RESEARCH ARTICLE

BMC Anesthesiology

Open Access

Effects of an abnormal mini-mental state examination score on postoperative outcomes in geriatric surgical patients: a meta-analysis



Shuang-jiao Cao[†], Dongxu- Chen[†], Lei Yang and Tao Zhu^{*}

Abstract

Background: Perioperative cognitive impairment (CI) following surgeries is prevalent in geriatric surgical population aged 60 and older. This meta-analysis was designed to investigate whether the Mini-Mental State Examination (MMSE) has prognostic value on adverse outcomes in aged surgical patients.

Methods: PubMed, Cochrane, Embase and Medline through the Ovid were searched. Meta-analyses were carried out for CI versus non-cognitive impairment (NCI). Quality of evidence was assessed by the GRADE approach.

Results: One randomized controlled trial, two retrospective cohort trials, and 18 prospective cohort trials were included in the meta-analysis. Perioperative diagnosis of CI by the MMSE had higher rates of patients suffering from postoperative delirium (POD) [odd ratio (OR) 5.02, 95% confidence interval (CI) 3.27, 7.71, P < 0.00001], in-hospital mortality (OR 7.51, 95% CI 2.17, 26.02, P = 0.001), mortality within 1 year (OR 2.53, 95% CI 1.95,3.29, P < 0.00001). Postoperative CI patients had no extended length of stay in orthopedic [standardized mean difference (SMD) -0.10, 95% CI -0.20, 0.17, P = 0.91)] nor rehabilitation wards ((SMD, 0.04; 95% CI, -0.23 to 0.31; P = 0.78).

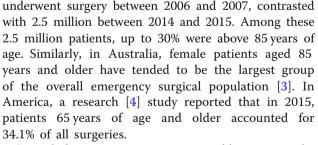
Conclusion: Older patients with perioperative CI were more likely to suffer from POD and mortality. The MMSE showed certain value on risk stratification and prognosis evaluation in geriatric surgical population.

Trial registration: PROSPERO CRD42018108739.

Keywords: MMSE, Geriatric, Outcomes, Postoperative delirium, Mortality, Meta-analysis

Backgrounds

Surgical safety and success rates have improved considerably as a result of remarkable medical breakthroughs, which has led to the extension of life expectancy and the rise in the number of aged patients undergoing surgical procedures [1]. Compared to 2017, the number of persons aged 60 and older is expected to more than double by 2050 and more than triple by 2100, thereby rising from 962 million globally in 2017 to 2.1 billion in 2050 and 3.1 billion in 2100 [2], which will necessitate an increasing demand for operations in aged individuals. In England, less than 1.5 million people over 75 years old



Comorbidities are more common in older patients, the ability for their bodies to compensate decreases, and some health problems may be underdiagnosed; these factors contribute to an inability to tolerate surgery [5]. Surgery in older patients presents medical workers with formidable challenges as they must weigh long-term benefits and risks carefully to make the best surgical

© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

^{*} Correspondence: xwtao_zhu@yahoo.com

[†]Shuang-jiao Cao and Dongxu- Chen contributed equally to this work. Department of Anesthesiology, West China Hospital, Sichuan University, Chengdu 610041, People's Republic of China

decisions. Previous studies have reported that both preexisting and new-onset cognitive impairment (CI) following surgeries, which has been observed in 19-83% of elderly patients varying with age and type of surgery [6, 7], was associated with an increased incidence of postoperative complications such as delirium. Other long-term problems and poor outcomes have also been of concern, such as mortality, impaired functional status, readmission, prolonged hospitalizations and increased expenses [3, 8–10]. The Mini-Mental State Examination (MMSE) is long established and the most widely used instrument to help clinicians detect cognitive impairment and grade the severity of cognitive change [11]. This tool is frequently applied in the perioperative period of older patients due to its adequate performance in a rule-out capacity [12]. However, several studies have confirmed that the MMSE score could be a predictor of adverse postoperative outcomes, but those findings are limited by the sample sizes or other confounders. Thus, whether the MMSE could be an indicator of the prognosis of geriatric surgical patients remains unclear. To our knowledge, there have been no reviews systematically exploring and quantifying the association between impaired cognition diagnosed by the MMSE and different postoperative outcomes.

We conducted this systematic review to investigate whether the MMSE used as a perioperative assessment tool has prognostic value on adverse clinical outcomes in aged surgical patients and to search for a better understanding and guidance for clinicians to evaluate and make optimal patient care decisions.

Methods

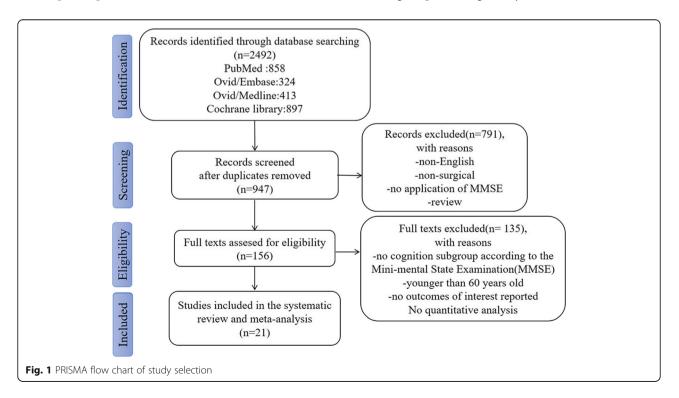
Data sources and search strategy

This systematic review was performed and presented following the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [13] and Assessing the Methodological Quality of Systematic Reviews (AMSTAR) guidelines. The protocol was registered with the International Prospective Register of Systematic Reviews (http://www.crd.york. ac.uk/PROSPERO/display_record.php?ID=CRD4201810 8739 RecordID = 108,739). A systematic literature search was conducted for studies published from 1995 to April 2018 by searching PubMed, the Cochrane Library, Embase and Medline through Ovid. The search terms combined medical subject headings and keywords related to the MMSE, surgery and outcomes. The sensitive search was performed by using the following terms: "cognition", "cognitive", "delirium", "complication", "outcome", "length of stay", "surgery", "operation", "operative", "procedure", "Mini-Mental State Examination", "MMSE". The literature search strategy is provided in Additional file 1: Material 1.

Study selection and data collection

This systematic review and meta-analysis included randomized controlled trials (RCTs), prospective cohort studies, and retrospective cohort studies. Eligible studies were included if they met the following criteria:

(1) Surgical patients aged 60 years and older.



Study	Design	Surgery type	Sample size	Participant	initiation of MMSE	Cutoff point of MMSE	Reported outcomes of interest
Beloosesky 2002 [26]	Prospective cohort study	Hip surgery	153	≥65 yr	Before surgery	24	In-hospital mortality
Bliemel 2015 [27]	Prospective cohort study	Hip surgery	399	≥60 yr	Before surgery	27	1-year mortality
Brouquet 2010 [<mark>25</mark>]	Prospective cohort study	Abdominal surgery	118	≥75 yr	Before surgery	26	POD
Guo 2014 [<mark>29</mark>]	Prospective cohort study	Hip surgery	244	>60 yr	Before surgery	24	Total hospitalization days (surgery and rehabilitation)
							1-month/6-months/1- year mortality
Häkkinen 2007 [<mark>28</mark>]	Prospective cohort study	Hip surgery	117	≥65 yr	After surgery	24	Length of stay in orthopedic /rehabilitation ward
							1-year mortality
Huusko 2000	RCT	Hip surgery	243	≥65 yr	After	24	3 month/1 year mortality
[19]					surgery		Length of stay in rehabilitation ward
Jones 2017	Prospective	Hip surgery	383	≥65 yr	After	18	6 months Mortality
[31]	cohort study				surgery		Length of stay in orthopedic wards
Kalisvaart 2006 [<mark>30]</mark>	Prospective cohort study	Hip surgery	603	≥70 yr	Before surgery	24	POD
Karni 2013 [<mark>33</mark>]	Prospective cohort study	Hip surgery	60	≥65 yr. Female	After surgery	24	Length of stay in rehabilitation ward
Kratz 2015 [32]	Prospective cohort study	general, abdominal, and trauma surgery	178	>70 yr	Before surgery	27	POD
Lee 2016 [7]		lumbar spine surgery	129	>65 yr	Before	24	POD
	cohort study				surgery		Length of stay in hospital
Moncada 2005 [<mark>35</mark>]	Prospective cohort study	Hip surgery	48	≥65 yr	After surgery	24	Length of stay in orthopedic /rehabilitation ward
							POD
Morghen 2011 [<mark>34</mark>]	Prospective cohort study	Hip surgery	386	≥65 yr	After surgery	24	Length of stay in rehabilitation ward
Osse 2012 [<mark>36</mark>]	Prospective cohort study	Cardiac surgery	125	≥70 yr	Before surgery	28	POD
Otano 2015 [<mark>38]</mark>	Prospective cohort study	Hip surgery	285	≥65 yr	After surgery	24	In-hospital mortality Length of stay in rehabilitation ward
Reissmüller 2006 [<mark>37</mark>]	Prospective cohort study	Cardiac surgery	107	≥60 yr	Before surgery	24	POD
Rolland 2004 [22]	Prospective cohort study	Hip surgery	61	≥70 yr	After surgery	20	Length of stay in rehabilitation ward
Ruggiero 2016 [21]	Prospective cohort study	Hip surgery	514	≥65 yr	After surgery	24	1-year mortality
Schaller2012 [23]	Prospective cohort study	Hip surgery	173	≥65 yr	After surgery	24	1 year mortality
Witlox 2009 [<mark>24</mark>]	Prospective cohort study	Hip surgery	76	≥75 yr	Before surgery	24	POD
Yukako 2016 [<mark>20</mark>]	Retrospective cohort study	Colorectal surgery	156	≥75 yr	Before surgery	24	POD

 Table 1
 Study characteristics.
 MMSE, the Mini-Mental State Examination; POD, postoperative delirium; RCT, randomized controlled trials

Table 2 Risk of bias for cohort trail	ls
---------------------------------------	----

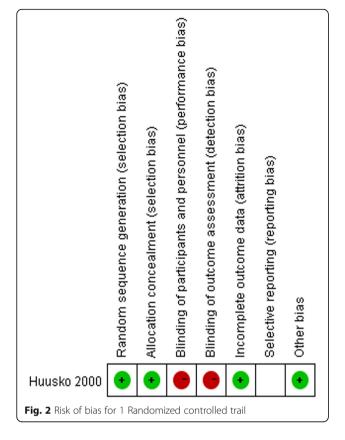
Study	Selection				Comparability	outcome			Score
	Represent- ativeness of the exposed cohort	Selection of the non- exposed cohort	Ascertainment of exposure	Demonstration that outcome of interest was not present at start of study	Comparabili-ty of Cohorts on the Basis of the Design or Analysis	Assess-ment of outco-me	Was follow-up long enough for out comes to occur	Adequacy of follow -up of cohorts	
Beloosesky 2002 [26]	*	*	*	*		*		*	6
Bliemel 2015 [27]	*	*	*	*	*	*	*	*	8
Brouquet 2010 [25]	*	*	*	*		*	*	*	7
Guo 2014 [<mark>29</mark>]	*	*	*	*	*	*	*	*	8
Häkkinen 2007 [28]	*	*	*	*	*	*	*	*	8
Huusko 2000 [19]	*	*	*	*	*	*	*	*	8
Jones 2017 [31]	*	*	*	*		*	*	*	7
Kalisvaart 2006 [<mark>30</mark>]		*	*	*	*	*		*	6
Karni 2013 [<mark>33</mark>]	*	*	*	*		*	*	*	7
Kratz 2015 [<mark>32</mark>]	*	*	*		*		*	*	6
Lee 2016 [7]	*	*	*	*		*	*	*	7
Moncada 2005 [35]	*	*	*	*	*	*		*	7
Morghen 2011 [34]	*	*	*	*	*	*	*	*	8
Osse 2012 [<mark>36</mark>]	*	*	*	*		*	*	*	7
Otano 2015 [<mark>38</mark>]	*	*	*	*	*	*	*	*	8
Reissmüller 2006 [37]	*	*	*	*		*	*	*	7
Rolland 2004 [22]	*	*	*	*	*	*	*	*	8
Ruggiero 2016 [21]	*	*	*	*		*	*	*	7
Schaller2012 [23]	*	*	*	*	*	*	*	*	8
Witlox 2009 [24]	*	*	*	*		*	*	*	7
Yukako 2016 [<mark>20</mark>]	*	*	*			*	*	*	6

- (2) The MMSE was applied during the preoperative or postoperative period or at admission to the rehabilitation ward/hospital/facility.
- (3) The outcomes of interest included all postoperative complications, especially postoperative delirium (POD), hospitalization days, mortality (in-hospital mortality and long-term mortality).
- (4) Quantitative data were reported to compare each MMSE group with outcomes.

The initial step was based on screening titles and abstracts to exclude irrelevant studies. Second, the full contents of potentially eligible studies were read. Additionally, data were extracted and collected. The extracted characteristics of the studies included author, published year, study design, type of surgery, sample size, patient age, initial time of MMSE assessment (before or after surgery), the cutoff point of the MMSE scores to define cognitive impairment, and reported outcomes along with their definitions and follow-up duration. Outcomes that had been observed in the same way in more than two studies were included in the meta-analysis. The number of events and the number of participants in each group were extracted for dichotomous outcomes. Mean, standard deviation and the number of participants were extracted for continuous outcomes. The screening and extraction were conducted by two authors separately. Discrepancies with regard to eligibility were determined by a third author.

Statistical analysis

We used Review Manager (RevMan, version 5.3 for Windows, Oxford, UK; The Cochrane Collaboration, 2008) to perform the meta-analysis, which included generating forest plots and testing for heterogeneity and overall effects. Dichotomous outcomes were analyzed by the Mantel-Haenszel method and odds ratios (ORs), while continuous outcomes were analyzed by the inverse variance method and standard mean difference (SMD). Random effects models were used for all analyses. Heterogeneity was assessed using the I² statistic. For this measure, 0 to 50%, 50 to 75%, and 75 to 100% represented low, moderate, and high levels of heterogeneity, respectively [14]. Moderate to high levels of heterogeneity ($I^2 > 50\%$) between studies were investigated by several subgroup analyses including preoperative and postoperative MMSE, length of stay in orthopedic wards and rehabilitation wards, in-hospital mortality and



1-year mortality. Sensitivity analyses were conducted to explore the impact of imputing nonsignificant results on pooled effects. *P* values of less than 0.05 indicated statistical significance. Funnel plots were generated in Review Manager, and Egger's test was performed in STATA15.0 (StataCorp LLC, Texas) to assess publication bias [15].

Quality assessment and risk of bias

Two authors assessed the risk of bias independently. Disagreement was resolved by consulting other authors. We used the Newcastle Ottawa Quality Assessment Scale [16] (NOS) for Cohort Studies (range of 0 to 9 stars), a tool for the critical appraisal of eligible cohort studies. We regarded a study that scored seven or more stars as high quality and five or less stars as poor quality. We used the Cochrane Risk of Bias tool, which allows an assessment of low, moderate, or high risk of bias [17] to analyze the quality of RCTs.

We used the GRADE (Grading of Recommendations, Assessment, Development, and Evaluation) [18] approach to rate the quality of the evidence for postoperative delirium, length of stay, readmission to hospital and admission to nursing home within 1 year. Evidence was judged as high, moderate, low and very low in consideration of risk of bias, inconsistency, indirectness, imprecision and other considerations. We used the GRADEpro GDT to generate the evidence profile.

Results

We identified a total of 2492 records. After removal of duplicates, we screened 947 titles and abstracts, of which 156 full text articles were selected for eligibility. Twenty-one studies met the inclusion criteria for the systematic review (Fig. 1).

The designs of the included studies were a randomized controlled trial [19](N = 1), retrospective cohort trials [7, 20](N = 2), and prospective cohort trials [21-38](N = 18). Two sets of data were extracted from the intervention group and control group of the RCT (Table 1). When defining cognitive impairment or cognitive intact, most of the included studies used the MMSE cutoff score of 24, while another 6 studies used a cutoff score of 18 [31], 20 [22], 26 [25], 27 [27, 32], or 28 [36]. Considering the inconsistency of the cutoff point, we grouped our study populations into CI and NCI based on a clear definition by every study instead of selecting a specific cutoff point.

Generally, most of the included studies were judged to be of moderate to high quality. The risk of bias concerns in all the cohort studies were frequently about the comparability of the CI and NCI groups. Other risks of bias included short follow-up durations and the presence of outcomes at the start of the studies. The RCT was at high risk of performance and detection bias and unclear risk of reporting bias (Table 2, Fig. 2).

Postoperative delirium (POD)

Ten studies [7, 20, 24, 25, 30, 32, 35–38] reported POD including 1411 patients with NCI and 543 with CI (Fig. 3). Patients with perioperative CI had a higher rate of POD compared with NCI patients [odds ratio (OR), 5.02; 95% confidence interval (CI), 3.27 to 7.71; P < 0.00001] (Fig. 3).

In a subgroup analysis, three studies [35, 37, 38] used the MMSE postoperatively, while eight [7, 20, 24, 25, 30, 32, 36, 37] studies used the MMSE preoperatively. One study [37] applied the MMSE both preoperatively and postoperatively. The rate of POD was higher in the preoperatively diagnosed CI group than in the preoperatively diagnosed NCI group (OR, 5.12; 95% CI, 3.46 to 7.59; P < 0.00001; $I^2 = 30\%$) (Fig. 3a). Postoperatively diagnosed CI did not increase the rate of POD (OR, 5.55; 95% CI, 0.91 to 33.88; P = 0.06; $I^2 =$ 84%), and substantial heterogeneity existed ($I^2 = 84\%$) (Fig. 3b), then we conducted a sensitivity analysis to explore the stability of the latter results. In the subgroup of postoperative MMSE for the outcome of postoperative delirium, we found that the postoperative CI group showed a higher rate (OR, 0.07; 95% CI, 0.01 to 0.45; P = 0.006; $I^2 = 66\%$) after excluding the

Study or Subaroup	CI Events	Total	NCI Events		Weight	Odds Ratio M-H, Random, 95% Cl			s Ratio dom, 95% Cl	
a preoperative MM										
Brouquet 2010	13	36	15	82	10.5%	2.52 [1.05, 6.09]			—	
Kalisvaart 2006	48	151	26	452		7.64 [4.52, 12.89]			_ _	
Kratz 2015	25	60	11	118		• • •			_	
Lee 2016	12	49	6	78		3.89 [1.35, 11.20]				
Osse 2012	.2	13	49	112		• • •				
Reissmüller 2006	4	4	21	103		• • •				
Witlox 2011	14	25	15	49	9.4%					
Yukako 2016	21	42	13	114						
Subtotal (95% CI)		380		1108		5.12 [3.46, 7.59]			•	
Total events	146		156			• • • •				
Heterogeneity: Tau ² = 0 Test for overall effect: 2	-		-	(P = 0	.19); I² = 3	30%				
b postoperative M	MSE									
Moncada 2005	4	18	8	30	6.4%	0.79 (0.20, 3.11)			+	
Otano 2015	89	135	39	179	15.1%	6.95 [4.20, 11.48]				
Reissmüller 2006	9	10	15	94		47.40 [5.59, 402.25]				→
Subtotal (95% CI)		163		303	24.8%	5.55 [0.91, 33.88]				-
Total events	102		62							
Heterogeneity: Tau ² = : Test for overall effect: 2	-		-	(P = 0	.002); l² =	84%				
Total (95% CI)		543		1411	100.0%	5.02 [3.27, 7.71]			•	
Total events	248		218							
									1 1	
Heterogeneity: Tau ² = 0	0.25; Chi	i ^z = 22.4	4, df = 1	0 (P = 1	0.01); I² =	55%	1 01	01	1 10	100
Heterogeneity: Tau ² = 0 Test for overall effect: 2	-		-	0 (P = 1	0.01); l² =	55%	0.01	0.1		100
	Z = 7.37 ((P < 0.0	0001)				0.01		1 10 Favours (NCI)	100
Test for overall effect: 2	Z = 7.37 ((P < 0.0	0001)			: 0%	0.01	Favours [CI]	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe	Z = 7.37 (rences:	(P < 0.01 Chi² = 0	0001) .01, df=	1 (P =	0.93), I² =	: 0% Odds Ratio	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe Study or Subgroup	Z = 7.37 (erences: log[O	(P < 0.0	0001) .01, df=	1 (P =	0.93), I² =	: 0%	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe <u>Study or Subgroup</u> C adjusted for ag	Z = 7.37 (erences: log[O	(P < 0.0) Chi ² = 0 dds Rat	0001) .01, df= <u>io]</u>	1 (P =	0.93), I² = Weight	: 0% Odds Ratio IV, Random, 95% Cl	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe Study or Subgroup C adjusted for ag Kratz 2015	Z = 7.37 (erences: log[O	(P < 0.00 Chi ² = 0 dds Rat 1.43	0001) .01, df= io] :03 0.4	1 (P = <u>SE</u> 551	0.93), I ² = <u>Weight</u> 12.1%	: 0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe Study or Subgroup C adjusted for ag Kratz 2015 Osse 2012	Z = 7.37 (erences: log[O	(P < 0.00 Chi ² = 0 dds Rat 1.43 1.45	0001) .01, df= io] :03 0.4 :86 0.4	1 (P = <u>SE 1</u> 551 646	0.93), I ² = <u>Weight</u> 12.1% 11.7%	: 0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe Study or Subgroup C adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006	Z = 7.37 (erences: log[O	(P < 0.00 Chi ² = 0 dds Rat 1.43 1.45 2.42	0001) .01, df= .03 0.4 .03 0.4 .03 0.4 .03 0.4	1 (P = <u>SE 1</u> 551 646 348	0.93), I ² = <u>Weight</u> 12.1% 11.7% 5.2%	: 0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe Study or Subgroup C adjusted for ag Kratz 2015 Osse 2012	Z = 7.37 (erences: log[O	(P < 0.00 Chi ² = 0 dds Rat 1.43 1.45 2.42	0001) .01, df= io] :03 0.4 :86 0.4	1 (P = <u>SE 1</u> 551 646 348	0.93), I ² = <u>Weight</u> 12.1% 11.7%	: 0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe Study or Subgroup C adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006	Z = 7.37 (erences: log[O	(P < 0.00 Chi ² = 0 dds Rat 1.43 1.45 2.42	0001) .01, df= .03 0.4 .03 0.4 .03 0.4 .03 0.4	1 (P = <u>SE 1</u> 551 646 348	0.93), I ² = <u>Weight</u> 12.1% 11.7% 5.2%	: 0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe Study or Subgroup C adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016	Z = 7.37 (rences: <u>log[O</u> je = 0.00; C	(P < 0.0(Chi ² = 0 d <u>ds Rat</u> 1.43 1.45 2.42 1.99 Chi ² = 2.	0001) .01, df = 03 0.4 86 0.4 48 0.7 161 0.4	1 (P = <u>SE</u> 551 646 348 483 3 (P =	0.93), I ² = <u>Weight</u> 12.1% 11.7% 5.2% 12.5% 41.5 %	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe <u>Study or Subgroup</u> <u>c</u> adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect	Z = 7.37 (rences: <u>log[O</u> je = 0.00; C	(P < 0.0(Chi ² = 0 d <u>ds Rat</u> 1.43 1.45 2.42 1.99 Chi ² = 2.	0001) .01, df = 03 0.4 86 0.4 48 0.7 161 0.4	1 (P = <u>SE</u> 551 646 348 483 3 (P =	0.93), I ² = <u>Weight</u> 12.1% 11.7% 5.2% 12.5% 41.5 %	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted	Z = 7.37 (rences: <u>log[O</u> je = 0.00; C	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0	0001) .01, df = .03 0.4 .86 0.4 .48 0.7 .61 0.4 02, df = .00001)	1 (P = <u>SE</u>) 551 646 348 483 3 (P =	0.93), ² = <u>Weight</u> 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); ² :	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0%	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010	Z = 7.37 (rences: <u>log[O</u> je = 0.00; C	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92	0001) .01, df= i0] 03 0.4 86 0.4 48 0.7 61 0.4 02, df= .00001)	1 (P = <u>SE 1</u> 551 646 348 483 3 (P =	0.93), ²= <u>Weight</u> 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); ²: 12.4%	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0% 2.52 [1.05, 6.09]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006	Z = 7.37 (rences: <u>log[O</u> je = 0.00; C	(P < 0.00 Chi ⁼ = 0 (1.43 1.45 2.42 1.99 Chi ⁼ = 2. 6 (P < 0 0.92 2.03	0001) .01, df= i0] 03 0.4 86 0.4 48 0.7 61 0.4 02, df= .00001) 61 0.4 28 0.2	1 (P = <u>SE 1</u> 1646 348 483 3 (P = 494 2671	0.93), *= Weight 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); *: 12.4% 27.0%	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0% 2.52 [1.05, 6.09] 7.64 [4.52, 12.89]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016	Z = 7.37 (rences: <u>log[O</u> je = 0.00; C	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35	0001) .01, df = .03 0.4 .86 0.4 .48 0.7 .61 0.4 02, df = .00001) .61 0.4 .28 0.2 .89 0.5	1 (P = <u>SE 1</u> 551 646 348 483 3 (P = 494 671 394	0.93), *= Weight 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); *: 12.4% 27.0% 9.1%	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0% 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009	Z = 7.37 (rences: <u>log[O</u> je = 0.00; C	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35	0001) .01, df= i0] 03 0.4 86 0.4 48 0.7 61 0.4 02, df= .00001) 61 0.4 28 0.2	1 (P = <u>SE 1</u> 551 646 348 483 3 (P = 494 671 394	0.93), ² = <u>Weight</u> 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); ² : 12.4% 27.0% 9.1% 10.1%	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0% 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009 Subtotal (95% CI)	Z = 7,37 (rences: log[O e = 0.00; C : Z = 6.9	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35 1.05	0001) .01, df = 103 0.4 103 0.4 186	1 (P = <u>SE</u>) 551 646 348 483 3 (P = 494 671 394 6083	0.93), *= Weight 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); *: 12.4% 27.0% 9.1% 10.1% 58.5%	2.52 [1.05, 6.09] 7.64 [4.52, 12.29] 2.89 [1.35, 11.20] 3.89 [1.35, 11.20] 5.61 [3.45, 9.11] 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81] 4.20 [2.30, 7.64]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009 Subtotal (95% CI) Heterogeneity: Tau ² =	Z = 7,37 (rences: log[O e = 0.00; C : Z = 6.9 = 0.19; C	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35 1.05 Chi ² = 6.	0001) .01, df = 103 0.4 186	1 (P = <u>SE</u>) 551 646 348 483 3 (P = 494 671 394 6083 3 (P =	0.93), *= Weight 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); *: 12.4% 27.0% 9.1% 10.1% 58.5%	2.52 [1.05, 6.09] 7.64 [4.52, 12.29] 2.89 [1.35, 11.20] 3.89 [1.35, 11.20] 5.61 [3.45, 9.11] 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81] 4.20 [2.30, 7.64]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: Z Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009 Subtotal (95% CI)	Z = 7,37 (rences: log[O e = 0.00; C : Z = 6.9 = 0.19; C	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35 1.05 Chi ² = 6.	0001) .01, df = 103 0.4 186	1 (P = <u>SE</u>) 551 646 348 483 3 (P = 494 671 394 6083 3 (P =	0.93), *= Weight 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); *: 12.4% 27.0% 9.1% 10.1% 58.5%	2.52 [1.05, 6.09] 7.64 [4.52, 12.29] 2.89 [1.35, 11.20] 3.89 [1.35, 11.20] 5.61 [3.45, 9.11] 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81] 4.20 [2.30, 7.64]	0.01	Favours (CI) Odds	Favours [NCI]	100 —
Test for overall effect: Z Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009 Subtotal (95% CI) Heterogeneity: Tau ² =	Z = 7,37 (rences: log[O e = 0.00; C : Z = 6.9 = 0.19; C	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35 1.05 Chi ² = 6.	0001) .01, df = 103 0.4 186	1 (P = <u>SE</u>) 551 646 348 3 (P = 494 671 394 6083 3 (P =	0.93), *= Weight 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); *: 12.4% 27.0% 9.1% 10.1% 58.5%	2.52 [1.05, 6.09] 7.64 [4.52, 12.29] 2.89 [1.35, 11.20] 3.89 [1.35, 11.20] 5.61 [3.45, 9.11] 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81] 4.20 [2.30, 7.64]	0.01	Favours (CI) Odds	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect	Z = 7,37 (rences: log[O e = 0.00; C : Z = 6.9 = 0.19; C : Z = 4.6	(P < 0.00 Chi ² = 0 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35 1.05 Chi ² = 6. 9 (P < 0	0001) .01, df = 103 0.4 186	1 (P = <u>SE</u>) 551 646 348 483 3 (P = 494 671 394 6083 3 (P =	0.93), ² = <u>Weight</u> 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); ² : 12.4% 27.0% 9.1% 10.1% 58.5% 0.10); ² : 100.0%	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0% 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81] 4.20 [2.30, 7.64] = 52% 5.01 [3.57, 7.04]		Favours [CI] Odds IV, Rando	Favours [NCI]	_
Test for overall effect: 2 Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect	Z = 7,37 (rences: log[O e = 0.00; C : Z = 6.9 = 0.19; C : Z = 4.6 = 0.04; C	(P < 0.00) Chi ² = 0 1.43 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35 1.05 Chi ² = 6. 9 (P < 0 Chi ² = 8.	0001) .01, df = 103 0.4 186	1 (P = <u>SE</u>) 551 646 348 3 (P = 494 671 394 6083 3 (P = 7 (P =	0.93), ² = <u>Weight</u> 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); ² : 12.4% 27.0% 9.1% 10.1% 58.5% 0.10); ² : 100.0%	0% Odds Ratio <u>IV, Random, 95% Cl</u> 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0% 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81] 4.20 [2.30, 7.64] = 52% 5.01 [3.57, 7.04]	0.05	Favours [CI] Odds IV, Rando	Favours [NCI]	100
Test for overall effect: 2 Test for subgroup diffe c adjusted for ag Kratz 2015 Osse 2012 Reissmüller 2006 Yukako 2016 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect d not adjusted Brouquet 2010 Kalisvaart 2006 Lee 2016 Witlox 2009 Subtotal (95% CI) Heterogeneity: Tau ² = Test for overall effect Total (95% CI) Heterogeneity: Tau ² =	Z = 7,37 (rences: log[O e = 0.00; C : Z = 6.9 = 0.19; C : Z = 4.6 = 0.04; C : Z = 9.3	(P < 0.00) Chi ² = 0 1.43 1.43 1.45 2.42 1.99 Chi ² = 2. 6 (P < 0 0.92 2.03 1.35 1.05 Chi ² = 6. 9 (P < 0 Chi ² = 8. 0 (P < 0	0001) .01, df = 103 0.4 186	1 (P = <u>SE</u>) 551 646 348 483 3 (P = 494 671 394 5083 3 (P = 7 (P =	0.93), *= Weight 12.1% 11.7% 5.2% 12.5% 41.5% 0.57); *: 12.4% 27.0% 9.1% 10.1% 58.5% 0.10); *: 100.0% 0.30); *:	• 0% Odds Ratio IV, Random, 95% Cl 4.18 [1.71, 10.20] 4.30 [1.73, 10.69] 11.30 [2.68, 47.70] 7.36 [3.06, 17.72] 5.61 [3.45, 9.11] = 0% 2.52 [1.05, 6.09] 7.64 [4.52, 12.89] 3.89 [1.35, 11.20] 2.88 [1.07, 7.81] 4.20 [2.30, 7.64] = 52% 5.01 [3.57, 7.04] = 17%	0.05	Favours [CI] Odds IV, Rando	Favours [NCI]	_

study "Moncada 2005" [35], and when the risk ratio (RR) or risk difference (RD) was calculated, the rates were higher (Table 3).

We carried out another subgroup analysis by whether the eight studies using MMSE preoperatively adjusted

for age to decrease the effects of age on postoperative delirium (Fig. 3c, d). After adjusting for age, the rate of POD was higher in the preoperatively diagnosed CI group than in the preoperatively diagnosed NCI group (OR, 5.61; 95% CI, 3.45 to 9.11).

comparison	Point estimate(95% CI)	Р	²
Primary analysis	OR 5.55(0.91, 33.88)	0.06	84%
Sensitivity analysis			
Exclude the study "Moncada 2005" [35]	OR 13.56(2.24, 81.97)	0.005	66%
Alter effect measure: Relative risk	RR 2.78(1.27, 6.05)	0.01	84%
Alter effect measure: Relative difference	RD 0.39(0.04, 0.74)	0.03	92%

Table 3 Sensitivity analysis of postoperative delirium in the postoperative subgroup. Figures are Mantel-Hanzel point estimates

Length of stay in hospitals

We included 8 studies [19, 22, 28, 31, 33–35, 38] consisting of 913 patients with NCI and 773 with CI for meta-analysis of postoperative MMSE and length of stay in hospital. CI did not increase length of stay (SMD, 0.01; 95% CI, – 0.18 to 0.20; P = 0.91) (Fig. 4). Three of the trials [28, 31, 35] reported length of stay in orthopedic wards or in the acute perioperative phase. Seven trials [19, 22, 28, 33–35, 38] reported stay length in geriatric wards or rehabilitation wards after the acute phase. We then conducted a subgroup analysis and found CI did not increase length of stay in in orthopedic wards (SMD, – 0.10; 95% CI, – 0.20 to 0.17; P = 0.91) nor in rehabilitation wards (SMD, 0.04; 95% CI, – 0.23 to 0.31; P = 0.78) (Fig. 4a, b).

Two studies [7, 29], using the MMSE preoperatively, defined length of stay in a confusing and inconsistent manner and could not be combined in the meta-analysis.

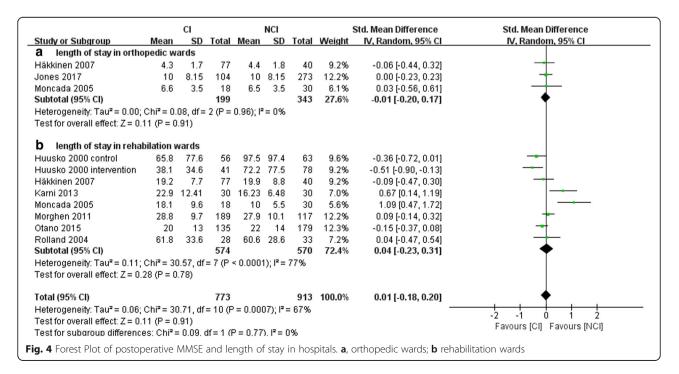
Mortality

Nine studies [19, 21, 23, 26–29, 31, 38] reported mortality including 1318 patients with NCI and 1204 with CI (Fig. 5). Patients with CI had a higher rate in mortality compared to those with NCI (OR, 2.65; 95% CI, 2.00 to 3.50; P < 0.00001). Heterogeneity between the trials was low (I² = 0%).

In the subgroup analysis, patients with preoperative CI [26, 27, 29] presented an increased rate of death. (OR, 2.52; 95% CI, 1.67 to 3.80; P < 0.0001). Similar results were also found in the postoperative CI [19, 21, 23, 28, 31, 38] group (OR, 2.72; 95% CI, 1.77 to 4.17; P < 0.00001) (Fig. 5a). The group with perioperative CI had a higher rate of in-hospital mortality [26, 38] (OR, 7.51; 95% CI, 2.17 to 26.02; P = 0.001) and mortality within one year [19, 21, 23, 27–29, 31] (OR, 2.53; 95% CI, 1.95 to 3.29; P < 0.00001;) (Fig. 5b).

Publication bias and quality of evidence

Quantitative synthesis of POD and mortality involved 11 and 10 sets of data, respectively; thus, we generated funnel plots. To exclude the existence of publication bias by visual inspection, we conducted the Egger test and found there was no evidence of publication bias for the outcomes of POD (p = 0.626) and mortality (p = 0.520) (Fig. 6).

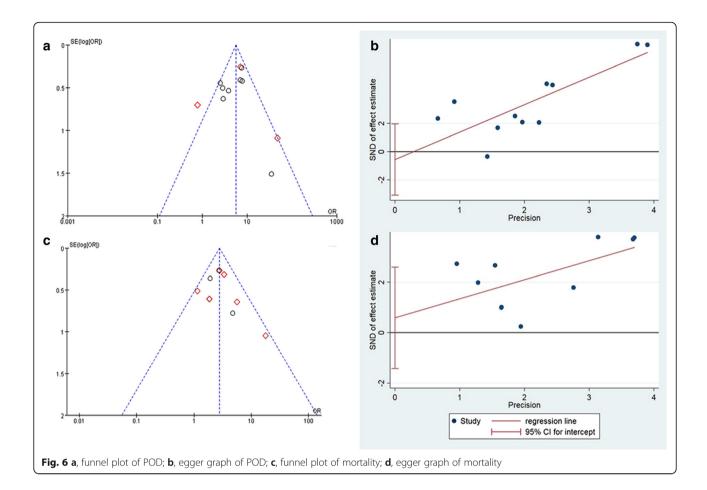


	CI		NCI			Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% Cl	M-H, Random, 95% Cl
preoperative mmse							
Beloosesky 2002	13	92	2	59	3.2%	4.69 [1.02, 21.60]	
Bliemel 2015	89	258	21	130	21.4%	2.73 [1.60, 4.66]	
Guo 2014	15	43	44	201	13.3%	1.91 [0.94, 3.89]	
Subtotal (95% CI)		393		390	37.9%	2.52 [1.67, 3.80]	
Total events	117		67				
Heterogeneity: Tau ² = 0.00			2 (P = 0.5	2); l ² =	0%		
Test for overall effect: Z = 4	.42 (P < 0.0	JOO1)					
postoperative MMSE							
Huusko 2000 control	10	63	8	56	7.1%	1.13 [0.41, 3.10]	
Huusko 2000 intervention	13	78	4	41	5.2%	1.85 (0.56, 6.09)	
Häkkinen 2007	13	77	4	40	5.2%	1.83 [0.55, 6.03]	
Jones 2017	24	104	23	279	16.6%	3.34 [1.79, 6.24]	
Otano 2015	12	135	1	179	1.8%	17.37 [2.23, 135.29]	
Ruggiero 2016	54	260	22	254	21.6%	2.76 [1.63, 4.70]	 -
Schaller 2012	17	94	3	79	4.6%	5.59 [1.57, 19.87]	
Subtotal (95% CI)		811		928	62.1%	2.72 [1.77, 4.17]	◆
Total events	143		65				
Heterogeneity: Tau ² = 0.10		0, df =	6 (P = 0.2	20); I ² =	30%		
Test for overall effect: Z = 4							
	CI	,	NCI			Odds Ratio	Odds Ratio
Study or Subgroup	Events	Total	Events	Total	Weight	M-H, Random, 95% CI	M-H, Random, 95% Cl
in-hospital mortality							
Beloosesky 2002	13	92	2	59	3.2%	4.69 [1.02, 21.60]	
			1	179	1.8%		
Otano 2015	12	135				17.37 12.23. 135.29	
Otano 2015 Subtotal (95% CI)	12	135 227				17.37 [2.23, 135.29] 7.51 [2.17, 26.02]	
Subtotal (95% CI)	12 25	227	3	238	5.0%	7.51 [2.17, 26.02]	
Subtotal (95% CI) Total events	25	227	3	238	5.0%		
Subtotal (95% CI)	25 ; Chi² = 1.0	227 2, df =	3	238	5.0%		
Subtotal (95% CI) Total events Heterogeneity: Tau² = 0.02 Test for overall effect: Z = 3	25 ; Chi² = 1.0 .18 (P = 0.0	227 2, df =	3	238	5.0%		
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes	25 ; Chi² = 1.0 .18 (P = 0.0	227 2, df = 001)	3 1 (P = 0.3	238 1); I ^z =	5.0 % 2%	7.51 [2.17, 26.02]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yea Bliemel 2015	25 ; Chi² = 1.0 .18 (P = 0.0 ar 89	227 2, df = 001) 258	3 1 (P = 0.3 21	238 1); I ^z = 130	5.0 % 2% 21.4%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014	25 ; Chi² = 1.0 .18 (P = 0.0 ar 89 15	227 2, df = 001) 258 43	3 1 (P = 0.3 21 44	238 1); I ² = 130 201	5.0% 2% 21.4% 13.3%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control	25 ; Chi ^z = 1.0 .18 (P = 0.0 ar 89 15 10	227 2, df = 001) 258 43 63	3 1 (P = 0.3 21 44 8	238 1); I ² = 130 201 56	5.0% 2% 21.4% 13.3% 7.1%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention	25 ; Chi ² = 1.0 .18 (P = 0.0 ar 89 15 10 13	227 2, df = 001) 258 43 63 78	3 1 (P = 0.3 21 44 8 4	238 1); I ² = 130 201 56 41	5.0% 2% 21.4% 13.3% 7.1% 5.2%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007	25 ; Chi² = 1.0 .18 (P = 0.0 ar 89 15 10 13 13	227 2, df = 001) 258 43 63 78 78 77	3 1 (P = 0.3 21 44 8 4 4 4	238 1); I ² = 130 201 56 41 40	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.03] 1.83 [0.55, 6.03]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yea Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017	25 ; Chi [≠] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24	227 2, df = 001) 258 43 63 78 77 104	3 1 (P = 0.3 21 44 8 4 4 23	238 (1); ² = 130 201 56 41 40 279	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6%	2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.03] 3.34 [1.79, 6.24]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016	25 ; Chi [≠] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54	227 2, df = 001) 258 43 63 78 77 104 260	3 1 (P = 0.3 21 44 8 4 4 23 22	238 (1); ² = 130 201 56 41 40 279 254	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6%	2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012	25 ; Chi [≠] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24	227 2, df = 001) 258 43 63 78 77 104 260 94	3 1 (P = 0.3 21 44 8 4 4 23	238 :1); ² = 130 201 56 41 40 279 254 79	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 4.6%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI)	25 ; Chi ² = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 7	227 2, df = 001) 258 43 63 78 77 104 260	3 1 (P = 0.3 21 44 8 4 4 23 22 3	238 (1); ² = 130 201 56 41 40 279 254	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6%	2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events	25 ; Chi [≥] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 13 24 54 17 235	227 2, df = 001) 258 43 63 78 77 104 260 94 977	3 1 (P = 0.3 21 44 8 4 23 22 3 129	238 (1); ² = 130 201 56 41 40 279 254 79 1080	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 4.6% 95.0%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI)	25 ; Chi [≠] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 17 235 ; Chi [≠] = 6.0	227 2, df = 001) 258 43 63 78 77 104 260 94 977 5, df =	3 1 (P = 0.3 21 44 8 4 23 22 3 129 7 (P = 0.5	238 (1); ² = 130 201 56 41 40 279 254 79 1080	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 4.6% 95.0%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.00 Test for overall effect: Z = 6	25 ; Chi [≠] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 17 235 ; Chi [≠] = 6.0	227 2, df = 001) 258 43 63 78 77 104 260 94 977 5, df = 00001)	3 1 (P = 0.3 21 44 8 4 23 22 3 129 7 (P = 0.5	238 (1); ² = 130 201 56 41 40 279 254 79 1080 (3); ² =	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 4.6% 95.0% 0%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87] 2.53 [1.95, 3.29]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.00 Test for overall effect: Z = 6 Total (95% CI)	25 ; Chi [₽] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 17 235 ; Chi [₽] = 6.0 .96 (P < 0.0	227 2, df = 001) 258 43 63 78 77 104 260 94 977 5, df =	3 21 44 8 4 23 22 3 129 7 (P = 0.5	238 (1); ² = 130 201 56 41 40 279 254 79 1080 (3); ² =	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 4.6% 95.0%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.00 Test for overall effect: Z = 6 Total (95% CI) Total events	25 ; Chi [≥] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 17 235 ; Chi [≥] = 6.0 .96 (P < 0.0	227 2, df = 001) 258 43 63 78 77 104 260 977 5, df = 00001) 1204	3 1 (P = 0.3 21 44 8 4 23 22 3 129 7 (P = 0.5	238 130 201 56 41 40 279 254 79 1080 3); ² = 1318	5.0% 2% 21.4% 13.3% 7.1% 5.2% 16.6% 21.6% 4.6% 95.0% 0%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87] 2.53 [1.95, 3.29]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Total events Heterogeneity: Tau ² = 0.02	25 ; Chi ^z = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 17 235 ; Chi ^z = 6.0 .96 (P < 0.0 ; Chi ^z = 10.	227 2, df = 001) 258 43 63 78 77 104 260 94 977 5, df = 00001) 1204 02, df =	3 1 (P = 0.3 21 44 8 4 4 22 3 129 7 (P = 0.5 132 = 9 (P = 0	238 130 201 56 41 40 279 254 79 1080 3); ² = 1318	5.0% 2% 21.4% 13.3% 7.1% 5.2% 16.6% 21.6% 4.6% 95.0% 0%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87] 2.53 [1.95, 3.29]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.00 Test for overall effect: Z = 6 Total (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 6	25 ; Chi [≥] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 54 17 235 ; Chi [≥] = 6.0 .96 (P < 0.0 260 ; Chi [≥] = 10. .84 (P < 0.0	227 2, df = 001) 258 43 63 78 77 104 260 94 977 5, df = 00001) 1204 02, df = 00001)	3 21 44 4 23 22 3 129 7 (P = 0.5	238 130 201 56 41 40 279 254 79 1080 33); ² = 1318 35); ² :	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 95.0% 0% 100.0% = 10%	2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87] 2.53 [1.95, 3.29] 2.65 [2.00, 3.50]	0.01 0.1 1 10 100 Favours [CI] Favours [NCI]
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: Z = 3 mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Total events Heterogeneity: Tau ² = 0.02	25 ; Chi [≥] = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 54 17 235 ; Chi [≥] = 6.0 .96 (P < 0.0 260 ; Chi [≥] = 10. .84 (P < 0.0	227 2, df = 001) 258 43 63 78 77 104 260 94 977 5, df = 00001) 1204 02, df = 00001)	3 21 44 4 23 22 3 129 7 (P = 0.5	238 130 201 56 41 40 279 254 79 1080 33); ² = 1318 35); ² :	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 95.0% 0% 100.0% = 10%	2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87] 2.53 [1.95, 3.29] 2.65 [2.00, 3.50]	
Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: $Z = 3$ mortality within 1 yes Bliemel 2015 Guo 2014 Huusko 2000 control Huusko 2000 intervention Häkkinen 2007 Jones 2017 Ruggiero 2016 Schaller 2012 Subtotal (95% CI) Total events Heterogeneity: Tau ² = 0.00 Test for overall effect: $Z = 6$ Total (95% CI) Total events Heterogeneity: Tau ² = 0.02 Test for overall effect: $Z = 6$ Test for overall effect: $Z = 6$	25 ; Chi ² = 1.0 .18 (P = 0.0 ar 89 15 10 13 13 24 54 54 17 235 ; Chi ² = 6.0 .96 (P < 0.0 ; Chi ² = 10. .84 (P < 0.0 pes: Chi ² =	227 2, df = 258 43 63 78 77 104 260 94 977 5, df = 00001) 2.82, df	3 21 21 4 4 8 4 4 23 22 3 7 (P = 0.5 7 (P = 0.5 7 (P = 0.5 7 (P = 0.5) 132 = 9 (P = 0) 132	238 1); ² = 130 201 56 41 40 279 254 79 1080 3); ² = 1318 35); ² ; 0.09),	5.0% 2% 21.4% 13.3% 7.1% 5.2% 5.2% 16.6% 21.6% 95.0% 0% 100.0% = 10%	7.51 [2.17, 26.02] 2.73 [1.60, 4.66] 1.91 [0.94, 3.89] 1.13 [0.41, 3.10] 1.85 [0.56, 6.09] 1.83 [0.55, 6.03] 3.34 [1.79, 6.24] 2.76 [1.63, 4.70] 5.59 [1.57, 19.87] 2.53 [1.95, 3.29] 2.65 [2.00, 3.50]	

Based on the GRADE approach, we found a high quality of evidence for POD in the preoperative subgroup. We found a moderate quality of evidence for POD in the postoperative subgroup and length of stay in orthopedic wards (Table 4). We did not rate the quality of evidence for mortality as we included one RCT and nine cohort studies in this outcome.

Discussion

The principal findings of our meta-analysis are that older patients with perioperative diagnoses of cognitive impairment by the MMSE had higher risk of postoperative delirium, in-hospital mortality and mortality within 1 year. We investigated the timing of assessment with the MMSE and the respective effect on adverse outcomes. Preoperative diagnosis of CI appeared to yield a more significant association with postoperative delirium than postoperatively diagnosed CI, however, according to the sensitivity analysis, the wide variance in the observed effect was due to high heterogeneity, primarily due to the inclusion of the study of Moncada's; therefore, definitive conclusions cannot be drawn. We suspected that as postoperative delirium is most common on the first and third postoperative days [39], the results of the meta-analysis may have been altered if the onset of delirium was earlier than the timing of postoperative diagnosis of CI; thus, preoperative use of the MMSE may be preferable to predict the incidence of postoperative delirium. Postoperative CI patients did not have an extended stay length in orthopedic wards or rehabilitation wards.



However, these results raised questions about insufficient sample sizes and heterogeneity in the eligible studies.

Our systematic review and meta-analysis was novel in providing data showing the predictive value of perioperative assessment by the MMSE on postoperative outcomes in older patients. We also took into consideration the timing of assessment of cognition, as preoperative cognitive impairment revealed a chronic agingrelated change while postoperative CI usually developed with acute onset and was confounded by surgery, anesthesia, medication and a stress response [40]. We summarized the data from 21 studies and used both the NOS and the Cochrane risk of bias tool to appraise the quality of selected cohort studies and an RCT, respectively. The methodological quality of studies in our review was fair to high. No significant bias of publication was observed in the report of postoperative delirium and mortality. These factors contributed to more powerful evidence than any single study or previous systematic review that failed to conduct a quantitative data synthesis.

However, there are some potential limitations should be considered. One important limitation of our study was the innate defects and the use of inconsistent cutoff scores for the MMSE across studies, which might have impacted the positive diagnosis rate of cognitive impairment. The MMSE has been validated widely around the world; however, it has been suggested that the MMSE does not perform better as a rule-out tool than a definitive diagnostic tool, which means that for those positive on the MMSE, a more detailed evaluation and inspection are required [12]. The MMSE is not the most efficient tool. Typically, it will take 8 min to complete an assessment in NCI individuals while taking 15 min or longer to evaluate CI patients [41]. Previous studies have indicated that the MMSE would be most suitable in specialist settings compared with community setting and primary care settings due to significant intraobserver differences [42]. The MMSE does not perform well enough in patients with mild cognitive impairment and early dementia [43]. Performance can be disrupted by education, age, language, ethnicity and cultural differences [44], thus optimal cutoff values change in different clinical settings. For instance, threshold values of 21, 23 and 24 are suggested in populations with primary school, high school and university education, respectively [12]. Therefore, we decided to group our study population based on a clear definition of CI by every investigator instead of selecting a specific cutoff score.

Certain	Certainty assessment						N⁰ of patients	ients	Effect		Certainty	Importance
Nº of studies		Risk of bias	Inconsistency	Indirectness	Imprecision	Study design Risk of Inconsistency Indirectness Imprecision Other considerations bias	Jo	NCI	Relative (95% CI)	Absolute (95% CI)		
POD - F	POD - preoperative MMSE	MSE										
Ø	observational not studies seri	not serious	not serious	not serious not	not serious	publication bias strongly suspected very strong association all plausible residual confounding would reduce the demonstrated effect ^a	146/380 (38.4%)	146/380 156/1108 (38.4%) (14.1%)	OR 5.12 (3.46 to 7.59)	315 more per 1000 (from 221 more to 414 more)	ФФФ НІСН	CRITICAL
POD - F	POD - postoperative MMSE	IMSE										
m	observational not studies seric	not serious	serious ^b	not serious serious ^c	serious ^c	very strong association all plausible residual confounding would reduce the demonstrated effect	102/163 62/303 (62.6%) (20.5%)	62/303 (20.5%)	OR 5.55 (0.91 to 33.88)	383 more per 1000 (from 15 fewer to 692 more)	000erate	CRITICAL
length	length of stay in orthopedic wards-postoperative MMSE	pedic wa.	rds-postoperativ	ve MMSE								
m	observational not studies seric	not serious	not serious	not serious not	not serious	all plausible residual confounding would reduce the demonstrated effect	199	343	I	SMD 0.01 lower (0.2 lower to 0.17 higher)	000erate Moderate	IMPORTANT
length	length of stay in rehabilitation wards -postoperative MMSE	ilitation v	vards -postoper.	ative MMSE								
9	observational not studies seric	not serious	serious ^d	not serious not	not serious	all plausible residual confounding would reduce the demonstrated effect	477	429	I	SMD 0.2 higher (0.1 lower to 0.49 higher)	00⊕⊕ NO1	IMPORTANT
<i>Cl</i> Confidenc <i>Explanations</i> ^a asymmetric ^b l ² = 84%	CI Confidence interval, OR Odds ratio, SMD Standardised mean difference Explanations ^a asymmetrical funnel plot ^b l ² = 84% ^c OR 5, 55(0,91,33,88)	R Odds ra ot	tio, <i>SMD</i> Standar	dised mean dif	ference							

Cao et al. BMC Anesthesiology

Page 10 of 13

Another limitation was that some of our included studies lacked clear discharge criteria and detailed rehabilitation interventions, making the outcome of length of stay less convincing. Selection and measurement biases existed as medical and rehabilitation strength varied from different districts and hospitals. Consequently, the results found in our study need to be interpreted with caution.

Most of our included studies have restricted enrollment or matched the two cognitive groups in terms of age, sex, education levels, and surgery types and so on, even though the variables in individual studies were different. There was an inconsistency distribution of age between the two cognitive groups in some studies (Additional file 2: Material 2), and age did have a negative effect on postoperative outcomes (Additional file 3: Material 3). We could reasonably have concluded that age may act as a confounder. Adjustment for a wide range of potential confounders in individual studies were listed in Additional file 4: Material 4. We carried out a subgroup analysis by whether studies adjusted for age to decrease the effects of age on postoperative delirium. we failed to do Subgroup analysis of other outcomes owing to the lack of adjusted ORs in original studies.

Delirium and dementia are among the most common causes of cognitive impairment in clinical settings, yet their interrelationship remains poorly understood and they are often either unrecognized or mistaken for each other [45]. Our study had limitations that warrant consideration. Cognitive status was measured using the general screening tool, MMSE, and did not specifically identify dementia or delirium., Almost all our included studies chose not to use the more specific concepts of delirium or dementia, as they frequently relied on second-hand observations or precise diagnostic methods instead of a single examination like MMSE, and thus were often under- or misdiagnosed, especially by nonpsychiatric staffs. Instead, they investigated the broader concept of cognitive impairment using MMSE. Thus we classified the patients into two cognitive groups and thereby minimize its false positive and negative results. We have not investigated whether acute or chronical cognitive impairment had different surgical outcomes as almost all our included studies did not provide baseline level of MMSE or a clear history of pre-existing dementia prior to admission or earlier. As a matter of fact, it was hard and exhausting for clinicians to get these medical history of patient especially geriatric ones in clinical settings. Previous studies proved that both chronic and acute cognitive impairment were independent risk factors for a poorer outcome after hip fracture [46-50]. Yet it still remained unclear which one would be involved with worse surgical outcomes. Beyond all question, dementia as well as delirium during an acute and intense stress, like surgical procedures or hip fracture, unveils the preexisting or subclinical frailty of the geriatric individual. It is critically important for Clinicians and researchers to screen for this part of surgical patients and pay attention for the presence of a frailty syndrome, which make sense of our study and the perioperative application of MMSE. Because of its wide acceptance as a general screening instrument for cognitive dysfunction, and because the test can be performed at the bedside within a relatively short period of time.

Some directions for future research should be drawn out from our findings. With the population aging, the demand for surgery in the elderly is growing. There are striking differences in the tolerance, recovery and clinical outcomes between cognitive impaired and cognitive intact older patients following surgery. The American College of Surgeons has provided both preoperative and postoperative rounding checklists for geriatric surgical patients, including a strong recommendation of the assessment and documentation of cognitive dysfunction [51, 52]. Our review further recommends that the MMSE may be used as a reliable preoperative screening tool as well as a postoperative follow-up index in geriatric surgical settings to optimize risk stratification, assess prognosis in this population, and provide indications for early and effective interventions. Moreover, further research is needed to search for more and better assessment instruments to help make clinical decisions for older patients.

Conclusions

We found that older patients with perioperative CI were more likely to suffer from postoperative delirium, in-hospital mortality and mortality within 1 year. The MMSE showed certain value on risk stratification and prognosis evaluation in the geriatric surgical population.

Additional files

Additional file 1: Search strategy for PubMed. (DOCX 13 kb)
Additional file 2: Age characteristics. (DOCX 15 kb)
Additional file 3: The association of age and outcomes of interest.
(DOCX 14 kb)
Additional file 4: Adjustment for possible confounders. (DOCX 14 kb)

Abbreviations

CI: Cognitive impairment; CI: Confidence interval; GRADE: Grading of Recommendations, Assessment, Development, and Evaluation; MMSE: Mini-Mental State Examination; NCI: non- Cognitive impairment; NOS: Newcastle Ottawa Quality Assessment Scale; OR: Odds ratio; POD: Postoperative delirium; RCT: Randomized controlled trial; RD: Risk difference; RR: Risk ratio; SMD: Standard mean difference

Acknowledgements

Not applicable.

Funding

This review did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Availability of data and materials

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

SJC and DXC carried out the study selection, data collection and analysis, they participanted in drafting the manuscript. SJC and DXC have contributed equally to this review. Disagreement was resolved by consulting LY and TZ. LY and TZ participanted in revising the manuscript critically. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 9 February 2019 Accepted: 18 April 2019 Published online: 15 May 2019

References

- Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. Lancet. 2017;389:1323–35.
- Bravo J, Henning S, Gu D, Pelletier F. World population projected to reach 9. 8 billion in 2050, and 11.2 billion in 2100 – says UN - United Nations sustainable development. Sustainable Development Goals. 2017:14 http:// www.un.org/sustainabledevelopment/blog/2017/06/world-populationprojected-to-reach-9-8-billion-in-2050-and-11-2-billion-in-2100-says-un/.
- Lin H-S, Watts JN, Peel NM, Hubbard RE. Frailty and post-operative outcomes in older surgical patients: a systematic review. BMC Geriatr. 2016; 16:157. https://doi.org/10.1186/s12877-016-0329-8.
- Deiner S, Westlake B, Dutton RP. Patterns of surgical care and complications in elderly adults. J Am Geriatr Soc. 2014;62:829–35.
- Makary MA, Segev DL, Pronovost PJ, Syin D, Bandeen-Roche K, Patel P, et al. Frailty as a predictor of surgical outcomes in older patients. J Am Coll Surg. 2010;210:901–8.
- Tse L, Bowering JB, Schwarz SKW, Moore RL, Sztramko R, Barr AM. Incidence and risk factors for impaired mobility in older cardiac surgery patients during the early postoperative period. Geriatr Gerontol Int. 2015;15:276–81. https://doi.org/10.1111/ggi.12269.
- Lee YS, Kim YB, Lee SH, Park YS, Park SW. The prevalence of undiagnosed presurgical cognitive impairment and its postsurgical clinical impact in older patients undergoing lumbar spine surgery. J Korean Neurosurg Soc. 2016;59:287–91.
- McRae PJ, Peel NM, Walker PJ, De Looze JWM, Mudge AM. Geriatric syndromes in individuals admitted to vascular and urology surgical units. J Am Geriatr Soc. 2014;62:1105–9.
- Franzo A, Francescutti C, Simon G. Risk factors correlated with postoperative mortality for hip fracture surgery in the elderly: a populationbased approach. Eur J Epidemiol. 2005;20:985–91.
- Muir SW, Yohannes B, Abebaw M. The impact of cognitive impairment on rehabilitation outcomes in elderly patients admitted with a femoral neck fracture: a systematic review. J Geriatr Phys Ther. 2009;32:24–32.
- Trzepacz PT, Hochstetler H, Wang S, Walker B, Saykin AJ. Relationship between the Montreal cognitive assessment and mini-mental state examination for assessment of mild cognitive impairment in older adults. BMC Geriatr. 2015.
- 12. Larner AJ. Cognitive screening instruments. London: Springer London; 2013. https://doi.org/10.1007/978-1-4471-2452-8.

- Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 2009;339 jul21 1:b2535–b2535.
- Sedgwick P. Meta-analyses: tests of heterogeneity. BMJ (Online) 2012;344: e3971–e3971. doi:https://doi.org/10.1136/bmj.e3971.
- van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a metaepidemiological study. BMC Med Res Methodol. 2014;14:70. https://doi.org/ 10.1186/1471-2288-14-70.
- Stang A. Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. Eur J Epidemiol. 2010;25:603–5.
- Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:1–9.
- 18. Kavanagh BP. The GRADE system for rating clinical guidelines. PLoS Med. 2009.
- Huusko TM, Karppi P, Avikainen V, Kautiainen H, Sulkava R. Randomised, clinically controlled trial of intensive geriatric rehabilitation in patients with hip fracture: subgroup analysis of patients with dementia. Bmj. 2000;321:1107–11.
- Mokutani Y, Mizushima T, Yamasaki M, Rakugi H, Doki Y, Mori M. Prediction of postoperative complications following elective surgery in elderly patients with colorectal Cancer using the comprehensive geriatric assessment. Dig Surg. 2016;33:470–7.
- Ruggiero C, Bonamassa L, Pelini L, Prioletta I, Cianferotti L, Metozzi A, et al. Early post-surgical cognitive dysfunction is a risk factor for mortality among hip fracture hospitalized older persons. Osteoporos Int. 2017;28:667–75. https://doi.org/10.1007/s00198-016-3784-3.
- Rolland Y, Pillard F, Lauwers-Cances V, Busquère F, Vellas B, Lafont C. Rehabilitation outcome of elderly patients with hip fracture and cognitive impairment. Disabil Rehabil. 2004;26:425–31.
- Schaller F, Sidelnikov E, Theiler R, Egli A, Staehelin HB, Dick W, et al. Mild to moderate cognitive impairment is a major risk factor for mortality and nursing home admission in the first year after hip fracture. Bone. 2012;51: 347–52. https://doi.org/10.1016/j.bone.2012.06.004.
- 24. Witlox J, Kalisvaart KJ, De Jonghe JFM, Verwey NA, Van Stijn MFM, Houdijk APJ, et al. Cerebrospinal fluid β -amyloid and tau are not associated with risk of delirium: a prospective cohort study in older adults with hip fracture. J Am Geriatr Soc. 2011;59:1260–7.
- Brouquet A, Cudennec T, Benoist S, Moulias S, Beauchet A, Penna C, et al. Impaired mobility, ASA status and administration of tramadol are risk factors for postoperative delirium in patients aged 75 years or more after major abdominal surgery. Ann Surg. 2010;251:759–65.
- Beloosesky Y, Grinblat J, Epelboym B, Weiss A, Grosman B, Hendel D. Functional gain of hip fracture patients in different cognitive and functional groups. Clin Rehabil. 2002;16:321–8. https://doi.org/10.1191/0269215502cr497oa.
- Bliemel C, Sielski R, Doering B, Dodel R, Balzer-Geldsetzer M, Ruchholtz S, et al. Pre-fracture quality of life predicts 1-year survival in elderly patients with hip fracture—development of a new scoring system. Osteoporos Int. 2016; 27:1979–87. https://doi.org/10.1007/s00198-015-3472-8.
- Häkkinen A, Heinonen M, Kautiainen H, Huusko T, Sulkava R, Karppi P. Effect of cognitive impairment on basic activities of daily living in hip fracture patients: a 1-year follow-up. Aging Clin Exp Res. 2007;19:139–44.
- Guo Y, Sun T, Wang X, Li S, Liu Z. Cognitive impairment and 1-year outcome in elderly patients with hip fracture. Med Sci Monit. 2014;20:1963– 8. https://doi.org/10.12659/MSM.892304.
- Kalisvaart KJ, Vreeswijk R, De Jonghe JFM, Van Der Ploeg T, Van Gool WA, Eikelenboom P. Risk factors and prediction of postoperative delirium in elderly hip-surgery patients: implementation and validation of a medical risk factor model. J Am Geriatr Soc. 2006;54:817–22.
- Jones CA, Jhangri GS, Feeny DH, Beaupre LA. Cognitive status at hospital admission: postoperative trajectory of functional recovery for hip fracture. J Gerontol A Biol Sci Med Sci. 2017;72:61–7.
- Kratz T, Heinrich M, Schlaub E, Diefenbacher A. Preventing postoperative delirium a prospective intervention with psychogeriatric liaison on surgical wards in a general hospital. Dtsch Arztebl Int. 2015;112:289–96.
- Karni S, Bentur N, Ratzon N. Participation and quality of life of cognitively impaired older women in Israel following hip fractures. Occup Ther Int. 2014;21:91–7.
- Morghen S, Gentile S, Ricci E, Guerini F, Bellelli G, Trabucchi M. Rehabilitation of older adults with hip fracture: cognitive function and walking abilities. J Am Geriatr Soc. 2011;59:1497–502.

- Moncada LV, Andersen RE, Franckowiak SC, Christmas C. The impact of cognitive impairment on short-term outcomes of hip fracture patients. Arch Gerontol Geriatr. 2006;43:45–52.
- Osse RJ, Fekkes D, Tulen JHM, Wierdsma AI, Bogers AJJC, Van Der Mast RC, et al. High preoperative plasma neopterin predicts delirium after cardiac surgery in older adults. J Am Geriatr Soc. 2012;60:661–8.
- Veliz-Reissmüller G, H AT, J van der L, Lindblom D, M EJ. Pre-operative mild cognitive dysfunction predicts risk for post-operative delirium after elective cardiac surgery. Aging Clin Exp Res. 2007;19:172–7.
- Uriz-Otano F, Uriz-Otano JI, Malafarina V. Factors associated with short-term functional recovery in elderly people with a hip fracture. Influence of cognitive impairment. J Am Med Dir Assoc. 2015;16:215–20. https://doi.org/ 10.1016/j.jamda.2014.09.009.
- Postoperative delirium in older adults. Best practice statement from the American geriatrics society. In: J am Coll Surg; 2015.
- Nadelson MR, Sanders RD, Avidan MS. Perioperative cognitive trajectory in adults. Br J Anaesth. 2014;112:440–51.
- Larner AJ. MMSE variants and subscores. In: Cognitive screening instruments. Cham: Springer International Publishing; 2017. p. 49–66. https://doi.org/10.1007/978-3-319-44775-9_4.
- 42. Creavin ST, Noel-Storr AH, Smailagic N, Giannakou A, Ewins E, Wisniewski S, et al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer's dementia and other dementias in asymptomatic and previously clinically unevaluated people aged over 65 years in community and primary care populations. In: Creavin ST, editor. Cochrane Database of Systematic Reviews. Chichester, UK: John Wiley & Sons, Ltd; 2014. doi:https://doi.org/10. 1002/14651858.CD011145.
- Schultz-Larsen K, Lomholt RK, Kreiner S. Mini-mental status examination: a short form of MMSE was as accurate as the original MMSE in predicting dementia. J Clin Epidemiol. 2007;60:260–7.
- Rosli R, Tan MP, Gray WK, Subramanian P, Chin AV. Cognitive assessment tools in Asia: a systematic review. Int Psychogeriatr. 2016;28:189–210.
- Leonard M, McInerney S, McFarland J, Condon C, Awan F, O'Connor M, et al. Comparison of cognitive and neuropsychiatric profiles in hospitalised elderly medical patients with delirium, dementia and comorbid deliriumdementia. BMJ Open. 2016.
- Olofsson B, Lundström M, Borssén B, Nyberg L, Gustafson Y. Delirium is associated with poor rehabilitation outcome in elderly patients treated for femoral neck fractures. Scand J Caring Sci. 2005.
- 47. Söderqvist A, Miedel R, Ponzer S, Tidermark J. The influence of cognitive function on outcome after a hip fracture. J Bone Jt Surg Ser A. 2006.
- Clague JE, Craddock E, Andrew G, Horan MA, Pendleton N. Predictors of outcome following hip fracture. Admission time predicts length of stay and in-hospital mortality. Injury. 2002.
- Kassahun WT. The effects of pre-existing dementia on surgical outcomes in emergent and nonemergent general surgical procedures: assessing differences in surgical risk with dementia. BMC Geriatr. 2018;18:1–9.
- 50. by Dove Press published. Dementia and delirium, the outcomes in elderly hip fracture patients; 2017. p. 421–30.
- Chow WB, Rosenthal RA, Merkow RP, Ko CY, Esnaola NF. Optimal preoperative assessment of the geriatric surgical patient: a best practices guideline from the American college of surgeons national surgical quality improvement program and the American geriatrics society. J Am Coll Surg. 2012;215:453–66.
- Mohanty S, Rosenthal RA, Russell MM, Neuman MD, Ko CY, Esnaola NF. Optimal perioperative Management of the Geriatric Patient: a best practices guideline from the American College of Surgeons NSQIP and the American Geriatrics Society. J Am Coll Surg. 2016;222:930–47. https://doi.org/10.1016/j. jamcollsurg.2015.12.026.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

