BJA

British Journal of Anaesthesia, 120(1): 37–50 (2018)

doi: 10.1016/j.bja.2017.09.002
Advance Access Publication Date: 24 November 2017
Review Article

Perioperative outcomes in the context of mode of anaesthesia for patients undergoing hip fracture surgery: systematic review and meta-analysis

C.M. O'Donnell[1,*], L. McLoughlin[1], C.C. Patterson[2], M. Clarke[2], K.C. McCourt[1], M.E. McBrien[1], D.F. McAuley[1,3] and M.O. Shields[1]

[1]Royal Victoria Hospital, Belfast Health and Social Care Trust, Grosvenor Road, Belfast BT12 6BA, UK, [2]Centre for Public Health, School of Medicine, Dentistry and Biomedical Sciences, Institute of Clinical Sciences, Block B, Queen's University of Belfast, Belfast BT12 6BA, UK and [3]Centre for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Wellcome–Wolfson Institute, Queen's University of Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK

*Corresponding author. E-mail: codonnell11@qub.ac.uk.

Abstract

Background: Previous meta-analyses on the anaesthetic management of patients undergoing surgery for hip fracture have focused on randomized trials. Furthermore, heterogeneity in outcome reporting across the studies has made it difficult to inform best practice guidelines for patient care.

Methods: This systematic review examined how perioperative outcomes were reported and defined in the context of comparing modes of anaesthesia for hip fracture surgery. Outcomes were included from randomised and non-randomised studies published between January 2000 and July 2017. Meta-analyses were performed for regional versus general anaesthesia, with sensitivity analyses performed for spinal versus general anaesthesia.

Results: By including data from 15 large observational studies in this meta-analysis, we have increased the number of patients for whom outcomes were assessed from approximately 3000 to 202 000. There was no significant difference in 30-day mortality [Odds ratio (OR) 1.15; 95% confidence interval (CI) 1.01, 1.32; I^2 87%; n=200 464], prevalence of pneumonia (OR 1.10; 95% CI 0.93, 1.30; I^2 43%; n=65 011), acute myocardial infarction (OR 0.96; 95% CI 0.88, 1.05; I^2 0%, n=64 904), delirium (OR 1.07; 95% CI 0.72, 1.58; I^2 93%, n=19 923) or renal failure (OR 0.94; 95% CI 0.54, 1.64; I^2 0%, n=27 873) for regional compared to general anaesthesia.

There was a small statistically significant difference for length of stay (standardized mean difference −0.03; 95% CI −0.05, −0.02; I^2 0%; n=78 711) favouring regional anaesthesia, which is unlikely to be clinically significant. Sensitivity analyses for the same outcomes examining spinal only *vs* general anaesthesia showed minor statistical significance for length of stay favouring spinal. We also present data highlighting the scale of the inconsistencies in reported outcomes across 32 studies, making evaluation in a standardized manner very difficult. As an example, mortality was reported in nine different ways throughout the studies.

Conclusions: We highlight the need for agreement on outcome definitions and for a minimum core outcome set to be measured and reported in hip fracture studies. This would strengthen the evidence-based approach to delivering optimal care.

Key words: anaesthesia; general; hip fractures; outcome measures; regional

Editorial decision: September 15, 2017; **Accepted:** September 15, 2017

Approximately 1.6 million people worldwide sustain a hip fracture each year,[1] with over 76 000 of these occurring in the UK.[1,2] The risk of dying of such a fracture within 30 days is 7.1%.[1] With 70% of patients being above the age of 80 yr, 30% experiencing a reduction in functionality after a hip fracture,[3] and 20% suffering serious complications during their acute hospital stay,[4] there is a huge burden placed on the health service. The median cost per patient presenting with a hip fracture in the UK in 2011 was around £9500 for the acute phase of their treatment, with 76% of this attributed to ward costs, 14% for theatre costs, and 10% for investigations.[5] Additional costs of the follow-on treatment would increase this further. District general hospitals spend between £3.6 and £4.8 million per yr on the acute management of patients with hip fracture[6]; however, research to date has produced insufficient evidence to guide a key element in the early care for these patients, namely the anaesthetic management for their hip fracture surgery. Studies evaluating regional compared to general anaesthesia have generated inconsistent results when looking at mortality as a primary outcome. Systematic reviews previously aiming to evaluate this question have been limited by the small number and generally low quality of randomized trials. They also have excluded more recent large observational studies.

This systematic review aimed to identify studies of patients undergoing emergency hip fracture surgery in the context of the type of anaesthesia administered. It aimed to explore how perioperative hip fracture outcomes were defined and reported across these studies.[7–9] In an attempt to use the large amount of available data from observational studies, meta-analyses were performed to compare various outcomes in the context of general *vs* regional anaesthesia, with sensitivity analyses for general anaesthesia *vs* spinal only anaesthesia, drawing on data from both randomized and non-randomized studies. In performing the meta-analyses, we recognized the inconsistency and heterogeneity in outcome reporting, causing difficulty in pooling results from the various trials and observational studies for comparison. Importantly, this limits the ability to evaluate interventions used in perioperative care for patients with a hip fracture and to provide best practice recommendations,[7,10,11] and attributes partly to the minimal change in 30-day mortality for hip fracture patients over the past 5 yr within the UK.

Methods

Initially in this systematic review a quantitative analysis addressing the intervention of regional or general anaesthesia for hip fracture surgery was performed following the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[12,13] A further sensitivity analysis for those cases wherein spinal anaesthesia was the sole technique *vs* general anaesthesia was also performed.

Secondly, based on a protocol previously registered on Prospero [33405],[14] the review identifies how outcome measures pertaining to mode of anaesthesia for hip fracture surgery have been reported.

Literature search

A search of Ovid EMBASE, Ovid MEDLINE, and the Cumulative Index to Nursing Allied Health Literature (CINAHL) databases, the Cochrane Central Register for Controlled Trials, Clinical Trials.gov, the ISRCTN registry, as well as grey literature for articles published from January 2000 to July 2017 was performed to focus on the most recent studies reflective of modern anaesthetic techniques. Search terms were applied to both subject headings and as keywords and restricted to human studies and English language only. Terms included:

- Hip Fracture or (hip adj5 fractur*)
- Femur Fracture or (femur adj5 fractur*)
- Anesthesia, General/'general an?esthe*'
- Anesthesia, Spinal/'neuraxial an?esthe*' 'regional an?esthe*'
- Postoperative Complications/
- Pneumonia/'pulmonary complication*' Pulmonary Atelectasis/hypoxi*
- Acute kidney injury/or kidney tubular necrosis, acute/renal Insufficiency/'renal failure*'/'kidney failure*'
- Myocardial infarction/or shock, cardiogenic/
- Delirium/(post?operative adj5 confus*)

Additional studies were identified by hand-searching the reference sections of all eligible studies and previously published review articles.[15]

Randomized trials and observational studies, both prospective and retrospective, were eligible for this review. Studies were included if they reported perioperative outcomes in the context of comparing modes of anaesthesia for hip fracture surgery regardless of the presence or specifics of any study intervention. For quantitative analysis, only the studies that examined regional compared with general anaesthesia were included. Those studies wherein a spinal anaesthetic was performed as a sole technique compared to a general anaesthetic were included in a subsequent sensitivity analysis. All eligible studies were included regardless of size and results restricted to full-text articles.

Study selection was based on independent screening of the titles and abstracts in the initial search by two investigators (C.O. and L.M.). Identified studies underwent a full text review by the same two reviewers working independently and in duplicate to assess eligibility. Disagreement regarding study eligibility was discussed and resolved through consultation with a third author (M.S.). Case reports and case series were excluded, as were studies concerning elective hip surgery.

Two authors (C.O. and L.M.) independently assessed the risk of bias of individual studies according to the Cochrane Risk of Bias Tool for randomized controlled trials (RCTs),[16] or the Cochrane Risk of Bias Tool for non-randomized studies for the observational studies.[17] Studies were assigned a low risk of bias, high risk of bias, or unclear risk of bias for each domain in the Cochrane tools as per Appendices 1 and 2.

The same two independent reviewers extracted data from the eligible full-text articles. This included study characteristics such as author, year of publication, country, study design, study participants, sample size, and intervention as presented in Table 1. Details relating to all outcome measures reported

Table 1 Study characteristics. CI, contraindication; CVA, cerebrovascular accident; GA, general anaesthesia; RA, regional anaesthesia. The term 'RA' included patients who received neuraxial blockade (epidural or spinal) with/without nerve block. RA* indicates studies where a breakdown of data was also provided for individual techniques, e.g. spinal only. 'No significant difference' indicates no difference in any of the reported outcomes within the study

Study (1st author year)	Country	Design	Participants, *n*	Intervention	Participant characteristics	Findings
Basques[28] 2015	USA	Retrospective observational	9842	GA *vs* spinal	≥70 yr	GA increased rate of adverse events Length of stay increased in spinal group
Biboulet[25] 2012	France	RCT	45	GA *vs* spinal	>75 yr ASA 3/4	Significant reduction in hypotension in RA group
Bjorkeland[24] 2010	Sweden	Prospective quasiexperimental	263	Multifactorial intervention programme to reduce delirium outcomes stratified based on anaesthesia type GA *vs* RA	>65 yr Cognitively intact on admission	No significant difference
Brox[39] 2016	USA	Retrospective observational	7585	GA *vs* RA	All operable hip fractures Exclusions: renal disease	No significant difference
Casati[47] 2003	Italy	RCT	30	GA *vs* spinal	> 65 yr Exclusions: CIs to spinal/GA or severe pulmonary/cardiac/ psychiatric disease	Greater blood loss GA group Pain scores better in RA group
Chu[43] 2015	Taiwan	Retrospective observational	182 307	GA *vs* RA	>65 yr	In-hospital mortality, CVA, and respiratory complications reduced with RA
Edelstein[44] 2014	USA	Prospective observational	921	Effect of delirium on morbidity categorized according to GA *vs* RA	> 65 yr Exclusions: cognitive impairment	GA predictor for delirium
Fields[29] 2015	USA	Retrospective observational	6133	GA *vs* spinal	All hip fracture patients	Significant difference in all cause morbidity favouring RA
Heidari[26] 2011	Iran	RCT	387	GA *vs* RA	>30 yr ASA1-3	Increased blood loss in GA group
Hoppenstein[48] 2005	Israel	RCT	60	GA *vs* spinal	ASA 1–3 >60 yr	No significant difference
Karaman[46] 2015	Turkey	Retrospective observational	308	GA *vs* RA	>65 yr Exclusions: multiple trauma, malignancy	Significant difference in mortality in favour of RA
Le-Wendling[31] 2012	Florida USA	Retrospective observational	308	GA *vs* RA	>65 yr Exclusions: patients with end stage renal disease	No significant difference
Llango[41] 2016	Australia	Prospective observational	318	GA *vs* spinal	All hip fractures	No significant difference
Messina[49] 2013	Italy	RCT	20	GA *vs* spinal	>75 yr	Statistically significant reduction in haemodynamic instability with spinal
Neuman[32] 2014	USA	Retrospective observational	56 729	GA *vs* RA	>50 yr Exclusions: multiple trauma, other surgical procedures and local anesthesia.	Length of stay shorter in matched RA group
O'Hara[33] 2000	USA	Retrospective observational	9425	GA *vs* RA	>60 yr Exclusions: patient refusal of blood transfusion or metastatic cancer	No significant difference

Continued

Table 1 Continued

Study (1st author year)	Country	Design	Participants, *n*	Intervention	Participant characteristics	Findings
Olofsson[20] 2004	Denmark	RCT double blinded	50	Low *vs* high dose spinal	≥ ASA 2	Significant difference in hypotension favouring low dose group.
Parker[27] 2015	UK	RCT	322	GA *vs* spinal	>49 yr	No significant difference
Patorno[45] 2014	USA	Retrospective observational	73 284	GA *vs* RA	>18 yr Exclusions: closed surgery, multiple injuries, day of admission surgery	No significant difference
Radcliff[38] 2008	USA	Retrospective observational	5683	GA *vs* RA	Males >65 yr Exclusions: admissions from institution	Increased 30-day mortality in GA group
Rashid[37] 2013	Pakistan	Retrospective observational	194	GA *vs* RA	All hip fractures	No significant difference
Seitz[23] 2011	Canada	Retrospective observational	11 787	Effect of cholinesterase inhibitors stratified on basis of anaesthetic type GA *vs* RA	>65 yr With dementia Exclusions: pathological fracture or palliative care	No significant difference
Seitz[4] 2014	Canada	Retrospective observational	6135	GA *vs* RA	Older patients With dementia Exclusions: major trauma, palliative care or pathological fracture	Higher intensive care unit admission for GA group
Sevtap[30] 2012	Turkey	Retrospective observational	185	GA *vs* RA*	>60 yr	No significant difference
Sieber[21] 2010	USA	RCT double blinded	114	Targeted deep *vs* light sedation during RA	>65 yr Exclusions: preoperative delirium and severe dementia	Delirium reduced in light sedation group
Shih[34] 2010	Taiwan	Retrospective observational	335	GA *vs* spinal	>80 yr	No significant difference mortality; morbidity greater in GA group
Tung[40] 2016	Taiwan	Retrospective observational	17 189	GA *vs* RA	>18 yr Exclusions: multiple hip fracture admissions, patients receiving both GA and RA	RA associated with reduced 30 day all cause and surgical site infection readmission
De Visme[22] 2000	France	RCT	29	Spinal *vs* lumbar and sciatic plexus block	>68 yr	No significant difference
White[35] 2014	UK	Retrospective observational	65 535	GA *vs* RA*	All hip fractures	No significant difference
White[3] 2016	UK	Prospective observational	11 085	GA *vs* RA*	All hip fractures	No significant difference
Whiting[42] 2015	USA	Retrospective observational	7764	GA *vs* RA*	All hip fractures	Significant increase in minor complications with RA
Wood[36] 2011	UK	Retrospective observational	1131	GA *vs* spinal	All hip fractures	Significant difference in hypotension favouring low dose spinal

within the studies, along with their various definitions were extracted. Outcomes that had been reported in the same way in at least two randomized or non-randomized studies comparing regional and general anaesthesia or spinal and general anaesthesia were included in the meta-analyses. Event rates were extracted for dichotomous outcomes and mean and standard deviation was extracted for continuous outcomes.

A full list of all reported outcomes from eligible studies was compiled (Supplementary Table). These included 30-day mortality, acute myocardial infarction, delirium, length of stay, pneumonia, and acute renal failure. The quality of evidence for each outcome was rated by the Grades of Recommendation, Assessment, Development and Evaluation (GRADE) Working Group approach.[18]

Review Manager software (RevMan for Mac, version 5.3; Cochrane Collaboration, Oxford, UK) was used to perform the meta-analyses. Where possible, the results of studies were pooled if at least two studies reported comparable outcomes. This allowed forest plots to be generated, statistical heterogeneity to be tested, and an overall estimate of the pooled effect for each outcome. A random effects model was used due to the differences in patient population and intervention. Similarity between studies was measured using the I^2 statistic to estimate the proportion of variation across studies that is due to heterogeneity rather than chance.[19]

Results

Initially, 440 studies were identified using the search methods described. After removal of duplicates, 408 titles and abstracts were screened (Fig. 1). Sixty-seven full text articles were assessed for eligibility, 43 of which were excluded. A further eight records found through checking the reference lists of included articles and review articles were added to the remaining 24. Therefore, 32 records were included in the narrative review, with 19 contributing to the quantitative analysis. These studies were published between January 2000 and July 2017.

Characteristics of the studies are shown in Table 1. The majority, 19 out of 32, were retrospective observational studies, nine were randomized trials, and four prospective cohort studies. Twenty-seven studies looked at outcomes relating to regional compared to general anaesthesia and 10 studies examined outcomes specifically for spinal compared with general anaesthesia. One study reported outcomes comparing high and low dose spinal anaesthesia,[20] one reported outcomes for deep *vs* light sedation[21] and another compared spinal *vs* lumbar and sciatic plexus blocks.[22] Of the remaining studies, two assessed outcomes for regional and general anaesthesia following either administration of cholinesterase inhibitors or the implementation of a delirium bundle in older adults with dementia undergoing hip fracture surgery.[23,24]

The general anaesthesia group encompasses 'general anaesthesia only' and 'general anaesthesia plus nerve block'. A total of three randomized trials[25–27] containing data on 752 patients and 16 observational studies[3,4,28–41] with data on 201 254 patients were included in analyses for general *vs* regional anaesthesia. Two randomized trials[25,27] containing data on 365 patients and 10 observational studies[3,4,28–31,34–36,41] with data on 96 813 patients were included in the sensitivity analyses for general *vs* spinal anaesthesia.

The risk of selection bias within the randomized trials was judged to be low. In terms of allocation concealment, half the trials did not report if or how this was achieved (Appendix 1). All the randomized trials were at high risk of performance and detection bias. This was based on lack of blinding of participants or personnel during the intervention and during outcome assessment or a lack of providing information on how this was achieved. The risks of attrition and reporting bias were judged to be low for the randomized trials (Appendix 1). The risk of major influence due to confounding factors was deemed high in all of the observational studies with the majority using statistical techniques such as propensity score matching and logistic regression in order to adjust for these (Appendix 2). The risk of bias in the classification and deviation from intended interventions and in the selection of reported results was low. However, selection bias was high or unclear in 10 out of 18 studies, with many having unclear criteria for inclusion and exclusion. There was a moderate risk of bias due to missing data in six out of 18 studies and an unclear to high risk due to non-standardized reporting of outcomes in five out of 18 studies (Appendix 2).

Meta-analysis: outcomes for regional vs general anaesthesia

Meta-analysis of 30-day mortality data for 14 studies, two randomized trials[25,27] and 12 observational studies,[3,4,28–30,32,33,35,36,38–40] revealed no significant difference between regional and general anaesthesia [odds ratio (OR) 1.15; 95% confidence interval (CI) 1.01, 1.32, I^2 87%; n=200 464]. The quality of evidence for this outcome was moderate according to the GRADE system.[18] The results for all meta-analysis are displayed in Fig. 2. Three studies, one randomized and two observational, showed no significant difference for 1-year mortality (OR 1.27; 95% CI 0.76, 2.12, I^2 78%, n=7956).

Nine studies, three randomized trials[25–27] and six observational studies,[4,28–30,33,40] representing a total of 31 322 patients receiving regional anaesthesia and 33 582 receiving general anaesthesia showed no significant difference in rate of acute myocardial infarction (OR 0.96; 95% CI 0.88, 1.05, I^2 0%, n=64 904).

There was no significant difference in the rates of pneumonia in patients receiving either regional or general anaesthesia in the meta-analysis of eight studies; two randomized trials[26,27] and six observational[4,28,29,33,34,40] studies (OR 1.10; 95% CI 0.93, 1.30; I^2 43%; n=65 011).

Six studies reported delirium in a way that was comparable; two randomized trials[26,27] and four observational studies.[3,33,34,41] There was no significant difference in the rate of delirium between regional and general anaesthesia (OR 1.07; 95% CI 0.72, 1.58; I^2 93%; n=19 923). There was substantial heterogeneity across these studies and between the results of the two randomized (I^2 70%) and four observational studies (I^2 95%).

Six studies were also able to be included in the meta-analysis for length of stay, one randomized[27] and five observational studies.[4,30–32,37] There was a small statistically significant difference between the two types of anaesthesia for this outcome (standardized mean difference −0.03; 95% CI −0.05, −0.02; I^2 0%; n=78,711) favouring regional. The five other studies that reported results for length of stay could not be included in the meta-analysis due to absence of statistical comparators. Two papers reported hospital stay as median values,[3,34] which could not be combined in meta-analyses, and two papers reported length of hospital stay for general and regional anaesthesia combined, and thus could not be

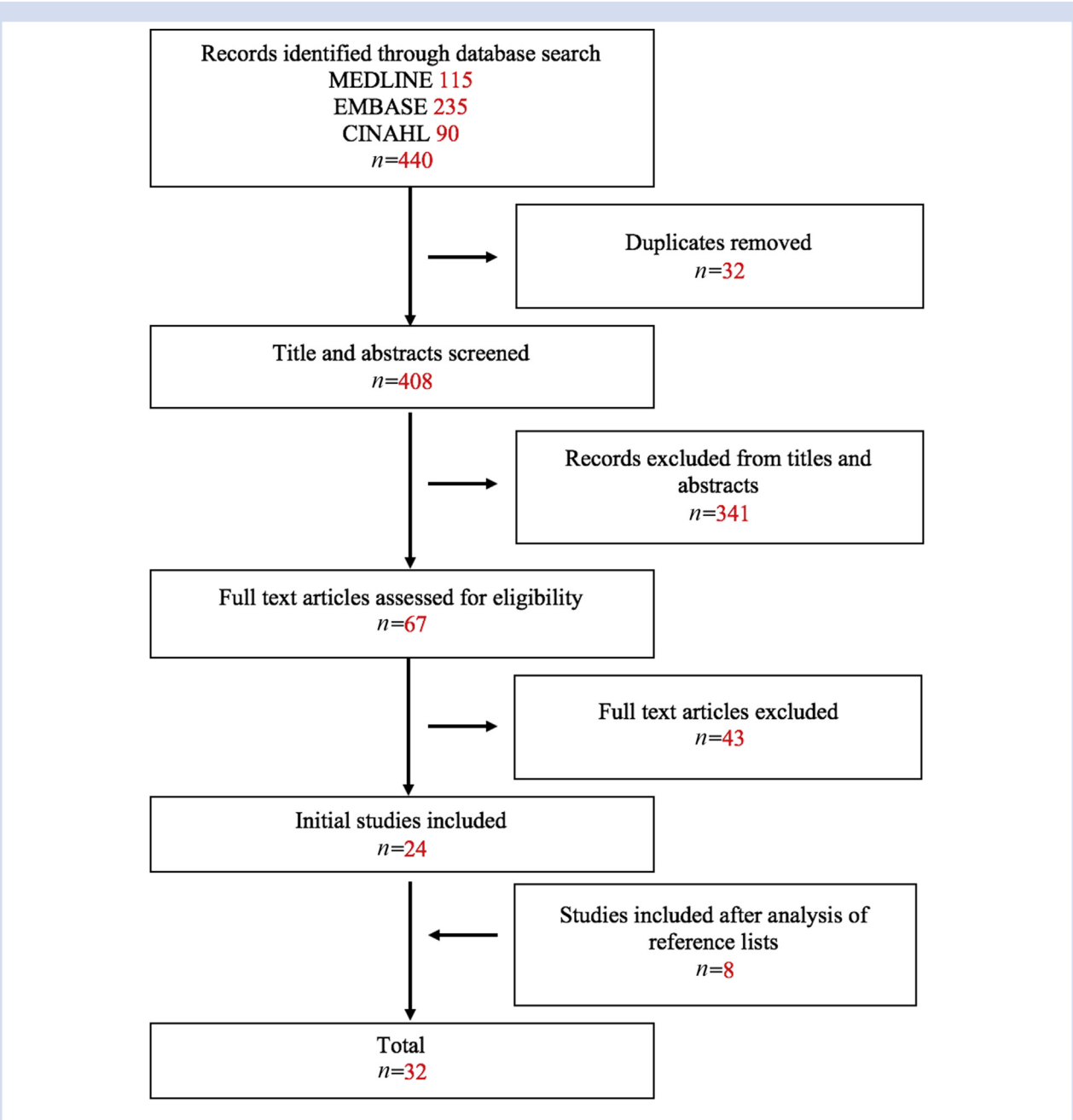


Fig 1. Search strategy.

compared.[41,44] Wood and White[36] reported 'postoperative length of stay', which was not comparable with 'hospital length of stay'. Despite the statistically significant result, such a small standard mean difference is unlikely to be of clinical significance.

Five studies reported acute renal failure in a way that could be comparable by meta-analysis, one randomized,[27] and four observational studies.[28,30,34,40] There was no significant difference between regional and general anaesthesia for this outcome (OR 0.94; 95% CI 0.54, 1.64; I^2 0%; n=27 873).

For outcomes such as hypotension, blood loss, and blood transfusion, it was not possible to perform meta-analyses because of the lack of uniformity in outcome reporting. Furthermore, there were too few comparable studies to examine mortality in the longer term, for example at 90 or 120 days. This considerable variation in reporting of outcomes, often lacking detail in definition as well as measurements at different time points, leads to difficulty comparing studies.

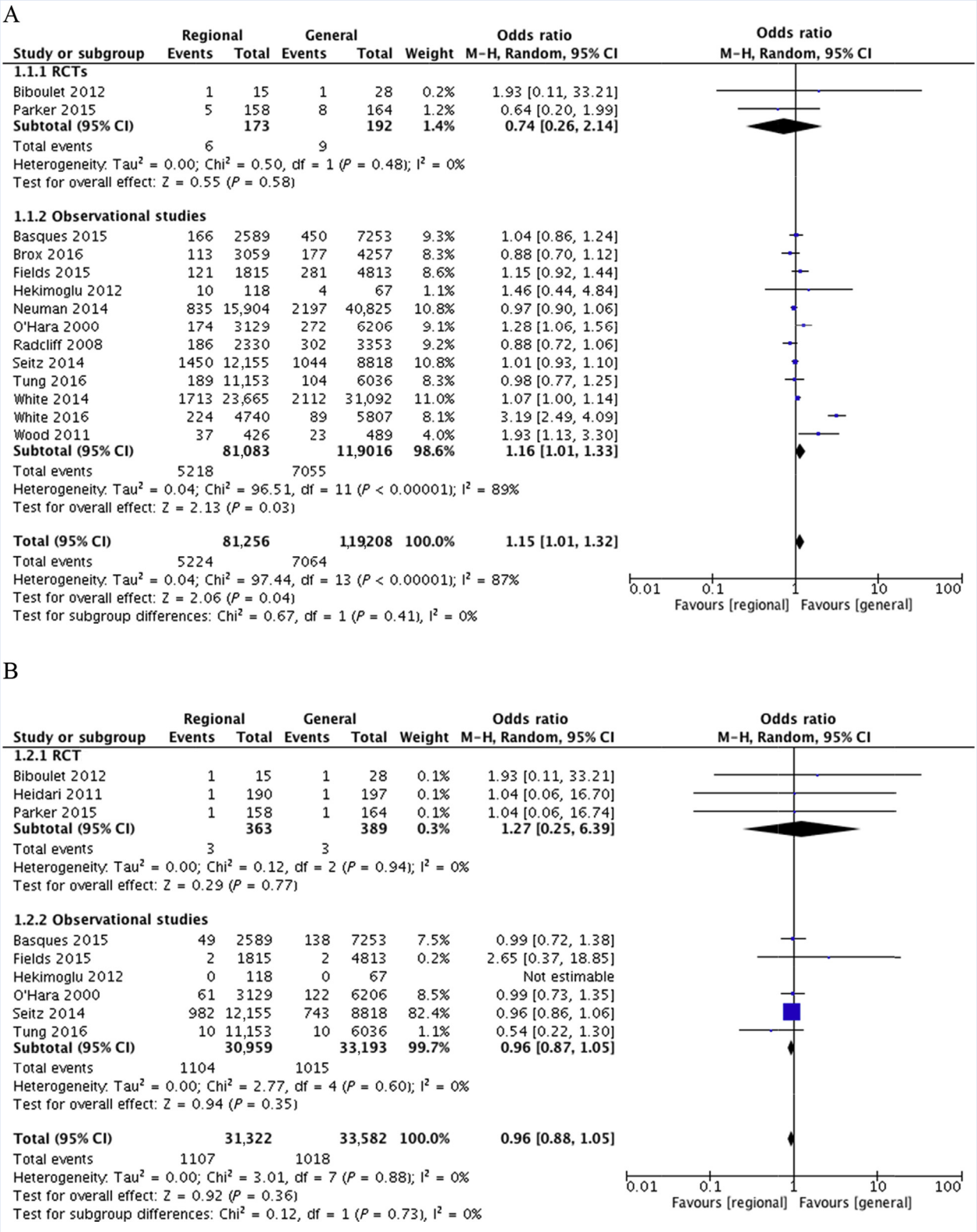
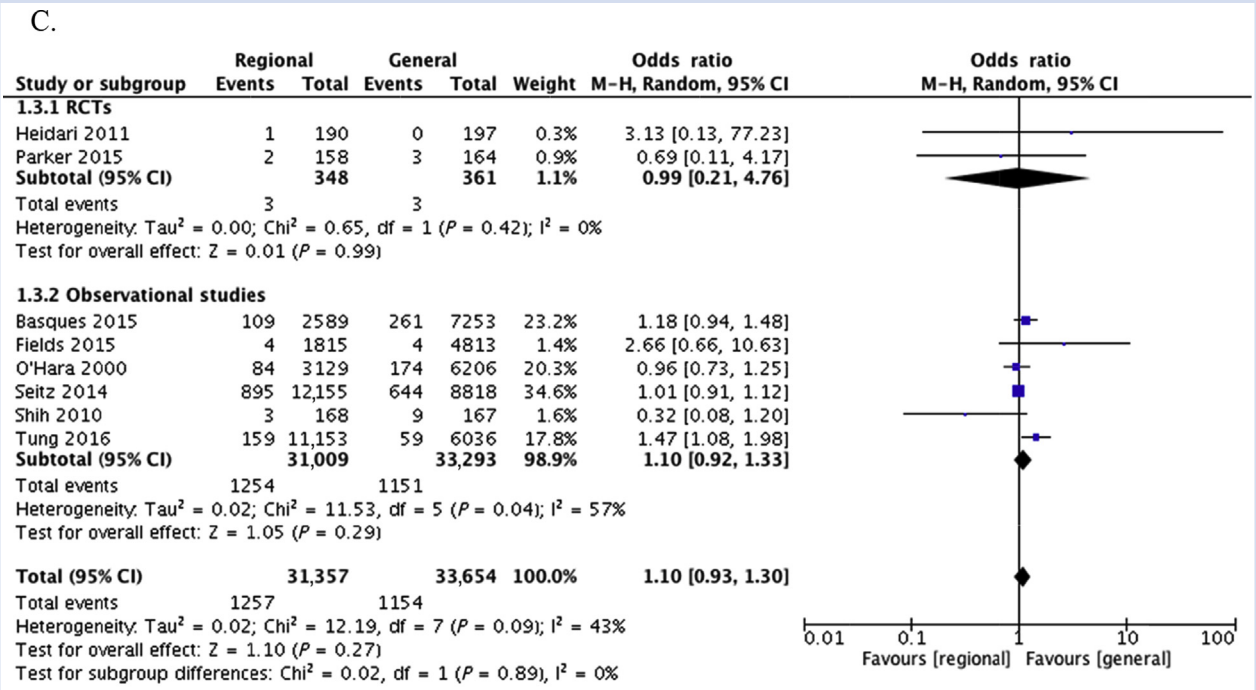


Fig 2. Forest plots showing pooled effect estimates for (a) mortality, (b) acute myocardial infarction, (c) pneumonia, (d) delirium, (e) acute renal failure, and (f) hospital length of stay when comparing regional with general anaesthesia.

C.



D.

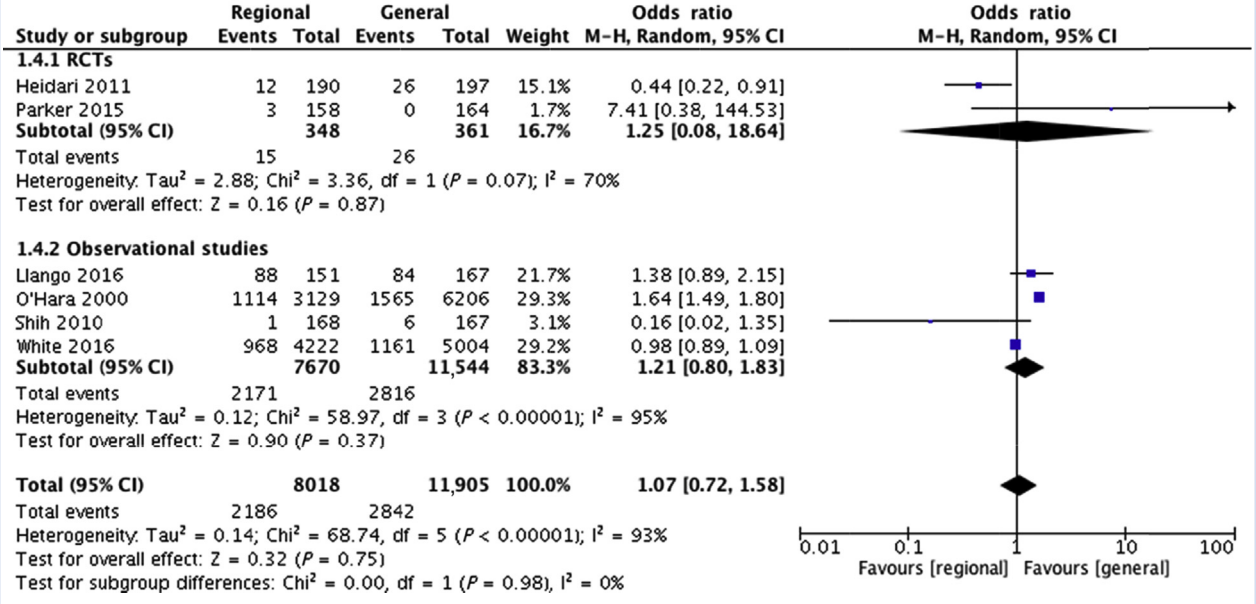


Fig 2. (*continued*).

Sensitivity analysis: outcomes for spinal only vs general anaesthesia

Two randomized[25,27] and seven observational studies[3,4,28–30,35,36] showed no significant difference for 30-day mortality (OR 1.34; 95% CI 1.07, 1.69; I^2 90%; n=96 217). For acute myocardial infarction, two randomized[25,27] and four observational studies[4,28–30] showed no significant difference between general and spinal anaesthesia (OR 0.96; 95% CI 0.87,1.06; I^2 0%; n=37 942). One randomized study[27] and four observational studies[4,28,29,34] showed no significant difference for prevalence of pneumonia (OR 1.06; CI 95% 0.85, 1.31; I^2 38%;

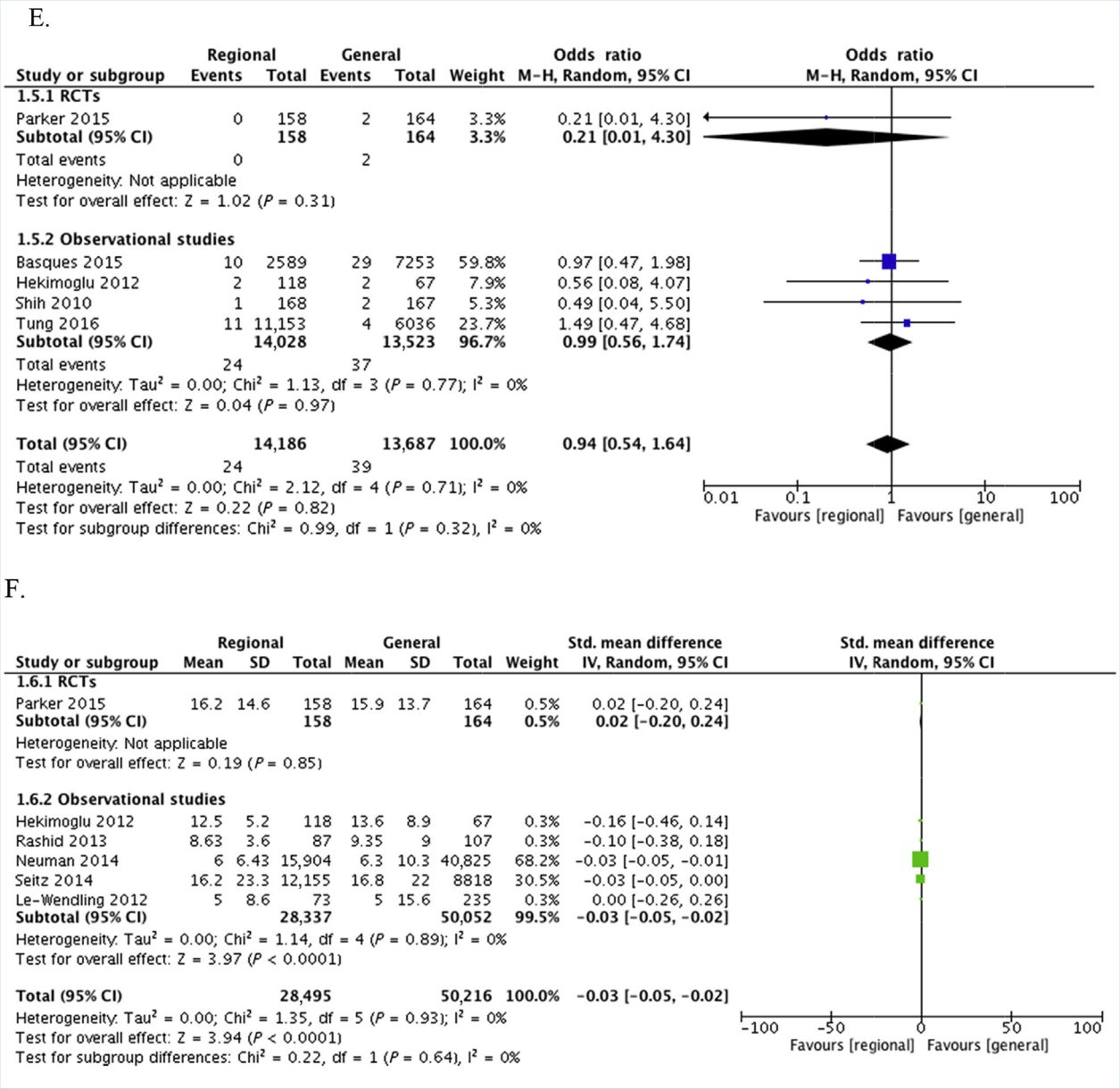


Fig 2. (*continued*).

n=38 100). One randomized study[27] and three observational studies[4,30,31] showed a statistically significant difference for hospital length of stay (standardized mean difference −0.01; 95% CI −0.14, 0.11; I^2 0%; n=1086) but once again, this is so small that it is likely to be clinically insignificant. For delirium, one randomized study[27] and two observational studies[34,41] showed no significant difference between general and spinal anaesthesia (OR 1.04; 95% CI 0.2, 5.33; I^2 61%; n=975). One randomized study[27] and two observational studies[28,30] showed no significant difference between the two modes of anaesthesia for acute renal failure (OR 0.90; 95% CI 0.47, 1.75; I^2 0%; n=10,298; Appendix 3).

Qualitative reporting of outcomes

Perioperative outcomes were reported in a non-standardized way across all studies included. The median number of outcomes per study was eight (range 1–21) as shown in Supplementary Table. Outcome measures tended to reflect the following domains; mortality, outcomes per organ system, and process of care. The most commonly reported outcome measure was 30-day mortality, which was reported in 18 (56%) of the studies.[3,4,23–25,27–30,32,33,35–40,42] The second most common outcome measure was myocardial infarction, reported in 14 (44%) of studies,[4,24–30,33,38,40,42–44] followed by pneumonia in 13

(41%) studies[4,23,24,26–29,33,34,38,40,42,44] and surgical site infection in 10 (31%) studies.[4,24,27–29,37,38,40,42,44] Outcome measures that were reported only once across the studies include myocardial ischaemia, hypoxia, reintubation, transient ischaemic attack, time from admission to medically fit for discharge, and mortality at 120 days (Supplementary Table). For most studies, outcomes were not defined as shown in Table 2.

Mortality

Mortality rates were reported in nine different ways throughout the 32 studies. After 30-day mortality, in-hospital mortality was the next most commonly reported, being used in 19% of the studies (n=6).[21,26,31,43–45] Four studies[27,39,41,44] reported 1-year mortality and two reported 'overall' mortality in an ill-defined manner.[34,46] Two studies reported mortality occurring in <5 days,[3,35] two reported mortality at 90 days[27,39] and a different two studies reported 7-day mortality.[30,33] Eight of the 32 studies reported mortality at more than one timepoint,[3,26,27,30,33,35,39,44] with Parker and colleagues reporting mortality at 30, 90, and 120 days, as well as at 1 year.[27]

Process of care

Other perioperative outcomes reported were length of stay, duration of surgical procedure and anaesthesia, and time from admission to the operating room and from admission until medically fit. They were reported with various frequencies within the studies.

Length of stay was reported in 11 (34%) studies,[23,27,30–32,34,36,37,41,44] with only two providing information on how this was defined and what time points were used.[3,27] White and colleagues[3] recorded the admission and surgery times, as well as 'time from surgery to discharge from the acute hospital'. Parker and colleagues[27] defined length of stay as the 'total hospital stay in days'. However, there was no specific information on whether this was the acute hospital stay or if movement to other care units was included. Time from operation until discharge was reported in six (19%) studies.[3,28,36–38,47] Basques and colleagues[28] reported postoperative length of stay as the number of days from operation until discharge, but provided no information regarding the total length of stay. Similarly, Casati and colleagues[47] recorded 'delay until discharge time' and Wood and White[36] clarified that 'hospital length of stay' was 'postoperative length of stay' only. One of the issues in comparing data for length of stay is the lack of definition and uncertainty that any two papers are comparable. Given that Wood and White[36] stated length of stay was actually a 'postoperative length of stay', they were excluded from the meta-analysis. With no definition, there is an assumption that 'hospital length of stay' includes pre- and postoperative time and thus papers which have described 'hospital length of stay' without elaboration have been included in quantitative analyses.

The less commonly reported outcome measures were time from admission to operation, reported in three studies,[24,26,31] time from admission until medically fit,[44] fasting times,[24] and recovery ward length of stay.[47] Hospital readmissions[28,31,38,40] and intensive care unit admissions[4,21,23,43] were reported in four (13%) and four (13%) studies respectively.

Table 2 Number of studies providing definitions for commonly reported outcomes.

Outcome	Studies reporting, n	Studies providing outcome definition n (%)
Acute myocardial infarction	14	1 (7)
Pneumonia	13	2 (15)
Delirium	9	3 (33)
Acute renal failure	7	1 (14)
Septic shock	6	1 (17)
Hospital length of stay	11	2 (18)

Perioperative morbidity

Supplementary Table illustrates the various perioperative complications, categorized by organ system, with a specific section for critical care interventions. There was no systematic approach to reporting these. Issues relating to analgesia, allergies, and gastrointestinal complications, which were reported less commonly, were categorized as 'other'.

All studies reported some form of cardiovascular complication, with myocardial infarction reported in 14 studies (44%).[4,24–30,33,38,40,42–44] Only one of these studies described how acute myocardial infarction was defined.[33] Eleven studies (34%) included more than one cardiovascular complication.[4,24–30,33,38,44] Heidari and colleagues[26] reported myocardial infarction and ischemia as separate entities with no definition as to how they were distinguished. The majority of studies did not mention any criteria upon which the diagnosis of acute myocardial infarction was made. The International Classification of Diseases codes were used in one study to identify those with acute myocardial infarctions; however, these do not provide definitions of what constituted the diagnosis.[43] Other methods described included a postoperative complication being reported if a consultant felt it was present,[26] or retrospective analysis of medical notes.[33] Both of these are subjective and prone to detection bias. Three (9%) studies reported 'any cardiac complication' with no further detail.[21,31,34] Hypotension was reported in seven (22%) studies[20,25,27,29,36,48,49] in various ways, including absolute and relative reductions in blood pressure and therefore it was not possible to undertake a meta-analysis for this outcome.

Pneumonia was the most commonly reported outcome classified within pulmonary complications, followed by venous thromboembolic events. Despite 13 studies reporting pneumonia as an outcome,[4,23,24,26–29,33,34,38,40,42,44] only two defined how the diagnosis was made. O'Hara and colleagues[33] confirmed a diagnosis of pneumonia if the patient had a chest radiograph with infiltrates and was commenced on antimicrobials. In contrast, Fields and colleagues[29] used the American College of Surgeons National Surgical Quality Improvement Program definition for pneumonia.[50] Venous thromboembolic event reporting included pulmonary embolism and deep vein thrombosis together in eight (25%) of the studies.[4,24,27,29,37,38,40,42] Pulmonary embolism was reported separately in seven studies[4,24,27,28,38,42,44]; however, the categorization of deep vein thrombosis and pulmonary embolism together in some studies, while others separated them, reduced the number of studies that could be included in meta-analyses. Respiratory failure was reported in five (16%) studies[23,30,34,40,43] and other pulmonary outcomes that were reported sporadically included pleural effusion, atelectasis, and exacerbation of chronic obstructive pulmonary disease.

Delirium[22,24,26,27,33,34,41,44,47] and cerebrovascular accident[24,25,27–29,34,40,42,43] were the two most commonly reported neurological outcomes, each reported in nine (28%) studies. Only 33% of studies that reported delirium clearly described the methods used to diagnose it. In each case, these methods varied by their use of different delirium assessment tools.[24,41,44] 'Cognitive issues' were reported in some studies; however, there was limited detail available to translate this into actual diagnoses.

Surgical site and urinary tract were the most commonly reported infections. Their rates were reported in 10 (31%)[23,24,27–29,37,38,40,42,44] and eight (25%) studies,[24,28,29,37,38,40,42,44] respectively. Septic shock was reported in six (19%) studies.[28–30,38,40,42] Wound dehiscence and graft failure was reported in four (13%)[28,29,38,42] and three studies,[28,29,38] respectively. The studies did not clearly define graft failure.

Seven (22%) studies[21,24,27–29,38,44] reported blood transfusion, five (16%) reported blood loss[20,24,34,37,49] and two studies reported a reduction in haemoglobin[26] and anaemia,[24] respectively. The different ways of reporting a similar outcome make meaningful pooling of these results impossible.

Similarly, renal outcomes were most commonly reported under the heading 'renal insufficiency' and 'renal complication', with no use of validated diagnostic criteria on how they had been defined. Seven (22%) studies reported acute renal failure[24,27–30,40,43] as an outcome. No study used the term *acute kidney injury*, which is formally diagnosed based on specific criteria.

Discussion

This study showed no significant differences in 30-day mortality or prevalence of myocardial infarction, pneumonia, delirium, or renal impairment in those patients with hip fracture undergoing surgery where either regional or general anaesthetic was used. Sensitivity analyses comparing spinal anaesthetic, as a sole technique with general anaesthesia, for the same outcomes, were also insignificant. There was a marginal statistical significance favouring regional and spinal anaesthesia when compared to general anaesthesia for hospital length of stay; however, this is not clinically significant as it results in a difference of only a small portion of an actual bed day.

This review is novel in including relevant outcomes from observational studies, increasing the number of patients from which outcomes were included from 2000–3000[51] to 202 000. Previous reviews have used data from mostly randomized studies, many of which dated back to the 1980s and had questionable relevance to modern anaesthetic practice.[51] The pooled, randomized data to date are inadequate in terms of power displayed by the wide 95% CIs displayed within the meta-analyses (Fig. 2). Powering a randomized study to detect a difference in 30-day mortality at the minimal acceptable power standard of 80% would require 8200 patients per group based on our calculations from Fig. 2a.

A recently published systematic review by Van Waesberghe and colleagues[52] also included observational studies but there were several methodological weaknesses such as inclusion of data from patients having elective total hip arthroplasties, comparing adjusted data with raw data as well as comparing means with medians for length of stay.

A limitation of including observational data is the possibility that their design could introduce bias.[53] The Cochrane Risk of Bias tools are well validated and this review reassuringly demonstrates a similar degree of heterogeneity and risk of bias between the observational and randomized studies (Appendices 1 and 2). A further limitation was the inability to include studies in the meta-analyses that provided adjusted data only. Adjusted estimates of effect, via various statistical approaches, aim to control for confounders and are non-comparable with crude estimates of effect. An example is a recently published study by Chu and colleagues,[43] which showed higher odds of in-hospital mortality in the general anaesthesia group when adjusted using propensity score–matching, (2.62 *vs* 2.13%; OR, 1.24; 95% CI, 1.15 to 1.35; $P<0.001$). Basques and colleagues[28] found a significantly shorter length of stay $P<0.001$ in the general anaesthesia group in a retrospective study of 9842 patients using bivariate analysis.

This is the first article to provide data showing substantial variation in the definitions of perioperative hip fracture outcomes across a wide number of studies. This is likely to be relevant to wider research areas. In the randomized and observational studies considered in this review, there was a large and equal degree of inconsistency in defining specific outcomes, if at all, and how they were reported. This left it possible to pool only limited numbers of outcomes in meta-analyses. For example, 'cardiac complications' were often reported under a generic umbrella term with no breakdown of the various components required to meet internationally agreed criteria.[26] Similarly, there was large heterogeneity in the definition of the study intervention. Regional anaesthesia encompassed a spectrum of interventions from neuraxial blockade to spinal and/or epidural anaesthesia, with or without a nerve blockade. Regional anaesthesia with sedation was documented as general anaesthesia[41] in one paper. However, deep sedation, resembling a general anaesthetic, may be simply categorized as regional anaesthesia. Only one study provided data on the number of patients having sedation while under spinal anaesthesia.[3] The patients within these studies receiving general anaesthesia commonly had a nerve block performed, which is true to practice,[23,28,36] thus for meta-analyses, the general anaesthesia group encompasses 'general only' and 'general plus nerve block'. Two studies[3,35] provided a breakdown of data for the categories spinal only or spinal plus nerve block and one study provided results for a group that received both general and regional anaesthesia.[39] Data from patients with such a combined technique has been handled differently in studies either by excluding it completely,[29,33] analysing it within the general anaesthesia group[41] or not describing how it was differentiated at all.

This review highlights that, although meta-analyses may be performed for a specific outcome in terms of how it is reported within studies, the investigators may not have captured the diagnoses in a standardized way. The lack of clarity in the definition of delirium[54,55] may account for the wide prevalence range for this outcome, varying from 0.9%[27] to 54%[41] in papers included within this review. Furthermore, a recent study by Numan *et al*[56] showed no agreement between experts when diagnosing delirium using well validated cognitive assessment tools in the same patient. Similarly, length of stay is one of the most widely reported outcomes, yet the definitions used for this across the studies has been inconsistent. It is therefore difficult to be certain that 'hospital length of stay' is actually measuring the same duration of time

across various healthcare systems. There is potential for many factors to affect a patient's journey from acute care phase to rehabilitation as well as ill-defined criteria guiding decisions on when to step down levels of care. In terms of getting definitions that are at least consistent, the practice in Wales and now England to report the combined acute and community care to give overall length of stay may be worthwhile.[1] This in now commonly referred to as the 'super-spell'.

The common domains that are addressed in this review give rise to the likelihood of obtaining agreement on outcome definitions and a core outcome set for future randomized studies examining the perioperative management of patients with hip fractures.[57] The search was restricted to a short but recent period of time, but we are confident that we have presented sufficient data to demonstrate variability in outcome reporting.

This review has shown that there is no difference in outcomes between patients who receive regional and general anaesthesia undergoing surgery for fractured neck of femur. The focus of future studies must shift from reporting ill-defined outcomes to creating research protocols that include well defined interventions and outcomes of importance to patients. The Patient Centred Outcome Measures after Major Surgery (P-COMMaS)[58] and the Regional versus General Anesthesia for Promoting Independence after Hip Fracture (REGAIN) study[59] in the USA are good examples of ongoing research where some of these principles have been considered.

In conclusion, randomized studies to date in the perioperative management of patients with hip fractures have many methodological flaws. Including data from large observational studies has added weight to the outcomes of the randomized studies but at present, due to enormous inconsistency in the choice of outcome measures and definition, it is difficult to draw many meaningful conclusions from these. Expert consensus on and implementation of standard outcome sets[10,11] is a fundamental requirement for comparing effects across studies and should now be a priority for all investigators.

Authors' contributions

Study design/data collection and analysing/writing paper: all authors.
Revising paper: all authors.

Acknowledgements

We are grateful to Mr Richard Fallis for the assistance in the literature search. Librarian, Queen's University Belfast.

Declaration of interest

None declared.

Funding

Department of Anaesthesia, Royal Victoria Hospital, Belfast, UK.

Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.bja.2017.09.002.

References

1. National Hip Fracture Database. *Annual report*. 2016. Available from: https://www.boa.ac.uk/wp-content/uploads/2016/11/NHFD-annual-report-2016_0.pdf. [Accessed 10 July 2017]
2. NHS National Services Scotland. *Hip fracture care pathway report*. Musculoskeletal audit. 2016. Available from: http://www.msk.scot.nhs.uk/documents/SHFA-Report-2016.pdf. [Accessed 12 July 2017]
3. White SM, Moppett IK, Griffiths R, et al. Secondary analysis of outcomes after 11,085 hip fracture operations from the prospective UK Anaesthesia Sprint Audit of Practice (ASAP-2). *Anaesthesia* 2016; **71**: 506–14
4. Seitz D, Gill S, Bell C, et al. Postoperative medical complications associated with anesthesia in older adults with dementia. *J Am Geriatr Soc* 2014; **62**: 2102–9
5. Sahota O, Morgan N, Moran CG. The direct cost of acute hip fracture care in care home residents in the UK. *Osteoporos Int* 2012; **23**: 917–20
6. Jameson S, Reed MR. Payment by results and coding practice in the national health service; the importance for orthopaedic surgeons. *J Bone Jt Surg Br* 2007; **89**: 1427–30
7. Clarke M, Williamson P. Core outcome sets and systematic reviews. *Syst Rev* 2016; **5**: 11
8. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014; **383**: 267–76
9. Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trial. *Br Med J* 2010; **340**, c869
10. Chan AW, Tetzlaff JM, Altman DG, et al. SPIRIT 2013 Statement: defining standard protocol items for clinical trials. *Ann Intern Med* 2013; **158**: 200–7
11. Hsu CC, Sanford BA. The delphi technique: making sense of consensus. PARE 2007; 10: 1–8. Available from: pareonline.net [Accessed 02 December 2015].
12. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009; **151**: 264–9
13. Booth A, Clarke M, Ghersi D, et al. An international registry of systematic-review protocols. *Lancet* 2011; **377**: 108–9
14. O'Donnell C, Clarke M, McAuley D, et al. Perioperative outcomes for patients undergoing hip fracture surgery: protocol for a systematic review. PROSPERO 2015; 33405. Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016033405.
15. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database Syst Rev* 2011; **2011**, MR000026
16. Higgins JPT, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Br Med J* 2011; **343**, D5928
17. Sterne JAC, Hernán MA, Reeves BC, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *Br Med J* 2016; **355**, I4919
18. Goldet G, Howick J. Understanding GRADE: an introduction. *J Evid Based Med* 2013; **6**: 50–4
19. Higgins JPT, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *Br Med J* 2003; **327**: 557–60

20. Olofsson C, Nygards EB, Bjersten AB, et al. Low-dose bupivacaine with sufentanil prevents hypotension after spinal anesthesia for hip repair in elderly patients. *Acta Anaesthesiol Scand* 2004; **48**: 1240–4
21. Sieber FE, Zakriya KJ, Gottschalk A, et al. Sedation depth during spinal anesthesia and the development of postoperative delirium in elderly patients undergoing hip fracture repair. *Mayo Clin Proc* 2010; **85**: 18–26
22. De Visme V, Picard F, Le Jouan R, et al. Combined lumbar and sacral plexus block compared with plain bupivacaine spinal anaesthesia for hip fractures in the elderly. *Reg Anes Pain Med* 2000; **25**: 158–62
23. Seitz DP, Gill SS, Gruneir A, et al. Effects of cholinesterase inhibitors on postoperative outcomes of older adults with dementia undergoing hip fracture surgery. *Am J Geriatr Psychiatry* 2011; **19**: 803–13
24. Bjorkelund KB, Hommel A, Thorngren K, et al. Reducing delirium in elderly patients with hip fracture: a multifactorial intervention study. *Acta Anaesthesiol Scand* 2010; **54**: 678–88
25. Biboulet P, Jourdan A, Van Haevre V, et al. Hemodynamic profile of target-controlled spinal anaesthesia compared with 2 target-controlled general anaesthesia techniques in elderly patients with cardiac comorbidities. *Reg Anes Pain Med* 2012; **37**: 433–40
26. Heidari SM, Soltani H, Hashemi SJ, et al. Comparative study of two anesthesia methods according to postoperative complications and one month mortality rate in the candidates of hip surgery. *J Res Med Sci* 2011; **16**: 323–30
27. Parker MJ, Griffiths R. General versus regional anaesthesia for hip fractures. A pilot randomised controlled trial of 322 patients. *Injury* 2015; **46**: 1562–6
28. Basques BA, Bohl DD, Golinvaux NS, et al. General versus spinal anaesthesia for patients aged 70 years and older with a fracture of the hip. *Bone Jt J* 2015; **97**: 689–95
29. Fields AC, Dieterich JD, Buterbaugh K, et al. Short-term complications in hip fracture surgery using spinal versus general anaesthesia. *Injury* 2015; **46**: 719–23
30. Sevtap HS, Nurettin H, Alkin C, et al. Comparison of different anesthetic techniques on postoperative outcomes in elderly patients with hip fracture. *Turkiye Klinikleri J Med Sci* 2012; **32**: 623–9
31. Le-Wendling L, Bihorac A, Baslanti T, et al. Regional anesthesia as compared to general anesthesia for surgery in geriatric patients with hip fracture: does it decrease morbidity, mortality and healthcare costs? Results of a single-centered study. *Pain Med* 2012; **13**: 948–56
32. Neuman MD, Rosenbaum PR, Ludwig JM, et al. Anesthesia technique, mortality, and length of stay after hip fracture surgery. *JAMA* 2014; **311**: 2508–17
33. O'Hara DA, Duff A, Berlin JA, et al. The effect of anesthetic technique on postoperative outcomes in hip fracture repair. *Anesthesiology* 2000; **92**: 947–57
34. Shih JY, Hsieh CH, Kang TW, et al. General versus spinal anesthesia: which is a risk factor for octogenarian hip fracture repair patients? *Int J Gerontol* 2010; **4**: 37–42
35. White SM, Moppett IK, Griffiths R. Outcome by mode of anaesthesia for hip fracture surgery. An observational audit of 65 535 patients in a national dataset. *Anaesthesia* 2014; **69**: 224–30
36. Wood RJ, White SM. Anaesthesia for 1131 patients undergoing proximal femoral fracture repair: a retrospective, observational study of effects of blood pressure, fluid administration and perioperative anaemia. *Anaesthesia* 2011; **66**: 1017–22
37. Rashid RH, Shah AA, Shakoor A, et al. Hip fracture surgery: does type of anesthesia matter? *Biomed Res Int* 2013; **2013**, 252356
38. Radcliff TA, Henderson WG, Stoner TJ, et al. Patient risk factors, operative care, and outcomes among older community-dwelling male veterans with hip fracture. *J Bone Jt Surg Am* 2008; **90**: 34–42
39. Brox WT, Chan PH, Cafri G, et al. Similar mortality with general or regional anesthesia in elderly hip fracture patients. *Acta Orthopaedica* 2016; **87**: 152–7
40. Tung YC, Hsu YH, Chang GM. The effect of anesthetic type on outcomes of hip fracture surgery: a nationwide population-based study. *Medicine* 2016; **95**, E3296
41. Ilango S, Pulle RC, Bell J, et al. General versus spinal anaesthesia and postoperative delirium in an orthogeriatric population. *Austral J Ageing* 2016; **35**: 42–7
42. Whiting PS, Molina CS, Greenberg SE, et al. Regional anaesthesia for hip fracture surgery is associated with significantly more peri-operative complications compared with general anaesthesia. *Int Orthop* 2015; **39**: 1321–7
43. Chu C, Weng S, Chen K, et al. Propensity score-matched comparison of postoperative adverse outcomes between geriatric patients given a general or a neuraxial anesthetic for hip surgery: a population-based study. *Anesthesiology* 2015; **123**: 136–47
44. Edelstein DM, Aharonoff GB, Karp A, et al. Effect of postoperative delirium on outcome after hip fracture. *Clin Orthop Rel Res* 2004; **422**: 195–200
45. Patorno E, Neuman MD, Schneeweiss S, et al. Comparative safety of anesthetic type for hip fracture surgery in adults: retrospective cohort study. *Br Med J* 2014; **348**: 4022
46. Karaman O, Ozkazanli G, Orak MM, et al. Factors affecting postoperative mortality in patients older than 65 years undergoing surgery for hip fracture. *Ulus Travma Acil Cerrahi Derq* 2015; **21**: 44–50
47. Casati A, Aldegheri G, Vinciguerra E, et al. Randomised comparison between sevoflurane anaesthesia and unilateral spinal anaesthesia in elderly patients undergoing orthopaedic surgery. *Eur J Anaesthesiol* 2003; **20**: 640–6
48. Hoppenstein D, Zohar E, Ramaty E, et al. The effects of general versus spinal anaesthesia on frontal cerebral oxygen saturation in geriatric patients undergoing emergency surgical fixation of the neck of femur. *J Clin Anes* 2005; **17**: 431–8
49. Messina A, Frassanito L, Colombo D, et al. Hemodynamic changes associated with spinal and general anaesthesia for hip fracture surgery in severe ASA III elderly population: a pilot trial. *Minerva Anestesiol* 2013; **79**: 1021–9
50. American College of Surgeons National Surgical Quality Improvement Program. *Participant use data file*. 2013. Available from: https://www.facs.org/~/media/files/quality%20programs/nsqip/ug12.ashx
51. Guay J, Parker MJ, Gajendragadkar PR, et al. Anaesthesia for hip fracture surgery in adults. *Cochrane Database Syst Rev* 2016, CD000521
52. Van Waesberghe J, Stevanovic A, Rossaint R, et al. General vs neuraxial anaesthesia in hip fracture patients: a systematic review and meta-analysis. *BMC Anesthesiology* 2017; **17**: 87

53. Deeks JJ, Dinnes J, D'Amico R, et al. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003; **7**: 1–173
54. Inouye SK, van Dyck CH, Alessi CA, et al. Clarifying confusion: the confusion assessment method. A new method for detection of delirium. *Ann Intern Med* 1990; **113**: 941–8
55. Aw D, Sahota O. Orthogeriatrics moving forward. *Age Ageing* 2014; **43**: 301–5
56. Numan T, Van den Boogard M, Kamper AM, et al. Recognition of delirium in postoperative elderly patients. A multicenter study. *J Am Geriatr Soc* 2017; **65**: 1932–8
57. McAuley D, Shields M, O'Donnell C. Identifying a core outcome set for evaluating perioperative morbidity in the hip fracture population COMET Initiative; 2015. Available from: http://www.comet-initiative.org/studies/details/757 [Accessed 30 October 2017].
58. National Institute of Academic Anaesthesia Health Services Research Centre. Patient-Centred Outcome Measures for Major Surgery (P-COMMaS). Available from: http://www.niaa.org.uk/HSRC-P-COMMaS [Accessed 22 August 2017].
59. Neuman MD, Ellenberg SS, Sieber FE, et al. Regional versus general anesthesia for promoting independence after Hip Fracture (REGAIN): protocol for a pragmatic, international multicentre trial. *BMJ Open* 2016; **6**, e013473

Handling editor: J.G. Hardman