures; TAVI, Transcatheter Aortic Valve Implantation; TICS, Telephone Interview for Cognitive Status.

[a] 9.1% nonfrail, 6.2% frail, no p value reported. Based on our own calculation form data abstraction, unadjusted values.

[b] Values given are for preoperative frailty score, not category.

[c] No adjusted P value.

[d] Based on our own calculation, abstracted from raw data. Unadjusted values.

[e]Odds ratio is for GFI score, not frailty

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

ROBINS-I Assessment of Study Bias for Included Studies

Bias Domain	Eide et al[22]	Itagaki et al[29]	Jung et al[25]	Khan et al[30]	Leung et al[26]	Mahanna et al[28]	Nomura et al[27]	Partridge et al[24]	Pol et al[23]
Due to confounding	Moderate	Low	Critical	Moderate	Low	Low	Moderate	Serious	Serious
Selection of participants	Low	Moderate	Moderate	No Info	Moderate	Low	Serious	Low	Low
Classification of interventions	Low	Low	Low	Low	Low	Low	Low	Low	Low
Deviation from intended interventions	Low	Low	Low	Low	Low	Low	Low	Low	Low
Missing data	Moderate	Low	Low	Low	Low	Low	Low	Moderate	Serious
Measurement of outcomes	Low	Low	Low	Low	Low	Low	Moderate	Low	Serious
Selection of the reported result	Moderate	Low	Low	Low	Low	Low	Low	Low	Low
Overall risk of bias	**Moderate**	**Moderate**	**Critical**	**Moderate**	**Moderate**	**Low**	**Serious**	**Serious**	**Serious**

Table 3.

Comparability of Confounding Variables Between Frail and Nonfrail Patients by Study

Study	Age	Percent Female	Baseline Cognition		Confounding Variable Comorbidity		Type of Surgery	Overall AHA Risk Level
	Group Comparison	Group Comparison	Group Comparison	Metric	Group Comparison	Metric	Group Comparison	
Jung et al[25]	$P = 0.0023$	$P = 0.0459$	$P < 0.001$	MoCA	$P = 0.0001$	EUROSCORE II	Same Procedure	High
Khan et al[30]	$P = 0.89$	$P = 0.08$	$P = 0.50$	MMSE	$P = 0.07$	ASA	$P = 0.18$	Intermediate
Mahanna-Gabrielli et al[28][a]	$P = 0.74$	$P = 0.05$	$P = 0.29$	MMSE	$P = 0.02$	ASA	$P = 0.053$	Intermediate
Nomura et al[27][a]	$P = 0.156$	$P = 0.230$	$P = 0.553$	Cognitive Z-score	$P = 0.003$	LESCORE*	$P = 0.028$	High
Partridge et al[24][b]	$P = 0.445$	$P = 0.068$	P 0.001	MoCA < 24	No summary score given	No summary score given	Not Reported	High

The following studies did not report confounding variables stratified by delirium and were not included in this table: Eide et al.[22] and Pol et al.[23]

Itagaki et al.[29] was not included because the report lacked summary confounding variables for frailty, non-stratified by minimal cognitive impairment.

Leung et al.[26] was not included because authors reported confounding variables based on frailty score, not frailty category.

Abbreviations: AHA, American Heart Association; ASA, American Society of Anesthesiologists; LESCORE, Logistic European System for Cardiac Operative Risk Evaluation; MMSE, Mini-Mental-State Examination; MoCA, Montreal Cognitive Assessment; TICS: Telephone Interview for Cognitive Status.

[a] Statistics are group comparisons of frail versus prefrail versus nonfrail

[b] P value is for percent of patients over age 75.