Cochrane Library

Cochrane Database of Systematic Reviews

Minimally invasive parathyroidectomy guided by intraoperative parathyroid hormone monitoring (IOPTH) and preoperative imaging versus bilateral neck exploration for primary hyperparathyroidism in adults (Review)

Ahmadieh H, Kreidieh O, Akl EA, El-Hajj Fuleihan G

Ahmadieh H, Kreidieh O, Akl EA, El-Hajj Fuleihan G.
Minimally invasive parathyroidectomy guided by intraoperative parathyroid hormone monitoring (IOPTH) and preoperative imaging versus bilateral neck exploration for primary hyperparathyroidism in adults.
Cochrane Database of Systematic Reviews 2020, Issue 10. Art. No.: CD010787.
DOI: 10.1002/14651858.CD010787.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Minimally invasive parathyroidectomy guided by intraoperative parathyroid hormone monitoring (IOPTH) and preoperative imaging versus bilateral neck exploration for primary hyperparathyroidism in adults

Hala Ahmadieh[1a], Omar Kreidieh[2b], Elie A Akl[3], Ghada El-Hajj Fuleihan[4]

[1]Department of Internal Medicine, Beirut Arab University, Beirut, Lebanon. [2]Department of Medicine, Division of Cardiovascular Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston MA, Massachusetts, USA. [3]Department of Internal Medicine, American University of Beirut Medical Center, Beirut, Lebanon. [4]Department of Internal Medicine, Calcium Metabolism and Osteoporosis Program, WHO Collaborating Center for Metabolic Bone Disorders, Division of Endocrinology, American University of Beirut Medical Centre, Beirut, Lebanon

[a]These authors contributed equally to this work.. [b]These authors contributed equally to this work.

Contact address: Ghada El-Hajj Fuleihan, gf01@aub.edu.lb.

Editorial group: Cochrane Metabolic and Endocrine Disorders Group.
Publication status and date: New, published in Issue 10, 2020.

Citation: Ahmadieh H, Kreidieh O, Akl EA, El-Hajj Fuleihan G. Minimally invasive parathyroidectomy guided by intraoperative parathyroid hormone monitoring (IOPTH) and preoperative imaging versus bilateral neck exploration for primary hyperparathyroidism in adults. *Cochrane Database of Systematic Reviews* 2020, Issue 10. Art. No.: CD010787. DOI: 10.1002/14651858.CD010787.pub2.

ABSTRACT

Background

Bilateral neck exploration (BNE) is the traditional approach to sporadic primary hyperparathyroidism. With the availability of the preoperative imaging techniques and intraoperative parathyroid hormone assays, minimally invasive parathyroidectomy (MIP) is fast becoming the favoured surgical approach.

Objectives

To assess the effects of minimally invasive parathyroidectomy (MIP) guided by preoperative imaging and intraoperative parathyroid hormone monitoring versus bilateral neck exploration (BNE) for the surgical management of primary hyperparathyroidism.

Search methods

We searched CENTRAL, MEDLINE, WHO ICTRP and ClinicalTrials.gov. The date of the last search of all databases was 21 October 2019. There were no language restrictions applied.

Selection criteria

We included randomised controlled trials comparing MIP to BNE for the treatment of sporadic primary hyperparathyroidism in persons undergoing surgery for the first time.

Data collection and analysis

Two review authors independently screened titles and abstracts for relevance. Two review authors independently screened for inclusion, extracted data and carried out risk of bias assessment. The content expert senior author resolved conflicts. We assessed studies for overall certainty of the evidence using the GRADE instrument. We conducted meta-analyses using a random-effects model and performed statistical analyses according to the guidelines in the latest version of the *Cochrane Handbook for Systematic Reviews of Interventions*.

Main results

We identified five eligible studies, all conducted in European university hospitals. They included 266 adults, 136 participants were randomised to MIP and 130 participants to BNE. Data were available for all participants post-surgery up to one year, with the exception of missing data for two participants in the MIP group and for one participant in the BNE group at one year. Nine participants in the MIP group and 11 participants in the BNE group had missing data at five years. No study had a low risk of bias in all risk of bias domains.

The risk ratio (RR) for success rate (eucalcaemia) at six months in the MIP group compared to the BNE group was 0.98 (95% confidence interval (CI) 0.94 to 1.03; P = 0.43; 5 studies, 266 participants; very low-certainty evidence). A total of 132/136 (97.1%) participants in the MIP group compared with 129/130 (99.2%) participants in the BNE group were judged as operative success. At five years, the RR was 0.94 (95% CI 0.83 to 1.08; P = 0.38; 1 study, 77 participants; very low-certainty evidence). A total of 34/38 (89.5%) participants in the MIP group compared with 37/39 (94.9%) participants in the BNE group were judged as operative success.

The RR for the total incidence of perioperative adverse events was 0.50, in favour of MIP (95% CI 0.33 to 0.76; P = 0.001; 5 studies, 236 participants; low-certainty evidence). Perioperative adverse events occurred in 23/136 (16.9%) participants in the MIP group compared with 44/130 (33.9%) participants in the BNE group. The 95% prediction interval ranged between 0.25 and 0.99. These adverse events included symptomatic hypocalcaemia, vocal cord palsy, bleeding, fever and infection. Fifteen of 104 (14.4%) participants experienced symptomatic hypocalcaemia in the MIP group compared with 26/98 (26.5%) participants in the BNE group. The RR for this event comparing MIP with BNE at two days was 0.54 (95% CI 0.32 to 0.92; P = 0.02; 4 studies, 202 participants). Statistical significance was lost in sensitivity analyses, with a 95% prediction interval ranging between 0.17 and 1.74. Five out of 133 (3.8%) participants in the MIP group experienced vocal cord paralysis compared with 2/128 (1.6%) participants in the BNE group. The RR for this event was 1.87 (95% CI 0.47 to 7.51; P = 0.38; 5 studies, 261 participants). The 95% prediction interval ranged between 0.20 and 17.87.

The effect on all-cause mortality was not explicitly reported and could not be adequately assessed (very low-certainty evidence). There was no clear difference for health-related quality of life between the treatment groups in two studies, but studies did not report numerical data (very low-certainty evidence). There was a possible treatment benefit for MIP compared to BNE in terms of cosmetic satisfaction (very low-certainty evidence).

The mean difference (MD) for duration of surgery comparing BNE with MIP was in favour of the MIP group (–18 minutes, 95% CI –31 to –6; P = 0.004; 3 studies, 171 participants; very low-certainty evidence). The 95% prediction interval ranged between –162 minutes and 126 minutes. The studies did not report length of hospital stay.

Four studies reported intraoperative conversion rate from MIP to open procedure information. Out of 115 included participants, there were 24 incidences of conversion, amounting to a conversion rate of 20.8%.

Authors' conclusions

The success rates of MIP and BNE at six months were comparable. There were similar results at five years, but these were only based on one study. The incidence of perioperative symptomatic hypocalcaemia was lower in the MIP compared to the BNE group, whereas the incidence of vocal cord paralysis tended to be higher. Our systematic review did not provide clear evidence for the superiority of MIP over BNE. However, it was limited by low-certainty to very low-certainty evidence.

PLAIN LANGUAGE SUMMARY

Minimally invasive parathyroidectomy versus bilateral neck exploration for primary hyperparathyroidism in adults

Review question

Is minimally invasive parathyroidectomy a better surgical treatment compared to classic bilateral neck exploration for people with sporadic primary hyperparathyroidism?

Background

Primary hyperparathyroidism is a condition where one or more of the four parathyroid glands (pea-sized glands located behind or in the thyroid gland in the neck) may enlarge and produce excess parathyroid hormone, a hormone that normally controls calcium and bone metabolism. Excess production of parathyroid hormone results in high blood calcium levels as calcium is drawn out of bones, resulting in increased risk of osteoporosis (weakened bones) and kidney stones. The word 'primary' means that this disorder originates in parathyroid glands and is mostly due to a benign excessive growth of parathyroid cells. Most but not all people with primary hyperparathyroidism have no symptoms. Surgery to remove the diseased parathyroid gland(s) (called parathyroidectomy) is the first-line therapy for people who develop symptoms, namely fractures and kidney stones. Minimally invasive parathyroidectomy is a shorter simpler procedure that uses scans to identify the diseased glands with potentially lower complication risk than bilateral neck exploration (where both sided of the neck are explored to identify which of the four glands are diseased).

Study characteristics

We identified five randomised controlled trials (clinical studies in which people are randomly assigned to one of two or more treatment groups) enrolling a total of 266 adults with primary hyperparathyroidism, who were assigned to one of two surgical techniques (136 participants to the minimally invasive parathyroidectomy group and 130 to the bilateral neck exploration group). One of the studies followed up participants up to five years, but the rest reported data until one year.

Key results

Within six months, operative success as measured by normal blood calcium levels after operation, was found in 97% of participants in the minimally invasive parathyroidectomy group compared with 99% in the bilateral neck exploration group. Five years after the surgery the proportions were 90% in the minimally invasive parathyroidectomy group compared to 95% in the bilateral neck exploration group. About 17% of participants in the minimally invasive parathyroidectomy group reported unwanted events around the time of the operation compared with 34% in the bilateral neck exploration group. These events consisted mostly of symptoms of low calcium levels (such as numbness, tingling cramps) occurring in 14% of the minimally invasive parathyroidectomy group and in 27% in the bilateral neck exploration group. A total of 5/133 (4%) participants in the minimally invasive parathyroidectomy group experienced vocal cord paralysis compared with 2/128 (2%) participants in the bilateral neck exploration group. Other events included bleeding, fever and infection, which were comparable in both groups. The effect on death from any cause was not explicitly reported. There were no clear differences for health-related quality of life between the treatment groups in two studies. There was a possible treatment benefit for minimally invasive parathyroidectomy compared to bilateral neck exploration in terms of cosmetic satisfaction. The duration of surgery was 18 minutes less for the minimally invasive parathyroidectomy technique compared with bilateral neck exploration. Four studies reported a switch from minimally invasive parathyroidectomy to bilateral neck exploration during the operation where 24/115 (21%) participants underwent the more extensive surgery.

Quality of the evidence

The quality of the evidence was low or very low mainly because of the small number of studies and participants.

This evidence is current to 21 October 2019.

SUMMARY OF FINDINGS

Summary of findings 1. Minimally invasive parathyroidectomy versus bilateral neck exploration for primary hyperparathyroidism in adults

Minimally invasive parathyroidectomy guided by intraoperative parathyroid hormone monitoring and preoperative imaging compared to bilateral neck exploration for primary hyperparathyroidism in adults

Participant: adults with primary hyperparathyroidism

Setting: hospitals

Intervention: minimally invasive parathyroidectomy

Comparison: bilateral neck exploration

Outcomes	Illustrative comparative risks (95% CI)		Relative effect (95% CI)	No of participants (studies)	Certainty of the evidence (GRADE)	Comments
	Risk with bilateral neck exploration	**Risk with** minimally invasive parathyroidectomy				
Success rate (eucalcaemia) (a) Follow-up: up to 6 months (b) Follow-up: up to 5 years	(a) **992 per 1000** (b) **949 per 1000**	(a) **972 per 1000** (933 to 1022) (b) **892 per 1000** (787 to 1025)	(a) **RR 0.98** (0.94 to 1.03) (b) **RR 0.94** (0.83 to 1.08)	(a) 266 (5) (b) 77 (1)	(a) / (b) ⊕⊖⊖⊖ **Very low** [a]	(b) 5-year follow-up: data available for 77/91 randomised participants.
Total incidence of perioperative adverse events (number) Follow-up: up to 48 hours postoperatively	**338 per 1000**	**169 per 1000** (112 to 257)	**RR 0.50** (0.33 to 0.76)	236 (5)	⊕⊕⊖⊖ **Low** [b]	—
All-cause mortality Follow-up: up to 5 years	See comment				⊕⊖⊖⊖ **Very low** [c]	No study explicitly reported on the occurrence of perioperative mortality. In 1 study with 5 years' follow-up there were 16 deaths, no data per intervention group were reported.
Health-related quality of life Follow-up: up to 6 months	See comment				⊕⊖⊖⊖ **Very low** [d]	2 studies reported that there were no clear differences between intervention

					groups; however, data were not presented.
Cosmetic satisfaction Follow-up: up to 1 year	See comment			⊕⊖⊖⊖ **Very low** [e]	3 studies reported some data associated with cosmetic satisfaction but measurements varied substantially; overall some findings indicated a benefit of minimally invasive parathyroidectomy.
Duration of surgery (time from skin incision to skin closure)	The mean duration of surgery ranged across control groups from 64 minutes to 82 minutes	The mean duration of surgery in the intervention groups was **18 minutes lower** (31 minutes lower to 6 minutes lower)	—	171 (3)	⊕⊖⊖⊖ **Very low** [f]	The 95% prediction interval ranged between –162 minutes and 126 minutes.
Length of hospital stay	Not reported				

CI: confidence interval; **RR:** risk ratio.

GRADE Working Group grades of evidence

High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate; the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited; the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate; the true effect is likely to be substantially different from the estimate of effect.

[a]Downgraded one level because of risk of bias in several risk of bias domains, one level because of inconsistency and indirectness (95% prediction interval ranging between 0.90 and 1.06 and surrogate outcome) and one level because of imprecision (low sample size, 95% CI consistent with benefit and harm) – see Appendix 13.
[b]Downgraded one level because of risk of bias in several risk of bias domains and one level because of imprecision (low sample size) – see Appendix 13.
[c]Downgraded one level because no study explicitly reported on this outcome and two levels because of serious imprecision (low sample size, low number of studies) – see Appendix 13.
[d]Downgraded one level because only two studies evaluated this outcome but did not report data and two levels because of serious imprecision (low sample size, low number of studies) – see Appendix 13.
[e]Downgraded one level because of risk of bias in several risk of bias domains, one level because of indirectness (surrogate outcome) and one level because of imprecision (low sample size, low number of studies) – see Appendix 13.
[f]Downgraded one level because of risk of bias in several risk of bias domains, one level because of inconsistency (95% prediction interval ranging between –162 minutes and 126 minutes) and one level because of imprecision (low sample size, low number of studies) – see Appendix 13.

BACKGROUND

Description of the condition

Primary hyperparathyroidism (PHPT) is a common disorder of bone metabolism and hypercalcaemia, occurring in 28 per 100,000 individuals yearly in the US. In the Mayo Clinic series, the incidence rose to 82 per 100,000 in the period 1974 to 1982, was down to 29.1 per 100,000 throughout the years 1983 to 1992, and 21.6 per 100,000 in 1992 to 2001 (Wermers 2006). It can occur at any age, with an increase in incidence after age 45 years. It is more common in women than men, and can occur in 2% of the postmenopausal women. The most common aetiology for the disease is a single gland adenoma in 80% to 85% of cases, with the rest of the cases constituting multiple gland hyperplasia (10% to 15%), double adenomas (2% to 5%), and carcinoma (1%) (Kaplan 1992; Kunstman 2012; Salti 1992).

The molecular basis for sporadic hyperparathyroidism remains largely unknown. Described abnormalities include gain-of-function in genes that stimulate parathyroid gland growth for sporadic tumours, such as Cyclin1, (Costa-Guda 2014) or loss of function mutations in genes that suppress tumour growth, such as Multiple Endocrine Neoplasia 1 (*MEN 1*) or *HRPT2* in sporadic and familial tumours (Westin 2009).

Since the development of adequate screening techniques, the disorder has evolved in high-income countries into a mostly asymptomatic disease, often detected as a laboratory abnormality with mild hypercalcaemia incidentally discovered via routine examinations. Conversely, in low- to middle income countries, classical presentations still prevail including bone pain, nephrolithiasis, nephrocalcinosis, bone loss, increased fractures and osteitis fibrosa cystica (Bilezikian 2000; Mishra 2001; Parfitt 1991). Skeletal effects seen through changes in bone mineral density (BMD) and histomorphometric analysis were originally thought to be most prominent at cortical bony sites. However, more recent studies have noted volumetric BMD loss at both sites and deterioration in bone structure at trabecular sites (Chen 2003; Pyram 2011; Stein 2013; Vu 2013). An increased fracture risk has been observed in both traditional cortical bony areas, such as the hip, the distal radius, as well as trabecular bony areas such as the vertebrae (Bilezikian 2014; Khosla 2002), but there are no randomised studies to validate such observations. A newly available US Food and Drug Administration (FDA)-approved dual-energy X-ray absorptiometry (DXA)-derived technology known as trabecular bone score (TBS) was evaluated in 24 postmenopausal women with PHPT and revealed deterioration in trabecular bone structure (Silva 2013; Stein 2013; Vu 2013).

Renal calcification or stones appear to be increased up to four-fold in people with PHPT compared to controls, occurring in up to 20% of people (Bilezikian 2014; Starup-Linde 2012; Suh 2008). Non-classical manifestations of PHPT include neurocognitive changes such as impaired concentration, increased depression and anxiety, decreased non-verbal learning process, difficulties in using direct memory, worse performance on tests of verbal memory, verbal fluency and visual constructive abilities (Babinska 2012; Coker 2005; Joborn 1988; McAllion 1989), and cardiovascular abnormalities such as hypertension, myocardial and vascular calcifications, and left ventricular hypertrophy with changes in endothelial function as well as increased vascular stiffness (Fitzpatrick 2008; Rubin 2005; Silverberg 2014; Walker 2012).

Description of the intervention

Parathyroidectomy is the only curative option in people with PHPT. In general, while surgery for hyperparathyroidism in the setting of chronic kidney disease often involves subtotal parathyroidectomy or total parathyroidectomy with autotransplantation (Welk 1987; Welsh 1984; White 1986). Surgery for PHPT aims to only resect the diseased gland(s) and, therefore, remove the source for excess parathyroid hormone (PTH) production. The goal is to thus decrease the incidence of nephrolithiasis and nephrocalcinosis, improve BMD, decrease fractures and possibly improve health-related quality of life. Even in asymptomatic individuals, there is some evidence that surgery improves BMD (Rubin 2008), and possibly functional capacity and health-related quality of life (Ramakant 2012). Therefore, surgery is indicated for all patients with symptomatic PHPT, and some patients with asymptomatic disease. The updated 2014 guidelines on the management of asymptomatic patients include age less than 50 years, serum calcium concentration of 1.0 mg/dL or more above upper limit of normal, creatinine clearance 60 mL/minute or less, BMD at any site with a DXA-derived T-score of –2.5 or less at spine, hip or forearm, vertebral fracture (documented by x-ray, computerised tomography (CT), magnetic resonance imaging (MRI), or vertebral fracture assessment (VFA)), a creatinine clearance less than 60 mL/minute, a 24-hour urine calcium more than 400 mg/day and an increased stone risk. Stone risk is assessed by biochemical stone risk analysis or the presence of nephrolithiasis or nephrocalcinosis identified by x-ray, ultrasound or CT scan (Bilezikian 2014).

Bilateral neck exploration

Bilateral neck exploration (BNE) is the traditional approach to primary sporadic hyperparathyroidism, and when performed by experienced surgeons results in curative rates of 95% to 98%, and is associated with low complication rates (Allendorf 2007). It remains the mainstay treatment for people with unlocalised pathology, familial or hereditary cases, or concomitant thyroid disease.

Minimally invasive parathyroidectomy

Minimally invasive parathyroidectomy (MIP) has largely replaced BNE (Greene 2009), due to its safety and effectiveness as well as possible lower costs and morbidity (Udelsman 2014). The exploration is done via direct visualisation of all parathyroid glands, and may be performed under local or general anaesthesia, through an open traditional incision, minimally invasive incision or even via a videoscopic approach (Alesina 2010; Allendorf 2007; Lo 1999; Lowney 2000; Udelsman 2014).

We will use the 2002 summary statement on asymptomatic hyperparathyroidism definition for MIP. It is a set of techniques employing preoperative imaging and intraoperative PTH assays (IOPTH) to limit surgical visualisation only to the suspected gland (Bilezikian 2002; Carneiro 2003). There are currently several variations for MIP techniques satisfying these criteria (Udelsman 2011; Udelsman 2014). In general, focused parathyroidectomy aims towards visualisation of just the suspected gland, whereas unilateral exploration visualises the entire side suspected to have a pathology. Exploration is either open or endoscopic. Two common endoscopic techniques are described in the literature, and offer advantages of magnified vision and tactile control (Henry 1999; Gracie 2012; Miccoli 1999). The technique suggested by Henry 1999 uses a more lateral approach avoiding dissection of the strap muscles and allowing possible direct visualisation of

the adenoma but may compromise the cardiorespiratory system through the frequent requirement of carbon dioxide insufflation in order to maintain an adequate working space. In contrast, the technique suggested by Miccoli and colleagues suggests a more medial approach, wherein gas insufflation is only maintained for a few minutes in order to allow for dissection of the strap muscles, after which a working space is maintained simply by using external retraction (Miccoli 1999). This 'gasless approach' promised to avoid emphysema, pneumomediastinum and neck swelling (Henry 1999; Henry 2008; Miccoli 1999; Miccoli 2011). Despite the nuances of slight differences in technique, the fundamental methods of all MIP are the same. Single gland disease is identified and localised by use of preoperative ultrasound, Sestamibi scan (Technetium (99mTc) nuclear medicine imaging), or both. A limited exploration then targets the suspicious side or gland in an attempt to avoid a cumbersome full neck exploration (Irvin 1991; Irvin 1994; Kuntsman 2013; Udelsman 2004). IOPTH monitoring is carried out generally through peripheral venous blood draws, with pre-incision, pregland ligation, and 5, 10 and 20 minutes postgland ligation measurements. Criteria used for evidence of adequate incision, and thus termination of surgery, vary widely. The most accepted, the Miami criterion, considers a decrease in PTH measurement by more than 50% from the highest baseline to the 10-minute postgland ligation value as evidence for adequate gland excision (Barczynski 2009; Carneiro 2003; Irvin 1993). In the event of inadequate decline using this criterion, surgery is then converted to a bilateral conventional technique, possibly due to location of abnormal gland on the opposite side or due to suspicion of multiglandular disease. MIP techniques are usually offered to people with preoperative localisation studies suggestive of single gland disease, in the absence of thyroid pathology, familial or hereditary hyperparathyroidism, or lithium intake. MIP techniques should only be implemented in centres that have sophisticated imaging and intraoperative PTH assays, and only performed by experienced endocrine surgeons (Udelsman 2014).

Adverse effects of the intervention

Apart from a theoretical concern for possibilities of subcutaneous emphysema following gas insufflation in the endoscopic technique described by Henry and colleagues (Henry 1999), both MIP and BNE have adverse events, which are rare for both interventions. These include anaesthesia-related or postoperative complications, or both, such as hypocalcaemia, vocal cord paralysis, haematomas and infections. One retrospective review of 656 parathyroid operations found a 3% complication rate for BNE and a 1.2% complication rate for MIP (Udelsman 2002). Haematomas are potentially life-threatening complications arising in 0.3% of all surgeries for PHPT, and may present with a variety of symptoms including neck pain, respiratory distress, dysphagia and wound drainage (Burkey 2001; Carty 2004). Hypoparathyroidism is commonly transient and presents with decreased calcium concentration in association with either mild symptoms of tingling or numbness, or more severe symptoms such as profound fatigue or carpopedal spasm. Other major complications, such as fever and severe hypocalcaemia are rare, unless after subtotal parathyroidectomy (Carty 2004). Such presentations are only transient and permanent hypoparathyroidism occurred in only 0.3% of 380 operations reviewed by Carty colleagues (Carty 2002). Similarly, permanent recurrent laryngeal nerve injury is very rare. In some reports, 0.2% of 1112 participants, and 0.7% of 401 bilaterally explored participants had permanent recurrent laryngeal nerve injury (Allendorf 2007; Udelsman 2002), and 0.3% of 380 participants and 0.4% of 255 participants receiving a minimally invasive surgery had similar injury (Carty 2002; Udelsman 2002). Differences in the incidence of adverse effects between MIP and BNE are controversial and the few randomised controlled trials (RCT) have differing findings. Apart from scar length differences, Slepavicius 2008 reported no significant differences in adverse events between intervention groups, Miccoli 1999 noted insignificant differences in one study and no adverse events in either group in another, while Bergenfelz and colleagues noted a greater incidence of severe hypocalcaemia in the bilateral group compared to MIP group (10% versus 0%) (Bergenfelz 2005; Miccoli 1999; Miccoli 2008; Rulli 2007; Slepavicius 2008).

How the intervention might work

The theoretical basis behind MIP is that most cases are caused by single gland pathology, and that modern tools can identify such pathology with acceptable accuracy, rendering the need for complete visualisation of both sides of the neck almost unnecessary. This may also decrease the rate of complications and thus decrease operative time and cost, although this has been debatable (Bergenfelz 2002; Miccoli 1999). Indeed, more than 80% of spontaneous PHPT is caused by a single solitary adenoma (Kunstman 2012; Ruda 2005). Improvements in imaging techniques have reached sensitivities between 71% to 80% for ultrasound imaging and a sensitivity greater than 90% for Sestamibi scanning (Arici 2001; Carty 1997; Dijkstra 2002; Hindié 2015; Ryan 1997; Ryan 2004). The advent of intraoperative PTH monitoring assays have optimised the accuracy of adequate resection, reaching accuracy greater than 96% for the Miami criterion (Barczynski 2009; Carneiro 2003), and encouraged the adoption of MIP by most surgeons. Indeed, by 2008, 68% of US surgeons were found to be practicing limited exploration techniques and only 10% practiced BNE exclusively (Greene 2009).

Why it is important to do this review

Despite the excellent short-term curative rate and low morbidity profile with both MIP and BNE in PHPT, there is still no consensus on the preferred option to-date (Udelsman 2014). The importance of our review stems from the large number of people undergoing these procedures yearly, without a definitive answer about the long-term success of MIP compared to BNE. The greatest uncertainty in this subject stems from concerns about increased long-term recurrence or missed multiglandular disease in people undergoing MIP. We are aware of only one RCT that had an extended follow-up for five years postoperatively. The investigators noted 4/47 recurrent cases in the MIP group compared to 2/44 in the conventional exploration group (Bergenfelz 2002). One retrospective analysis by a high-volume group found that the long-term failure rate of the unilateral approach was 11 times higher than that of conventional bilateral surgery. However, that group is known for not using IOPTH monitoring (Norman 2012). Another group noted a higher than 8% recurrence rate after eight years of follow-up in people undergoing MIP as opposed to 0% in the open parathyroidectomy group (Schneider 2012). The differences were not significant because of the small number of participants followed up for this duration. Identifying further publications with long-term follow-up may quell concerns, re-ignite them, or highlight the urgent need for long-term outcome studies.

This review will help characterise the respective risks and benefits of each type of surgery more clearly, both in the short term and longer term. Few RCTs have been published on this topic, often with differing conclusions (e.g. see 'Adverse effects of the intervention'). A properly conducted systematic review and meta-analysis helps us provide an objective risk-benefit assessment for both techniques, on prespecified participant outcomes, in both short-term and long-term follow-up. This will serve to empower doctors and participants alike in making better educated choices. Furthermore, we only identified one systematic review comparing both approaches on the topic (Gracie 2012). The Gracie 2012 review had several limitations, and among which it excluded unilateral exploration. We also found the review deficient in key criteria reflecting a systematic methodological quality. On AMSTAR, a validated tool used for assessing methodological quality of systematic reviews (Shea 2007), the previous review scored low, with only two out of a possible 11 tool items scoring positively. Inclusion and exclusion criteria were provided in a table format but the type of studies, languages and publication status (grey literature) were not discussed. The methods of study selection and data abstraction were also not entirely clear, and there was no indication of whether these processes occurred in duplicate. While the search methodology was presented in the paper, the initial results, excluded studies and reasons for exclusions were not. No meta-analysis was done for any of the outcomes of interest, and there was no assessment of risk of bias in the included studies. The review also did not address the greatest source of uncertainty in the topic, that is, long-term success rate. We believe that inclusion of this outcome is extremely important in order to truly gauge the long-term impact of potentially leaving residual enlarged tissue in people who undergo minimally invasive techniques, relying on localisation studies and IOPTH monitoring. Our review aims to address the above limitations in detail.

OBJECTIVES

To assess the effects of minimally invasive parathyroidectomy (MIP) guided by preoperative imaging and intraoperative parathyroid hormone monitoring versus bilateral neck exploration (BNE) for the surgical management of primary hyperparathyroidism.

METHODS

Criteria for considering studies for this review

Types of studies

We included RCTs comparing MIP to BNE for people with sporadic PHPT presenting to surgery for the first time, as specified in our protocol (Kreidieh 2013).

Types of participants

We included studies of adults with PHPT presenting for first-time parathyroidectomy.

We excluded studies of participants:

- undergoing repeat surgeries;
- with secondary hyperparathyroidism;
- with tertiary hyperparathyroidism;
- with parathyroid carcinoma;
- with increased risk of multiglandular disease (i.e. children or people with genetic predispositions to have hyperparathyroidism such as people with multiple endocrine neoplasia);
- with different types of hyperparathyroidism with no separate reporting of results by type;
- with elevated mean creatinine at entry into study.

Diagnostic criteria for considered conditions

Hyperparathyroidism

We used the following definitions of hyperparathyroidism (Bilezikian 2002).

- Elevated or inappropriately normal PTH level, with serum calcium levels above normal reference range.
- Elevated serum PTH with normal calcium levels, after exclusion of secondary causes for elevation of PTH (mainly decreased calcium intake, vitamin D deficiency, renal insufficiency, hypercalciuria of renal origin).

Sporadic primary hyperparathyroidism

- PHPT.
- Exclusion of secondary causes such as renal insufficiency, vitamin D deficiency and familial hyperparathyroidism.

Types of interventions

We planned to investigate the following comparisons of intervention versus comparator.

Intervention

- MIP (open unilateral parathyroidectomy) guided by IOPTH and preoperative imaging.
- MIP (open focused parathyroidectomy) guided by IOPTH and preoperative imaging.
- MIP (endoscopic unilateral parathyroidectomy) guided by IOPTH and preoperative imaging.
- MIP (endoscopic focused parathyroidectomy) guided by IOPTH and preoperative imaging.

Comparator

- BNE regardless of use of and results of any operative adjuncts.

Preoperative imaging

Localisation may be done prior to or after randomisation. We accepted either case, but downgraded the certainty of the evidence for studies with localisation procedures done prior to randomisation due to indirectness. This is because participants receiving BNE in the general population do not routinely receive preoperative imaging.

We accepted imaging results as suggestive of single gland disease, multiple gland disease and inconclusive as determined by an expert radiologist or surgeon, as reported in the paper.

Accepted preoperative localisation procedures included at least one of the following:

- ultrasound imaging using a 5 MHz, 7.5 MHz or 10 MHz transducer;
- technetium 99m-Sestamibi scanning using single isotope dual phase scan, dual isotope subtraction scan and three-

dimensional single-photon emission computerised tomography (SPECT) imaging scan;
- thallium technetium scanning.

Intraoperative parathyroid hormone monitoring

We accepted the use of a second- or third-generation rapid PTH assay intraoperatively for confirmation of adequate gland resection (Eastell 2014), per a commonly accepted criterion, such as the Miami criterion, of a fall in serum PTH of 50% at 10 minutes postgland excision. This fall is from the highest PTH value of either a preskin incision baseline or a pregland excision baseline (Barczynski 2009).

Types of outcome measures

We did not exclude a study if it failed to report one or several of our primary or secondary outcome measures. If the study reported none of our primary or secondary outcomes, we did not include the study but provided some basic information in the Characteristics of studies awaiting classification table.

We investigated the following outcomes using the methods and time points specified below.

Primary outcomes

- Success rate.
- Total incidence of perioperative adverse events.

Secondary outcomes

- Specific adverse events.
- Conversion rate from minimally invasive to open procedure.
- Postoperative increase in PTH with eucalcaemia.
- All-cause mortality.
- Health-related quality of life.
- Cosmetic satisfaction.
- Bone fracture rate.
- Nephrolithiasis rate.
- Absence from work.
- Duration of surgery.
- Length of hospital stay.
- Socioeconomic effects.

Method and timing of outcome measurement

- Success rate: defined by authors as eucalcaemia or hypocalcaemia. We defined short-term success rate as within six months of surgery, medium-term success rate as between six months and five years and long-term success rate as at five years or more after surgery.
- Total incidence of perioperative adverse events: defined by authors as adverse events occurring within 48 hours of surgery.
- Specific adverse events:
  * bleeding events (identified as such in the included study or one requiring transfer to an intensive care unit or requiring blood transfusion within 48 hours of surgery);
  * infection within one month of surgery;
  * hypocalcaemia within 48 hours, one month and six months of surgery (including transient, permanent, severe and mild hypocalcaemia as defined in the adverse events section of our protocol; Kreidieh 2013). We differentiated symptomatic hypocalcaemia, in which participants exhibit typical symptoms of hypocalcaemia, from biochemical hypocalcaemia, referring to any participant with calcium levels below the lower limit of normal for the laboratory in which the measurements were made. Transient hypocalcaemia was defined as hypocalcaemia resolving within six months of surgery, while permanent hypocalcaemia referred to hypocalcaemia persisting longer than six months postoperatively (Mehrabi 2012);
  * postoperative pain using a validated pain score such as the visual analogue scale at 48 hours postoperatively;
  * vocal cord paralysis identified as such in the study in the postoperative period (usually 48 hours);
  * anaesthesia-related complications identified as such in the study occurring intraoperatively.
- Conversion rate from minimally invasive to open procedure: defined as the proportion of participants who were planned to have a MIP but were converted to a BNE intraoperatively.
- Postoperative increase in PTH with eucalcaemia: defined as a within normal serum level of calcium, with an above normal level of PTH at short term (within six months of surgery), medium term (between six months and five years after surgery) and long term (at five years or more after surgery) (Ning 2009).
- Health-related quality of life: measured by a validated instrument such as the medical outcomes study 36-item Short Form Health Survey (SF-36) at one month and six months postoperatively.
- Cosmetic satisfaction: measured using a validated instrument such as a Holander Scale within 48 hours of surgery and at six months postoperatively.
- Bone fracture rate: we considered bone loss as evidenced by a decrease in BMD as a surrogate for fracture rate in case data on fracture rate were not readily available. The minimal important difference in each centre was defined by the centre-specific quality assurance protocol, if defined in the study. If it was not available, we considered as significant any decrease in BMD that exceeded 5% at any skeletal site, considering a precision of up to 2.5%, as recommended by the International Society of Clinical Densitometry (ISCD) (Baim 2008). However, this could have still resulted from random error in centres that did not report centre-specific precision or could not abide by the ISCD quality assurance measures. For the pooled estimate, we considered the minimally important difference as the highest obtained from all studies. Both outcomes were considered within one and five years of surgery.
- Nephrolithiasis rate: defined as percentage of participants having an incidence of nephrolithiasis within five years of surgery.
- Absence from work: defined as the number of days of work missed and determined by study authors to have been caused by the surgery. Timing was not applicable to this item.
- Duration of surgery: defined as time from skin incision to skin closure during surgery. Timing was not applicable to this item.
- Length of hospital stay: defined as the number of days of hospitalisations prior to and following first admission for surgery. Timing was not applicable to this item.
- Socioeconomic effects were not prespecified in the protocol but were detailed in the manuscripts, as provided in the

respective studies. These included direct costs defined as admission/readmission rates, mean length of stay, in-hospital charges, visits to general practitioner, accident/emergency visits; medication consumption; or indirect costs defined as resources lost due to illness by the participant or their family member.

Search methods for identification of studies

Electronic searches

We searched the following sources from inception to 21 October 2019 and placed no restrictions on the language of publication.

- Cochrane Central Register of Controlled Trials (CENTRAL; via Cochrane Register of Studies Online) (searched 21 October 2019).
- Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations and Daily 1946 to October 18, 2019 (searched 21 October 2019).
- ClinicalTrials.gov (searched 21 October 2019).
- World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) Search Portal (apps.who.int/trialsearch/) (searched 21 October 2019).

For detailed search strategies, see Appendix 1.

Searching other resources

We tried to identify other potentially eligible studies or ancillary publications by searching the reference lists of retrieved included studies, systematic or other reviews, meta-analyses and health-technology assessment reports. We contacted study authors of included studies to request missing data and to identify any further studies that we may have missed.

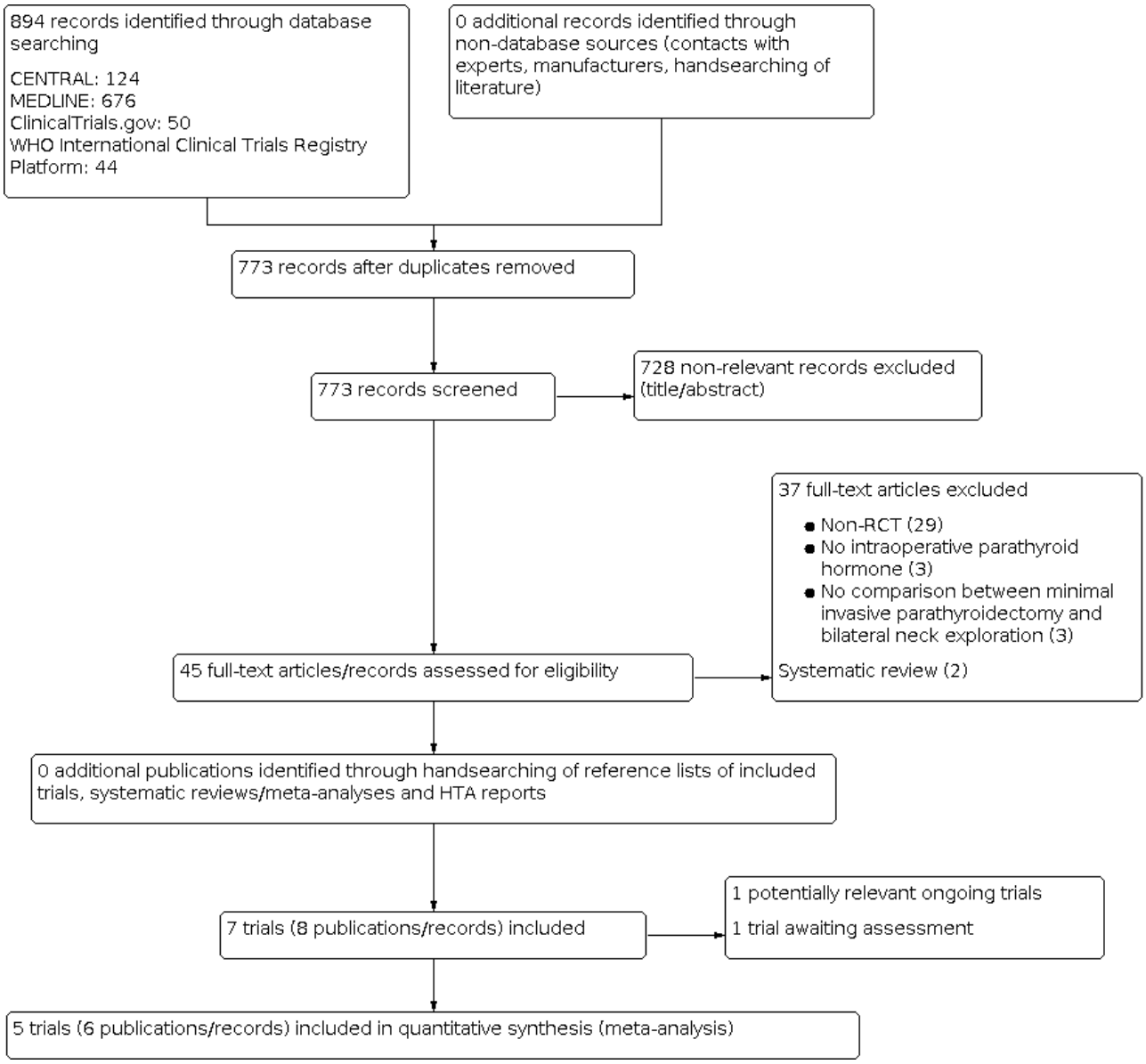
Data collection and analysis

Selection of studies

Two review authors (OK, HA) independently scanned the abstract, title, or both, of every record retrieved in the literature searches, to determine which studies we should assess further. We obtained the full text of all potentially relevant records. We resolved disagreements through consensus or by recourse to a third review author (GEHF and EA). If we could not resolve a disagreement, we categorised the study as a 'Study awaiting classification' and contacted the study authors for clarification. We present an adapted PRISMA flow diagram to shown the process of study selection (Liberati 2009).

We obtained full-text articles from the database searches available at the American University of Beirut. We aimed to translate any studies available in languages other than English, Arabic or French (Figure 1).

Figure 1. Study flow diagram. HTA: health technology assessment.



Data extraction and management

For studies that fulfilled our inclusion criteria, two review authors (OK, HA) independently extracted key participant and intervention characteristics and outcome data. We reported data on outcomes and adverse events using standardised data extraction sheets from the Cochrane Metabolic and Endocrine Disorders (CMED) Group. We resolved disagreements by discussion or, if required, by consultation senior review authors (GEHF and EA) (for details see Table 1; Appendix 2; Appendix 3; Appendix 4; Appendix 5; Appendix 6; Appendix 7; Appendix 8; Appendix 9; Appendix 10; Appendix 11; Appendix 12; Appendix 13; Appendix 14).

We provided information including study identifier for potentially relevant ongoing trials in the Characteristics of ongoing studies table and in Appendix 7 'Matrix of study endpoints (publications and trial documents)'. We tried to identify the protocol for each included study and report in Appendix 7 primary, secondary, and other outcomes in comparison with data in publications.

We sent an email to all authors of included studies to enquire whether they were willing to answer questions regarding their studies. We presented the results of this survey in Appendix 15. We thereafter sought relevant missing information on the study from the primary study author(s), if required.

Dealing with duplicate and companion publications

In the event of duplicate publications, companion documents or multiple reports of a primary study, we maximised the yield of information by collating all available data and used the most complete data set aggregated across all known publications. We listed duplicate publications, companion documents, multiple reports of a primary study and trial documents of included trials (such as trial registry information) as secondary references under

the study ID of the included study. Furthermore, we also listed duplicate publications, companion documents, multiple reports of a study and trial documents of excluded trials (such as trial registry information) as secondary references under the study ID of the excluded study. We attempted to resolve any remaining uncertainties by contacting the authors whenever possible.

Data from clinical trials registers

If data from included studies were available as study results in clinical trials registers, such as ClinicalTrials.gov, or similar sources, we made full use of this information and extracted the data. If there was also a full publication of the study, we collated and critically appraised all available data. If an included study was marked as a completed study in a clinical trial register but no additional information (study results, publication, or both) was available, we added this study to the Characteristics of studies awaiting classification table.

Assessment of risk of bias in included studies

Two review authors (HA, OK) independently assessed the risk of bias for each included study. We resolved disagreements by consensus, or by consultation with a senior review author (GEHF). In the case of disagreement, we consulted the remainder of the review authors team and made a judgement based on consensus. If adequate information was not available from the publications, study protocols or other sources, we contacted study authors to request missing data on 'Risk of bias' items.

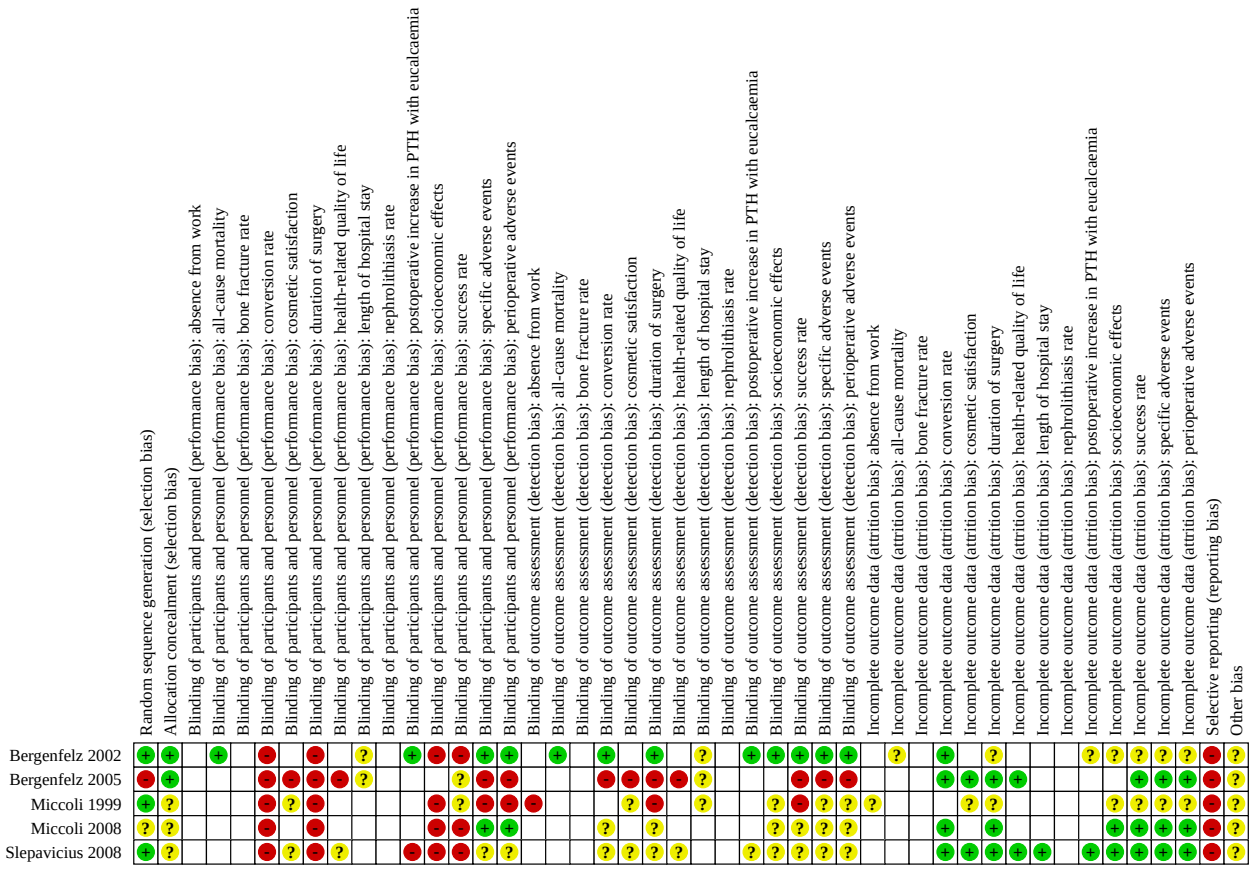
We used the Cochrane 'Risk of bias' assessment tool (Higgins 2011a; Higgins 2017), assigning assessments of low, high or unclear risk of bias (for details, see Appendix 2; Appendix 3). We evaluated individual bias items as described in the *Cochrane Handbook for Systematic Reviews of Interventions,* according to the criteria and associated categorisations contained therein *(* Higgins 2017).

Summary assessment of risk of bias

We presented a 'Risk of bias' graph (Figure 2) and a 'Risk of bias' summary figure (Figure 3).

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies (blank cells indicate that the particular outcome was not measured in some studies).

Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study (blank cells indicate that the study did not measure that particular outcome).



We distinguished between self-reported, investigator-assessed and adjudicated outcome measures.

We defined the following outcomes as self-reported by participants.

- Total incidence of perioperative adverse events.
- Postoperative pain.
- Health-related quality of life.
- Cosmetic satisfaction.

We required the following outcomes to be investigator-assessed and objectively measured by a study personnel.

- Success rate.
- Specific adverse events.
- Conversion rate from minimally invasive to open surgery.
- Postoperative increase in PTH with eucalcaemia.
- All-cause mortality.
- Bone fracture rate.
- Nephrolithiasis rate.
- Duration of surgery.
- Length of hospital stay.
- Socioeconomic effects.

Risk of bias for a study across outcomes

Some 'Risk of bias' domains, such as selection bias (sequence generation and allocation sequence concealment), affect the risk of bias across all outcome measures in a study. In case of high risk of selection bias, we marked all endpoints investigated in the associated study as high risk. Otherwise, we did not perform a summary assessment of the risk of bias across all outcomes for a study.

Risk of bias for an outcome within a study and across domains

We assessed the risk of bias for an outcome measure by including all entries relevant to that outcome (i.e. both study-level entries and outcome-specific entries). We considered low risk of bias to denote a low risk of bias for all key domains, unclear risk to denote an unclear risk of bias for one or more key domains, and high risk to denote a high risk of bias for one or more key domains.

Risk of bias for an outcome across studies and across domains

These are the main summary assessments that we incorporated into our judgements about the certainty of the evidence in the

'Summary of finding' table. We defined outcomes as at low risk of bias when most information came from studies at low risk of bias, unclear risk when most information came from studies at low or unclear risk of bias, and high risk when a sufficient proportion of information came from studies at high risk of bias.

We graded the overall certainty of the evidence for each outcome using the GRADE approach. The approach classified the certainty of the evidence into four categories: high, moderate, low and very low. It took into account study design and the following factors: risk of bias, imprecision, inconsistency, indirectness, publication bias, large effect size, dose-response effect and confounding (Guyatt 2011; Meader 2014).

Measures of treatment effect

We expressed dichotomous data as risk ratio (RR) or hazard ratio (HR) with 95% confidence intervals (CI). We expressed continuous data as mean differences (MDs) with 95% CI when studies used the same scale and standardised mean differences (SMD) with 95% CI when studies used difference scales (Deeks 2017; Hozo 2005; Riley 2011).

Unit of analysis issues

The unit of analysis was the individual participant for specific outcomes (clarification: the meta-analyses is based on group-level data). We took into account the level at which randomisations occurred, such as cluster-randomised studies and multiple observations for the same outcome Higgins 2011b).

Dealing with missing data

We attempted to obtain relevant missing data from authors of included studies. If unsuccessful, we used a complete-case approach in the main analysis. We then conducted sensitivity analyses using plausible assumptions about the outcomes of participants with missing outcome data to test the robustness of statistically significant results.

For both continuous and dichotomous data, we imputed plausible treatment effects using progressively stringent criteria, as outlined by Akl and Ebrahim (Akl 2013; Ebrahim 2013).

Assessment of heterogeneity

In the event of substantial clinical or methodological heterogeneity, we did not report study results as a pooled effect estimate in the meta-analysis. We identified heterogeneity (inconsistency) by visual inspection of the forest plots and by using a standard Chi^2 test with a significance level = 0.1 (Deeks 2017). In view of the low power of this test, we also considered the I^2 statistic, which quantifies inconsistency across studies to assess the impact of heterogeneity on the meta-analysis (Higgins 2002; Higgins 2003).

Had we found heterogeneity, we would have attempted to determine possible reasons for this by examining individual study and subgroup characteristics.

Assessment of reporting biases

If we included 10 or more studies that investigated a particular outcome, we used funnel plots to assess small-study effects. Several explanations may account for funnel plot asymmetry, including true heterogeneity of effect with respect to study size, poor methodological design (and hence small-study bias) and publication bias (Sterne 2017). Therefore, we interpreted the results carefully (Sterne 2011).

Data synthesis

We conducted meta-analyses using a random-effects model (Wood 2008). In addition, we performed statistical analyses according to the statistical guidelines referenced in the latest version of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011b). In the event of substantial clinical, methodological or statistical heterogeneity, we considered whether to report the study results as meta-analytically pooled effect estimates.

We planned to undertake (or display) a meta-analysis only if participants, interventions, comparisons and outcomes were judged to be sufficiently similar to ensure an answer was clinically meaningful. Unless good evidence showed homogeneous effects across studies, we primarily summarised low risk of bias data using a random-effects model (Wood 2008). We interpreted random-effects meta-analyses with due consideration to the whole distribution of effects, ideally by presenting a prediction interval (Borenstein 2017a; Borenstein 2017b; Higgins 2009). A prediction interval specifies a predicted range for the true treatment effect in an individual study (Riley 2011). For rare events such as event rates below 1%, we used Peto's odds ratio method, provided that there was no substantial imbalance between intervention and comparator group sizes and intervention effects were not exceptionally large. In addition, we performed statistical analyses according to the statistical guidelines presented in the *Cochrane Handbook for Systematic Reviews of Interventions* (Deeks 2017).

Subgroup analysis and investigation of heterogeneity

We expected some characteristics to introduce clinical heterogeneity, and we planned to carry out the following subgroup analysis including investigation of interactions (Altman 2003).

- Surgeon speciality.
- Academic versus non-academic setting.
- High-volume versus low-volume groups.
- IOPTH criteria used.

Sensitivity analysis

We planned to performed sensitivity analyses to explore the influence of some factors (when applicable) on effect sizes by restricting analysis to the following.

- Published studies.
- Effect of risk of bias, as specified in the Assessment of risk of bias in included studies section.
- Taking into account missing data information.
- Very long or large studies to establish the extent to which they dominated the results.
- Use of the following filters: diagnostic criteria, imputation, language of publication, source of funding (industry versus other) or country.

We tested the robustness of the results by repeating analysis using different measures of effect size (e.g. RR, OR, etc.) and different statistical models (fixed-effect and random-effects models).

Certainty of the evidence

We presented the overall certainty of the evidence for each outcome specified below according to the GRADE approach, which takes into account issues related to internal validity (risk of bias, inconsistency, imprecision, publication bias) and to external validity, such as directness of results. Two review authors (HA, OK) independently rated the certainty of the evidence for each outcome.

We included Appendix 13 entitled 'Checklist to aid consistency and reproducibility of GRADE assessments', to help with standardisation of the 'Summary of findings' table (Meader 2014). Alternatively, we used the GRADEpro GDT software and presented evidence profile tables as an appendix (GRADEpro GDT 2014). We presented results for the outcomes as described in the Types of outcome measures section. If meta-analysis was not possible, we presented the results narratively in the 'Summary of findings' table. We justified all decisions to downgrade the certainty of the evidence by using footnotes, and we made comments to aid the reader's understanding of the Cochrane Review when necessary.

'Summary of findings' table

We presented a summary of the evidence in Summary of findings 1. This provided key information about the best estimate of the magnitude of effect, in relative terms and as absolute differences for each relevant comparison of alternative management strategies, numbers of participants and studies that addressed each important outcome, and a rating of overall confidence in effect estimates for each outcome. We created Summary of findings 1 using the methods described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Schünemann 2017) along with Review Manager 5 (Review Manager 2014).

The intervention presented in the 'Summary of findings' table was MIP and the comparator was BNE.

We reported the following outcomes, listed according to priority.

- Success rate.
- Total incidence of perioperative adverse events.
- All-cause mortality.
- Health-related quality of life.
- Cosmetic satisfaction.
- Duration of surgery.
- Length of hospital stay.

RESULTS

Description of studies

For a detailed description of studies, see Table 1, Characteristics of included studies table and Characteristics of excluded studies table.

Results of the search

Search in the electronic databases identified 773 articles after removal of duplicates. From reading titles and abstracts, we assessed 45 full-text articles for eligibility after initial screening. After full-text screening, we excluded 37 publications because they did not meet our inclusion criteria (see Characteristics of excluded studies table), leaving six publications reporting on five studies that were included in this review (see Characteristics of included studies table; Figure 1).

Included studies

A detailed description of the characteristics of included studies is included in the Characteristics of included studies table and Appendix 4; Appendix 5; Appendix 6; Appendix 9. The following is a succinct overview.

All five retrieved studies were written in English (Bergenfelz 2002; Bergenfelz 2005; Miccoli 1999; Miccoli 2008; Slepavicius 2008). The studies were published between 1999 and 2008, and all were conducted in European university hospitals (Germany, Italy, Lithuania, Sweden).

Source of data

All data were obtained from the published literature. We requested Information from all study authors, a reply was only received from Dr Bergenfelz via email correspondence (Bergenfelz 2002; Bergenfelz 2005).

Comparisons

Four studies employed a focused surgical technique in the MIP group, wherein the surgeon only targeted the suspicious gland identified on imaging (Bergenfelz 2005; Miccoli 1999; Miccoli 2008; Slepavicius 2008). Bergenfelz 2002 employed a unilateral technique where surgery was started on the side indicated by the preoperative scintigram or on the left side whenever the scintigram failed to localise any enlarged parathyroid glands. In participants with non-localising preoperative imaging, Slepavicius 2008 performed bilateral intraoperative internal jugular vein PTH sampling to guide the surgery (Appendix 4).

Other observed differences included anaesthesia, use of a videoscopic technique, IOPTH criteria and criteria for conversion of surgery.

Overview of study populations

The five studies included 266 participants, 136 participants were randomised to MIP and 130 to BNE. Data were available for all participants postsurgery up until one year but data were missing for two participants in the MIP group and for one in the BNE group at one year and for nine participants in the MIP group and for 11 in the BNE group at five years (Table 1).

The mean age in four studies was 62 years, with a male to female ratio of 1:5 (Bergenfelz 2002; Bergenfelz 2005; Miccoli 1999; Miccoli 2008). In the study by Slepavicius and colleagues, authors reported an age range of study participants between 18 and 90 years, but did not specify gender (Slepavicius 2008). The mean preoperative serum calcium level for included studies was 2.83 mmol/L (11.32 mg/dL), with a mean PTH level of 22.75 pmol/L (206.81 pg/mL). Only one study reported that 13/91 participants were asymptomatic prior to surgery (Bergenfelz 2002). None of the studies reported the proportion of participants who had fractures or kidney stones at baseline (Appendix 5; Appendix 6).

Two studies reported on gender distribution in each treatment group, and had a male to female ratio of 1:4.6 in the MIP group and 1:4 in the BNE groups. In the three studies reporting age by

treatment group, the intervention group age was 61.6 years old on average, while the control group was 64 years old (Bergenfelz 2002; Bergenfelz 2005; Miccoli 2008). The mean calcium concentration was 2.84 mmol/L (11.36 mg/dL) in the intervention group and 2.82 mmol/L (11.28 mg/dL) in the control group in all five studies. The mean PTH in the MIP group was 20.32 pmol/L (191.1 pg/mL), and 20.05 pmol/L (188.5 pg/mL) in the BNE group (Appendix 10; Appendix 11).

Criteria for entry into the individual studies are outlined in the Characteristics of included studies table.

Inclusion criteria were not explicitly stated in two studies (Miccoli 2008; Slepavicius 2008). One study included all participants with PHPT based on a serum calcium level of more than 2.60 mmol/L (10.4 mg/dL), a serum PTH level of more than 3.5 pmol/L (33.0 pg/mL), and a serum creatinine level of less than 200 µmol/L (2.26 mg/dL) (Bergenfelz 2002). The other two studies included only those participants with biochemical confirmed (Miccoli 1999), or sporadic hyperparathyroidism (Bergenfelz 2005), and concomitant suspicion for single gland disease on preoperative imaging (Bergenfelz 2005; Miccoli 1999).

Exclusion criteria were clearly delineated by all studies except in Miccoli 2008. Those studies excluded participants with prior history of neck surgery, those who had indications or anticipation of thyroidectomy, a family history of PHPT such as multiple endocrine neoplasia (MEN 1, MEN 2), hereditary PHPT and anticipated or planned simultaneous thyroid operations. Three studies also mentioned that participants were only allowed to participate if they could comprehend the information given to them (Bergenfelz 2002; Bergenfelz 2005; Slepavicius 2008). Two studies excluded pregnant or breastfeeding women, and people with hypercalcaemic crisis (Bergenfelz 2002; Slepavicius 2008). Slepavicius 2008 also excluded people with severe concomitant pathology making surgical treatment impossible, while Bergenfelz 2005 excluded people with allergy to drugs used for local anaesthesia.

Study design

Studies were RCTs employing a parallel-group superiority design. The study protocols differed in randomisation timing. Two studies randomised participants into treatment groups (MIP or BNE) even before preoperative imaging determined the likelihood of a single gland disease, and participants randomised to the BNE group did not undergo any preoperative imaging (Bergenfelz 2002; Slepavicius 2008). The other three studies randomised participants into MIP or BNE only after preoperative diagnostic studies were performed (Bergenfelz 2005; Miccoli 1999; Miccoli 2008), but only two studies explicitly specified that they included only participants with suspicion of single gland disease on preoperative imaging prior to the randomisation process (Bergenfelz 2005; Miccoli 1999).

Slepavicius 2008 was run at two centres (Department of Abdominal and Endocrine Surgery of Klaipeda University Hospital and Second Department of Abdominal Surgery of Vilnius University Hospital "Santariskiu Klinikos", Vilnius, Lithuania). All other studies were run at a single university hospital.

In terms of blinding, no study was double-blinded for participants and personnel, two studies were single-blinded for participants (Bergenfelz 2002; Miccoli 2008), and the other three studies did not define blinding (Bergenfelz 2005; Miccoli 1999; Slepavicius 2008). Only one study blinded outcome assessors (Bergenfelz 2005).

Studies were performed between 1996 and 2008.

The duration of follow-up ranged from six months to five years. No study was terminated early.

Settings

All studies were conducted in European university hospitals (Germany, Italy, Lithuania, Sweden). Surgeries were all performed in a hospital setting.

Interventions

Imaging techniques varied among studies as well prior to surgeries. Two studies used both ultrasound and Sestamibi scanning (Miccoli 2008; Slepavicius 2008), two studies only used Sestamibi scanning (Bergenfelz 2002; Bergenfelz 2005), and one study only used ultrasound (Miccoli 1999).

Outcomes

All studies reported some of our outcomes, but in some, there was no clear specification of which outcomes were primary or secondary (Appendix 7).

Excluded studies

We excluded 37 publications because they did not meet our inclusion criteria (see Characteristics of excluded studies table).

Studies awaiting classification

We found one study presently published as a conference poster that is awaiting classification (see Characteristics of studies awaiting classification table).

Ongoing studies

We found one ongoing study that was last updated on 24 June 2005 with no information available about the publication of the results of this trial (see Characteristics of ongoing studies table).

Risk of bias in included studies

For details on risk of bias of included studies see Characteristics of included studies table.

For an overview of review authors' judgements about each risk of bias item for individual studies and across all studies see Figure 2 and Figure 3. In general, the risk of bias was unclear to high in most cases.

Allocation

We judged three studies at low risk for selection bias regarding random sequence generation (Bergenfelz 2002; Miccoli 1999; Slepavicius 2008). Bergenfelz 2002 mentioned that they utilised block randomisation through computer software and concealed envelopes. Slepavicius 2008 randomised participants according to a double random principle. Miccoli 1999 randomly divided participants into two groups by flipping a coin. Miccoli 2008 provided no information about random sequence generation. Bergenfelz 2005 answered our query by email and confirmed randomisation was through block randomisation by computer software; however, this randomisation was not balanced.

Therefore, we assessed the risk of bias as high for this study (Bergenfelz 2005).

As for allocation concealment, we judged two studies at low risk of selection bias, because they confirmed the use of sequentially numbered, concealed, opaque envelopes (Bergenfelz 2002; Bergenfelz 2005). Miccoli 1999 utilised coin flipping, but it was unclear if this was done prior to participant presentation or after inclusion, which would determine whether there was adequate allocation concealment. There was insufficient detail to allow a definite judgement in Miccoli 2008. Slepavicius 2008 used envelopes but there was no information about opaqueness of envelopes and accordingly, we judged the risk of bias as unclear.

Blinding

We assessed the risk of blinding of participants and personnel combined, and that of outcome assessors separately. For unreported outcomes, the risk of bias was automatically judged as unclear. Two studies did not provide sufficient information about the blinding of participants (Miccoli 1999; Slepavicius 2008). One study mentioned that participants were not blinded (Bergenfelz 2005). For the two studies that had blinding of their participants, we judged the risk of performance and detection bias to be high for certain outcomes, such as conversion rate from minimally invasive to open parathyroidectomy and duration of surgery, since these outcomes are dependent on the surgeon, who in this instance, was not blinded. Miccoli 1999 and Slepavicius 2008 provided insufficient information on the blinding of participants, but surgeons were not blinded. Hence for the self-reported outcomes that depended on the participants, the risk of bias was judged as unclear. For the outcomes that depended on the surgeon, such as conversion rate or duration of surgery, the risk of bias was judged to be high. Bergenfelz 2005 mentioned that participants were not blinded. Based on our judgement, the risk of performance bias and detection bias was high for the self-reported outcomes (Bergenfelz 2005).

We could obtain adequate information about blinding of outcome assessors in only two studies by email reply to our queries (Bergenfelz 2002; Bergenfelz 2005). Bergenfelz 2002 blinded outcome assessors, and therefore the risk of bias was low for adjudicated/investigator-assessed and self-reported outcomes. Bergenfelz 2005 did not blind outcome assessors, hence we considered the risk of bias to be high for both adjudicated/investigator-assessed outcomes and self-reported outcomes. The other studies did not mention blinding of outcome assessors. In Miccoli 1999, randomisation was done after a scintigraphy preoperatively identified a single adenoma, and therefore, for the outcomes that may have been affected by this approach, we judged the risk of bias to be high and for the other outcomes not affected by this approach, we judged the risk of bias as unclear. In the other two studies, we judged the risk of bias as unclear for both adjudicated/investigator-assessed outcomes and self-reported outcomes, due to the insufficient information on the blinding of the outcome assessors (Miccoli 2008; Slepavicius 2008).

Incomplete outcome data

Outcome data were incomplete in three studies: Bergenfelz 2002 had 47 participants in the MIP group and 44 in the BNE group at study entry, 45 in the MIP and 43 in the BNE at one year, and 38 in the MIP and 33 in the BNE at five years (Bergenfelz 2002). The study authors provided insufficient information regarding missing data and did not clarify how missing data were handled. Therefore, the risk of bias was unclear for outcomes reported at one and five years. However, for the outcomes where data were available for all the participants in the perioperative period, we judged the risk of bias as low. For the perioperative and specific adverse events, which included hypocalcaemia symptoms, reported in the first 24 hours and at six weeks, some data were missing at both times (e.g. at six weeks, data were only available in 19/39 participants (49%) in the BNE group and 12/43 in the MIP group (28%)). The proportion of participants with missing data was 5/43 (11.6%) in the BNE group and 1/44 (2%) in the MIP group, showing disparate attrition rates between the two groups and the study authors did not clarify how missing data were handled. Therefore, we judged the risk of bias for this outcome as unclear (Bergenfelz 2002). Two participants in Miccoli 1999 had multiglandular disease that was discovered during surgery and were excluded from the study. We judged the risk of bias as unclear for all outcomes, as there was insufficient information to assess whether missing data in combination with the method used to handle missing data were likely to induce bias. Slepavicius 2008 excluded three participants who had a conversion from MIP to BNE and two participants who had hyperplasia, which in both instances were unlikely to have had an effect on the outcomes. Therefore, we considered the risk of attrition bias to be low for all outcomes.

Selective reporting

We identified no published protocols for any of the included studies, therefore no study could be assessed as having a low risk of reporting bias. Instead, we relied on comparisons between outcomes listed in the 'Methods' section to those reported in the 'Results' section for each study, and whether the results reported corresponded to the outcomes described under 'Methods' and were reported adequately. We judged all studies at high risk of reporting bias for at least one outcome measure (Appendix 8).

Other potential sources of bias

Four out of five studies provided no details of surgeon familiarity and experience with each surgical technique, and may be a source of bias whenever surgeons were more familiar with one of the tested techniques (Bergenfelz 2002; Miccoli 1999; Miccoli 2008; Slepavicius 2008). Only Bergenfelz 2005 specified that all surgeons performing operations were experienced, but did not detail the level of experience and familiarity with each technique in particular. At least two included studies tested techniques that were introduced by the principle investigator of the study (Miccoli 1999; Miccoli 2008).

None of the studies mentioned sources of funding.

The presence of conflict of interest could not be excluded as a result of those two factors. Therefore, we determined risk of bias as unclear for all those studies.

Effects of interventions

See: **Summary of findings 1** Minimally invasive parathyroidectomy versus bilateral neck exploration for primary hyperparathyroidism in adults

Baseline characteristics

For details of baseline characteristics, see Appendix 5 and Appendix 6.

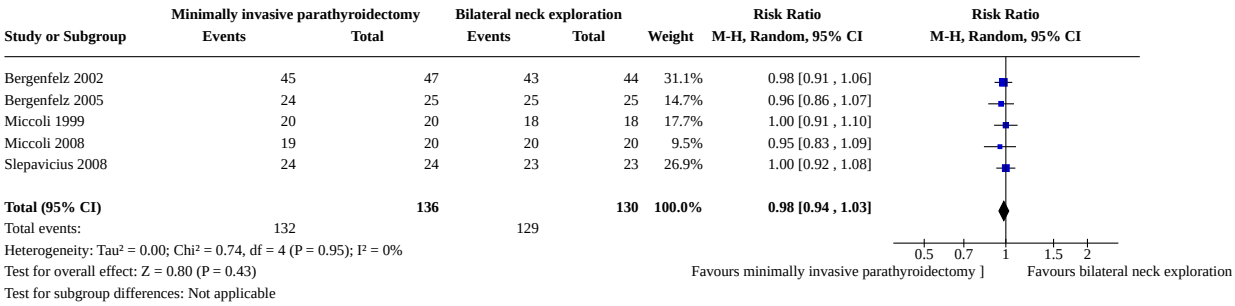
Primary outcomes

Success rate at six months and five years

All five studies reported success rate (eucalcaemia) within six months, one at six weeks postoperatively (Bergenfelz 2002), and another four at six months postoperatively (Bergenfelz 2005; Miccoli 1999; Miccoli 2008; Slepavicius 2008). One study had follow-up data at one and five years postoperatively (Bergenfelz 2002). A total of 132/136 (97.1%) participants in the MIP group compared with 129/130 (99.2%) participants in the BNE group were judged as operative success. The RR of success in the MIP group compared to the BNE group up to six months was 0.98 (95% CI 0.94 to 1.03; P = 0.43; 5 studies, 266 participants; very low-certainty evidence; Analysis 1.1). The 95% prediction interval ranged between 0.90 and 1.06.

Findings are summarised in Figure 4. One article did not report on success rate in the main text but mentioned in the abstract that "No cases of persistent PHPT were present in either group" (Miccoli 1999).

Figure 4. Forest plot of comparison: 1 Minimally invasive parathyroidectomy versus bilateral neck exploration, outcome: 1.1 Success rate at six months.

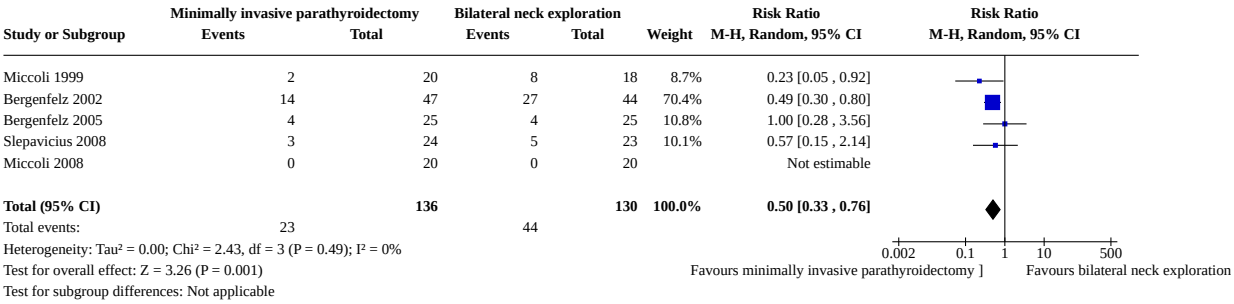


No study reported on the operative success rates between six months and five years. Only one study reported on the success rate at five years postoperatively (Bergenfelz 2002). Of 73 participants with deducible conclusions about operative success at five years, 34/38 (89.5%) participants in the MIP group compared with 37/39 (94.9%) participants in the BNE group were judged as operative success (Analysis 1.2). Six participants had persistent or recurrent disease, four in the MIP group and two in the BNE group. Interestingly, three of the four failures in the unilateral group went on to undergo a bilateral surgery. Two participants in the unilateral group and one in the bilateral group who failed surgery had mutations in the gene for MEN (Bergenfelz 2002).

Total incidence of perioperative adverse events

Five studies reported the total incidence of perioperative adverse events (Bergenfelz 2002; Bergenfelz 2005; Miccoli 1999; Miccoli 2008; Slepavicius 2008; Figure 5). These events occurred in 23/136 (16.9%) participants in the MIP group compared with 44/130 (33.9%) participants in the BNE group. The RR was 0.50, in favour of MIP (95% CI 0.33 to 0.76; P = 0.001; 5 studies, 236 participants; low-certainty evidence). The 95% prediction interval ranged between 0.25 and 0.99. Perioperative adverse events included symptomatic hypocalcaemia, vocal cord palsy, bleeding, fever, infection and others. Miccoli 1999 reported a total incidence of perioperative adverse events (including fever, hypocalcaemia and recurrent laryngeal nerve palsy) occurring in 2/20 (10%) participants in the MIP group and in 8/18 (44.4%) participants in the BNE group. Miccoli 2008 did not report any postoperative adverse complications including haemorrhage, laryngeal nerve palsy or hypocalcaemia. Bergenfelz 2002 reported a total of perioperative adverse events occurring in 14/47 (29.8%) participants in the MIP group compared to 27/44 (61.4%) participants in the BNE group including severe biochemical and symptomatic hypocalcaemia. In addition, 2/47 (4.2%) participants in the MIP group and 5/44 (11.4%) participants in the BNE group had significant complications including tracheal oedema, paresis of recurrent laryngeal nerve, bleeding and serious hypocalcaemia (Bergenfelz 2002). Bergenfelz 2005 reported the occurrence of vocal cord palsy in one participant in the MIP group and the drainage of a wound seroma occurring in one participant in the BNE group. Furthermore, three participants in the MIP group and three in the BNE group reported hypocalcaemia postoperatively. Slepavicius 2008 reported that two (9.6%) participants in the MIP group and four (19%) participants in the BNE group sustained postoperative symptomatic hypocalcaemia. Furthermore, vocal cord palsy occurred in two participants, one each in both groups (Slepavicius 2008). Details on reported adverse events are found in Appendix 10; Appendix 11; and Appendix 12.

Figure 5. Forest plot of comparison: 1 Minimally invasive parathyroidectomy versus bilateral neck exploration, outcome: 1.3 Total incidence of perioperative adverse events.



Secondary outcomes

Specific adverse events

Bleeding events

Two studies reported bleeding events within 48 hours of the intervention (Bergenfelz 2002; Miccoli 2008), and were provided by email reply for one study (Bergenfelz 2005). There was just one reported bleeding event for 91 participants in the three studies. The RR comparing MIP with BNE was 0.31 (95% CI 0.01 to 7.47; P = 0.47; 2 studies, 131 participants; Analysis 1.4).

Infection

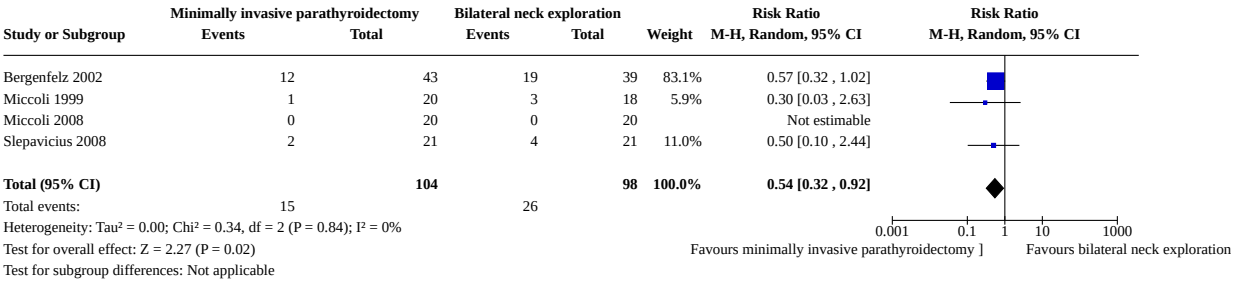
There was one infection reported in the BNE group and none in the MIP group within 48 hours of the intervention in the postoperative period in each of the two studies assessing this outcome (Bergenfelz 2005; Miccoli 1999).

Hypocalcaemia

Five studies reported hypocalcaemia and symptomatic hypocalcaemia. Only four of the studies reported symptomatic hypocalcaemia within 48 hours of surgery (Bergenfelz 2002; Miccoli 1999; Miccoli 2008; Slepavicius 2008), whereas one study reported on participants needing calcium supplementation, an intervention that did not necessarily reflect the incidence of symptomatic hypocalcaemia (Bergenfelz 2005).

Comparing MIP with BNE showed a RR of 0.54 in favour of MIP (95% CI 0.32 to 0.92; P = 0.02; 4 studies, 202 participants; Analysis 1.5). The 95% prediction interval ranged between 0.17 and 1.74. In the MIP group, 15/104 (14.4%) participants experienced symptomatic hypocalcaemia compared with 26/98 (26.5%) participants in the BNE group. Findings are summarised in Figure 6.
However, we must note that if the reported cases who received calcium supplements were an accurate surrogate for the development of symptomatic hypocalcaemia, the results would have been different, with a RR of 0.74 (95% CI 0.42 to 1.31; data not shown). A main argument for the exclusion of Bergenfelz 2005 was that administration of calcium was defined in the 'Methods' section of the publication to be acceptable for all participants with either symptomatic or biochemical hypocalcaemia. Furthermore, the reasons for participants requesting calcium supplementation were not clearly elucidated.

Figure 6. Forest plot of comparison: 1 Minimally invasive parathyroidectomy versus bilateral neck exploration, outcome: 1.3 Symptomatic hypocalcaemia.



Permanent hypocalcaemia or hypocalcaemia persisting for longer than six months occurred in only one participant, belonging to the BNE group (RR 0.33, 95% CI 0.01 to 7.81; P = 0.46; 5 studies, 236 participants; Analysis 1.6).

Postoperative pain

Slepavicius 2008 reported postoperative pain using a 100-point visual analogue scale (0 indicating pain was absent to 100 indicating unbearable pain) at four, eight, 16, 24, 36 and 48 hours after surgery, while Miccoli 1999 reported pain using a 10-point scale (1 indicating no pain to 10 indicating worst pain ever) at 12, 24 and 48 hours after surgery, Miccoli 1999 showed greater pain in the BNE group, but data could not be pooled because of missing data on standard deviations (SD). Bergenfelz 2002 reported pain at one, two, three and four days after surgery, using an undefined visual analogue pain scale supposedly with 0 indicating no pain. Bergenfelz 2005 reported that after surgery,10 participants in the BNE group compared with seven participants in the MIP

group required analgesia for pain. At our prespecified 48-hour postoperative time point, the SMD of the visual analogue scales comparing MIP with BNE was –0.70 (95% CI –1.69 to 0.28; P = 0.16; 2 studies, 133 participants; Analysis 1.7; random-effects model). The fixed-effect model showed an SMD of –0.51 in favour of MIP (95% CI –0.86 to –0.16).

Vocal cord paralysis

Four studies reported the incidence of vocal cord paralysis within 48 hours (Bergenfelz 2005; Miccoli 1999; Miccoli 2008; Slepavicius 2008), and one study before discharge (Bergenfelz 2002). Seven cases (five in MIP and two in BNE) resolved at one month postoperatively except for one case in Miccoli 1999 that persisted at six months postoperatively. Comparing MIP with BNE showed a RR of 1.87 (95% CI 0.47 to 7.51; P = 0.38; 5 studies, 261 participants; Analysis 1.8). The 95% prediction interval ranged between 0.20 and 17.87. A total of 5/133 (3.8%) participants in the MIP group compared with 2/128 (1.6%) participants in the BNE group experienced vocal cord paralysis.

Conversion rate from minimally invasive to open procedure

Four studies reported intraoperative conversion rate information (Bergenfelz 2002; Bergenfelz 2005; Miccoli 2008; Slepavicius 2008). Out of 115 included patients, there were 24 incidences of conversion in total, amounting to a conversion rate of 20.8%.

Postoperative increase in parathyroid hormone with eucalcaemia

Two studies reported postoperative eucalcaemic hyperparathyroidism (Bergenfelz 2002; Slepavicius 2008). A total of 13/68 (19.1%) participants in the MIP group compared with 16/65 (24.6%) participants in the BNE group showed a postoperative increase in PTH with eucalcaemia. The RR of eucalcaemic hyperparathyroidism comparing MIP with BNE was 0.81 (95% CI 0.43 to 1.53; P = 0.51; 2 studies, 133 participants; Analysis 1.9).

All-cause mortality

No study explicitly reported on the occurrence of perioperative mortality; however, complete data reporting on all patients in four studies led us to deduce that there were no instances of perioperative mortality within six months (Bergenfelz 2005; Miccoli 1999; Miccoli 2008; Slepavicius 2008). Bergenfelz 2002 reported on all participants at six weeks postoperatively, and reported two deaths during the follow-up period of one year without specifying in which treatment group these deaths had happened. Fourteen other deaths occurred in the same study within the five-year follow-up period, but neither the cause of death nor in which treatment group these had happened was stated (Bergenfelz 2002). The overall certainty of the evidence was very low.

Health-related quality of life

One study reported health-related quality of life only using the 36-item Short Form (SF-36) (Slepavicius 2008) (see Appendix 14). The authors found no difference between the treatment groups, but did not present data (very low-certainty evidence). Similarly, in an email response, Dr Bergenfelz indicated that an SF-36 survey was used in one study but the difference was judged to be clinically unimportant and was not published (Bergenfelz 2005).

Cosmetic satisfaction

Slepavicius 2008 reported cosmetic satisfaction using a modified Holander scale (ranging from 0 to 7, with 0 indicating optimal result and 1 to 7 suboptimal result), describing the overall cosmetic appearance of the wound at two days, one month, six months and one year postoperatively. There was a statistically significant difference in favour of the MIP group at the first three time points, with a score of 2 at two days, 1.4 at one month and 1.6 at six months for the MIP group compared to a score of 3.9 at two days, 3.4 at one month and 2.5 at six months for the BNE group. However, this difference became statistically non-significant one year postoperatively.

Miccoli 1999 assessed cosmetic satisfaction using personal opinions by physicians about the aesthetics of the scar with an undefined 10-point score at one month, three months and six months, postoperatively. Patient satisfaction was higher in the MIP group at all three time points.

Neither study provided SDs, and we were thus unable to pool the results. The consistency in findings among studies suggest that there was a treatment benefit for MIP as compared to BNE in terms of cosmetic satisfaction (very low-certainty evidence).

Bergenfelz 2005 mentioned in his study that all participants in both groups stated that they were satisfied with the cosmetic result of the surgery, but did not report the scale used.

Some studies reported on surrogate outcomes related to cosmetic satisfaction that we had not identified as outcomes in our protocol. Slepavicius 2008 reported on scar length using a flexible tape at 12 months postoperatively, with median scar length being shorter in the MIP group (1.9 cm) than in the BNE group (8 cm).

Bone fracture rate

None of the studies reported bone fracture rates.

Nephrolithiasis rate

None of the studies reported nephrolithiasis rates.

Absence from work

Miccoli 1999 reported that the postoperative inactivity period was shorter in participants treated with MIP (mean 2 (SD 5.5) days in MIP group versus 16 (SD 6) days in BNE group) but did not specifically mention the number of days of work missed due to surgery.

Duration of surgery

All five studies reported on the duration of surgery from skin incision to wound closure. Data from Miccoli 2008 and Bergenfelz 2005 could not be included in our meta-analysis. Bergenfelz 2005 reported a difference in the duration of surgery in favour of MIP, but reported the data as median and range (41 minutes, 19 minutes to 120 minutes) in the MIP group versus 63 minutes (35 minutes to 110 minutes) in the BNE group)). Miccoli 2008 reported that there was no statistically significant difference in operative time between treatment groups, with the mean duration of the MIP group being 33 minutes as opposed to 32 minutes in the BNE group (SDs were not reported).

In the three other studies, comparison of duration of surgery showed a benefit for the MIP group compared to the BNE group (MD

–18 minutes, 95% CI –31 to –6; P = 0.004; 3 studies, 171 participants; very low-certainty evidence; Analysis 1.10). The 95% prediction interval ranged between –162 minutes and 126 minutes.

Length of hospital stay

In Bergenfelz 2005, participants were asked to stay for four days for postoperative observation period. In Bergenfelz 2002, it was mentioned that after the first postoperative day the patient stayed at the hotel of the hospital and made individual visits on postoperative days two to four. Two studies discharged participants in the second day after surgery and it was not mentioned if there was a difference between the groups in the length of hospital stay (Miccoli 1999; Slepavicius 2008). One study mentioned that participants were discharged on the first postoperative day without stating if there was a difference between the two groups (Miccoli 2008).

Socioeconomic effects

Four studies reported on costs incurred from MIP and BNE. Bergenfelz 2002 calculated the costs from official in-hospital charges obtained from different departments, and reported a trend towards extra cost in the MIP group compared with the BNE group (mean cost: USD 2258 (SD 509) with MIP versus USD 2097 (SD 505) with BNE). Slepavicius 2008 found a difference in costs of the procedures in favour of BNE (EUR 1428 with MIP versus EUR 1166 with BNE; P < 0.05). Miccoli 2008 compared costs of specific procedures but not of hospital charges and stated that MIP was relatively cheaper than BNE, with the difference in costs mainly resulting from the increased duration of operating room use. Miccoli 1999 reported that the overall costs of BNE was USD 1910 compared with USD 1720 in favour of MIP, but stated that the difference was minimal. Both Bergenfelz 2002 and Slepavicius 2008 performed preoperative localisation in the MIP group exclusively, and the costs for this localisation made up the greatest proportion of the difference in costs observed by other studies. Miccoli 1999 differed by including only participants with suspicion for single adenoma on preoperative imaging in their study. Both MIP and BNE groups had preoperative imaging prior to randomisation and the costs of these procedure were not included in the cost calculation. We performed no meta-analyses because only one study provided SDs (Bergenfelz 2002).

Subgroup analyses

We did not perform subgroups analyses due to the small number of studies and study participants.

Sensitivity analyses

We used commonly suggested approaches for dealing with participants with missing outcome data in the meta-analyses (Akl 2015). In the primary meta-analysis, we used a complete-case analysis (i.e. we excluded participants with missing outcome data).

When the primary meta-analysis resulted in a statistically significant result, we tested its robustness by running sensitivity meta-analyses. For those sensitivity meta-analyses, we applied increasingly stringent but plausible assumptions about the outcomes of participants with missing data defined a priori (Akl 2013; Ebrahim 2013)

We conducted sensitivity analysis for hypocalcaemia and duration of surgery.

Hypocalcaemia

The sensitivity analysis found that the pooled effect estimate lost statistical significance for about half of these analyses. See Table 2 for detailed description of all the assumptions used.

Similarly, the sensitivity analysis based on restricting analyses to studies with low risk of bias (Miccoli 2008; Slepavicius 2008) showed that the pooled effect estimate lost statistical significance (Table 2).

Duration of surgery

The sensitivity analysis found that the pooled effect estimate did not lose statistical significance for any of these analyses (Table 2).

Similarly, the sensitivity analysis based on restricting to studies with low risk of bias (Slepavicius 2008) showed that the pooled effect estimate did not lose statistical significance (Table 2).

Assessment of reporting bias

We did not draw funnel plots due to limited number of studies (five).

DISCUSSION

Summary of main results

Our review identified five studies comparing MIP to BNE. MIP was associated with a shorter duration of surgery and lower incidence of symptomatic hypocalcaemia; however, the latter finding lost significance in the sensitivity analyses. There was a tendency towards greater cosmetic satisfaction, lower total adverse events and increased vocal cord paralysis in the MIP group. There were clear differences in eucalcaemic hyperparathyroidism, all-cause mortality and health-related quality of life. Information on bleeding events, infections or aggregated costs was limited either due to a very low incidence of events or lack of meta-analysable data.

Overall completeness and applicability of evidence

We conducted a comprehensive review of three major databases and reviewed each of them for potentially eligible studies. Identified studies encompassed a wide range of currently utilised minimally invasive and BNE techniques. We believe that we have adequately summarised all pertinent randomised controlled data available to date, but several important patient-related outcomes, such as fracture rate, nephrolithiasis rate, missed days from work and length of hospitalisation, were not reported in the published literature. There was insufficient power to detect other patient-related outcomes, such as bleeding rate, mortality, short-term success rate and importantly, long-term success rate. Furthermore, all studies are from European centres, and this limits the applicability to non-European populations or low- and middle-income countries.

Two types of indirectness pertaining mainly to patient populations and intervention designs were encountered in all studies (Guyatt 2011), and directly challenge the applicability of the evidence provided to everyday practice. Two studies included participants with suspicion of single gland disease. They therefore aimed to answer the question of optimal surgery in people who have had prior imaging, while our original aim was to explore the optimal strategy in the general population presenting with sporadic PHPT without any a priori knowledge of any imaging. Furthermore, three studies performed preoperative imaging even on people who

eventually received BNE (Bergenfelz 2005; Miccoli 1999; Miccoli 2008). Surgeon knowledge of preoperative imaging findings may affect intraoperative decisions about which glands to excise, the thoroughness of the exploration and the timing of terminating the surgery. Conversely, the other two study designs employed randomised to MIP or BNE first, and then proceeded with localisation regardless of results (Bergenfelz 2005; Slepavicius 2008). We suggest a more direct and appropriate protocol that would provide an objective risk-benefit assessment of both techniques, on prespecified patient outcomes, in the short term as well as in the long term.

Quality of the evidence

Details about study methodology were limited and brief. In general, the scarcity of published studies coupled with the small sample size of each contributed to imprecision in observed treatment effects for most outcomes of interest (see Summary of findings 1). There was serious risk of bias, especially for self-reported outcomes. Coupled with observed indirectness in answering our review's questions, these factors meant that the certainty of the available evidence ranged generally from low to very low (see Summary of findings 1).

Potential biases in the review process

The main limitations in this review were the indirectness and the small number of retrieved studies. There were fewer than 400 events in all cumulative outcomes of interest, limiting precision and power to detect relevant differences between the two procedures. Furthermore, both MIP and BNE appeared to be safe procedures, with infrequent adverse effects and failures. Detecting differences between procedures would require large populations with the preferred randomisation approach as detailed above. However, this would be difficult to compile in RCTs in the near future.

Agreements and disagreements with other studies or reviews

Other reviews

We could not identify any meta-analysis on the topic within the time frame under consideration. We only identified one other review discussing this topic (Gracie 2012). The review agreed with our findings that MIP and BNE were comparable in terms of success rate and complication rate. However, the review cited advantages in operative duration, learning curve, and cost-analyses and recommend treatment with MIP for people with solitary parathyroid adenomas and those were based on data or expert opinion. Interestingly, the learning curve and cost-effectiveness analyses were not outcomes of any of their reviewed papers. Three studies evaluated costs but not cost-effectiveness. With regards to costs, our review found that, contrary to the Gracie 2012 findings, strategies employing BNE in all participants without exposing them to preoperative imaging are likely to be less costly than those that aim for MIP in participants employing localising preoperative scans (Gracie 2012). Our review confirms the observation that MIP takes less time, but we considered the observed MD of less than 20 minutes to be clinically irrelevant, especially since the CI for the finding included a time saving of about only six minutes.

Non-randomised controlled trial data

We found no non-RCT publications comparing nephrolithiasis rates, absenteeism from work or even cosmetic satisfaction.

Success rates

We found no difference in success rate between MIP and BNE at six months. In our search, most retrieved non-RCT studies found no difference in success rates within six months between MIP and BNE, namely 97% in MIP and 99% in BNE (Adler 2008; Baliski 2008; Beyer 2007; Chen 1999; Genc 2003; Grant 2005; Irvin 2004). Success rates were similarly high to those observed in our review, and in fact were 100% in several studies (Adler 2008; Baliski 2008; Genc 2003). Others reported varying success rates, above 90%, namely 97% in MIP and 94% in BNE (Irvin 2004), 98% in MIP and 94% in BNE (Boggs 1999), 97% for both MIP and BNE (Grant 2005), 99% in MIP and 100% in BNE (Beyer 2007), 98% in MIP and 94% in BNE (Carneiro 2000), and 100% in MIP and 97.3% in BNE (Chen 1999).

One study described differences in outcomes between treatment strategies. Bergenfelz 2007 presented findings from a large audit of parathyroid surgeries performed in Scandinavian countries between 2004 and 2006. They showed in a multivariate analysis an increased chance for alleviation of hypercalcaemia in individuals who underwent unilateral or focused surgery as compared to those undergoing BNE. The audit showed that the overall cure rate was lower than reported in the literature and that of hypocalcaemia to be somewhat higher in MIP compared with BNE.

Viewed in total however, we consider the findings from the above retrospective and non-randomised prospective studies to agree with our own conclusions. Fewer studies looked at recurrence beyond six months. Beyer 2007 reported a long-term success rate of 107/109 participants in the BNE group, with both failures occurring after the six months' time interval, compared to 109/111 participants in the MIP group. Of note, the MIP had a significantly shorter median follow-up of 3.7 (SD 5.6) months. Boggs 1999 and Carneiro 2000 found in a subset of participants having long-term follow-up and successful initial surgery, adjudicated at the six month time point, that 5/176 participants in the BNE group had recurrence with a mean follow-up of 9.3 years (6 to 313 months); this was compared with 2/144 participants at 2.3 years (6 to 85 months) median follow-up in the MIP group. The number of events, combined with the differences in follow-up time between MIP and BNE, did not allow for a confident conclusion to be made about long-term success rate.

Adverse events

In agreement with our findings, three non-RCT studies reported no substantial differences in total complication rates between the two surgical approaches. Specifically, adverse event rates were 6.3% (Adler 2008), 8.9% (Baliski 2008), and 2.2% (Chen 1999) in the BNE group compared to 3.1% (Adler 2008), 5.3% (Baliski 2008), and 0% (Chen 1999) in the MIP group.

Postoperative increase in parathyroid hormone with eucalcaemia

We only found two RCTs reporting on eucalcaemic hyperparathyroidism with a RR of 0.81 (95% CI 0.43 to 1.53) in MIP compared with BNE (Bergenfelz 2002; Slepavicius 2008). Similarly, several non-RCTs found no substantial differences in the incidence of eucalcaemic hyperparathyroidism following surgery. In two studies, the incidence of eucalcaemic hyperparathyroidism was 19/176 participants (Beyer 2007) and 34/109 participants (Carneiro 2000) in the BNE group, compared to 19/144 participants (Beyer 2007) and 50/111 participants (Carneiro 2000) in the MIP group.

All-cause mortality

None of our included studies and very few studies in other literature explicitly stated the number of fatalities from surgery. We propose that this is mainly because of an expectation for a safe parathyroid surgery. Nonetheless, two non-RCTs specifically reported having had no mortalities in 198 MIPs and 290 BNEs procedures (Adler 2008; Agrawal 2014).

Health-related quality of life

Two RCTs found no substantial difference in health-related quality of life between the different treatment groups using the SF-36 questionnaire. In contrast, Adler 2008 in a non-RCT measured similar SF-36 health survey outcomes, and noted improvements at a one-week time point in four scales for MIP (vitality, role-emotional, mental health and mental component summary) versus just two scales in the BNE (vitality and general health). At the one-year interval, MIP participants improved significantly in eight scales (physical functioning, role-physical, bodily pain, vitality, social functioning, role-emotional, mental health and the mental component summary scales) compared to only four scales (general health, vitality, mental health and the mental component summary scales) in the BNE group (Adler 2008).

Bone fracture rate

We found no fracture rate or bone fragility data stratified by treatment group in non-RCT literature.

Duration of surgery

In our meta-analysis, MIP had a shorter duration of surgery than BNE. However, in several non-RCTs there was no significant difference in duration of surgery between groups. Baliski 2008 reported a mean procedure time of 93 minutes in 56 participants in the BNE group compared to 74 minutes in 19 participants in the MIP group. Agrawal 2014 also found similar durations, with the mean duration being 44 minutes in 93 participants receiving MIP successfully compared to a mean duration of 56 minutes in 20 participants receiving BNE. Of note, the duration of surgery for seven participants who were planned to receive MIP but had a BNE instead was considered. Beyer 2007 found a longer procedure time for participants receiving BNE without IOPTH (129 minutes) when compared to those with MIP and IOPTH (119 minutes). However, this difference may have been caused by the significantly greater proportion of people undergoing concomitant thyroid surgery in the BNE group. When these participants were excluded, the authors found no persistent significant difference in operative duration (Beyer 2007).

Length of hospital stay

In one of the included studies, participants were asked to stay for four days for a postoperative observation period (Bergenfelz 2005). In Bergenfelz 2002, it was mentioned that after the first postoperative day individual visits on postoperative days two to four were made in the patient hotel of the hospital. Slepavicius 2008 discharged patients on the second day after surgery and it was not mentioned if there was a difference between the groups in the length of hospital stay.

There was some variability in length of hospital stay in the literature retrieved, and this may pertain to physician preferences and practices. In one study, the mean length of hospital stay for a total of 33 participants undergoing MIP was 0.3 days compared with 1.8 days for BNE (Chen 1999). However, as several other studies did not report a relevant difference among treatment groups with regards to mean hospitalisation times for MIP being 0.2 days (Adler 2008), 1.06 days (Baliski 2008), and one day (Grant 2005), compared to 0.9 days (Adler 2008), 1.2 days (Baliski 2008), and 1 day (Grant 2005) for BNE.

Socioeconomic effects

RCT findings about cost efficiency were variable, with one reporting a significant cost saving for BNE (Slepavicius 2008), another a non-significant difference (Bergenfelz 2002), and a third a relative cost-effectiveness for MIP (Miccoli 1999). This variability may be explained by difference in hospital charges for imaging/procedures, as well as differences in the surgical protocol. Similarly, other literature produced variable results. It seems that cost-effectiveness of a surgical approach may be more an institutional than a generalisable characteristic. Baliski 2008 used unit costs from the St. Paul's Hospital Cost Model (SPHCM) in conjunction with retrospectively collected data about hospital length of stay and intraoperative complications. Total costs were USD 4524 for BNE compared to USD 4961 for MIP, translating to a cost of USD 28,439 per complication avoided by using MIP. Diagnostic testing and operative duration were found to be the main determinants of differences in cost, with diagnostics in particular costing USD 750 more for MIP compared to BNE. Even when changing parameters such as success rates and unit costs to other values cited in the literature, the smallest cost per complication avoided for MIP was USD 11,599 (Baliski 2008). In contrast in Beyer 2007, participants undergoing MIP surgery with IOPTH had lower mean charges (USD 3667) when compared to those undergoing BNE without IOPTH (USD 4787) and those undergoing BNE with IOPTH (USD 4272) . Similarly, in Chen 1999, the mean total hospital charge for 33 participants who underwent MIP was USD 3174 compared with USD 6328 for BNE (Beyer 2007).

Hypocalcaemia

All five studies reported symptomatic hypocalcaemia. Only four of the studies reported symptomatic hypocalcaemia within 48 hours of surgery. One study reported on participants' needs for calcium supplementation, which we judged as insufficient information to accurately represent the incidence of symptomatic hypocalcaemia (Bergenfelz 2005). The incidence of perioperative hypocalcaemia was significantly lower in the MIP group compared to the BNE group, but it was lost on sensitivity analyses. Adler 2008 noted a slightly increased incidence of transient hypocalcaemia in the BNE group (3.8%) as compared to the MIP group (2.6%) in their postoperative results reported within two weeks, and it was mentioned that those participants required oral calcium supplements for a total duration of two weeks. However, all their complications results were not substantially different between the groups.

AUTHORS' CONCLUSIONS

Implications for practice

Based on very low-certainty evidence, success rates between minimally invasive parathyroidectomy (MIP) or bilateral neck exploration (BNE), either in short-term or long-term follow-up were comparable in people with primary hyperparathyroidism undergoing parathyroid surgery for the first time. There was very

low-certainty evidence suggesting that MIP was associated with shorter duration of surgery.

Implications for research

Perhaps the greatest value of this systematic review is that it highlights the gap in knowledge for two frequently used surgical interventions, although surgeons today seem to favour MIP over BNE, with no firm evidence to do so. There were only five randomised controlled studies about the subject, and only one had a long-term follow-up revealing a concerning tendency for recurrence in the MIP group. We highlight two important areas future research must focus on. An effort must be made to include a greater number of participants in order to optimise power in large multicentre, and possibly multinational studies, and such future RCTs must aim to implement designs that seek to emulate current practice. Since all patients with negative imaging should undergo BNE, we suggest a design wherein patients are randomised to either receive preoperative imaging or surgery without imaging. Those who receive imaging would then undergo MIP if findings suggest single gland disease, or BNE if the scan is found to be non-localising. All patients who are randomised to surgery without imaging would receive BNE.

ACKNOWLEDGEMENTS

We wish to thank the staff of the Cochrane Metabolic and Endocrine Disorders group for their comments and feedback, with special thanks to Assistant Managing Editor Gudrun Paletta for the valuable comments and prompt response to our queries and Maria-Inti Metzendorf for developing the search strategies.

REFERENCES

References to studies included in this review

Bergenfelz 2002 *{published data only}*

* Bergenfelz A, Lindblom P, Tibblin S, Westerdahl J. Unilateral versus bilateral neck exploration for primary hyperparathyroidism. *Annals of Surgery* 2002;**236**(5):543-51.

Westerdahl J, Bergenfelz A. Unilateral versus bilateral neck exploration for primary hyperparathyroidism: five-year follow-up of a randomized controlled trial. *Annals of Surgery* 2007;**246**(6):976-80.

Bergenfelz 2005 *{published data only}*

Bergenfelz A, Kamigiesser V, Zielke A, Nies C, Rothmund M. Conventional bilateral cervical exploration versus open minimally invasive parathyroidectomy under local anaesthesia for primary hyperparathyroidism. *British Journal of Surgery* 2005;**92**(2):190-7.

Miccoli 1999 *{published data only}*

Miccoli P, Bendinelli C, Berti P, Vignali E, Pinchera A, Marcocci C. Video-assisted versus conventional parathyroidectomy in primary hyperparathyroidism: a prospective randomized study. *Surgery* 1999;**126**(6):1117-22.

Miccoli 2008 *{published data only}*

Miccoli P, Berti P, Materazzi G, Ambrosini CE, Fregoli L, Donatini G. Endoscopic bilateral neck exploration versus quick intraoperative parathormone assay (qPTHa) during endoscopic parathyroidectomy: a prospective randomized trial. *Surgical Endoscopy* 2008;**22**(2):398-400.

Slepavicius 2008 *{published data only}*

Slepavicius A, Beisa V, Janusonis V, Strupas K. Focused versus conventional parathyroidectomy for primary hyperparathyroidism: a prospective, randomized, blinded trial. *Langenbeck's Archives of Surgery* 2008;**393**(5):659-66.

References to studies excluded from this review

Aarum 2007 *{published data only}*

Aarum S, Nordenström J, Reihnér E, Zedenius J, Jacobsson H, Danielsson R, et al. Operation for primary hyperparathyroidism: the new versus the old order. A randomised controlled trial of preoperative localisation. *Scandinavian Journal of Surgery* 2007;**96**(1):26-30.

Adler 2008 *{published data only}*

Adler JT, Sippel RS, Chen H. The influence of surgical approach on quality of life after parathyroid surgery. *Annals of Surgical Oncology* 2008;**15**(6):1559-65.

Baliski 2008 *{published data only}*

Baliski C, Nosyk B, Melck A, Bugis S, Rosenberg F, Anis AH. The cost-effectiveness of three strategies for the surgical treatment of symptomatic primary hyperparathyroidism. *Annals of Surgical Oncology* 2008;**15**(10):2653-60.

Barczyński 2006 *{published data only}*

Barczyński M, Cichoń S, Konturek A, Cichoń W. Minimally invasive video-assisted parathyroidectomy versus open minimally invasive parathyroidectomy for a solitary parathyroid adenoma: a prospective, randomized, blinded trial. *World Journal of Surgery* 2006;**30**(5):721-31.

Beyer 2007 *{published data only}*

Beyer TD, Solorzano CC, Starr F, Nilubol N, Prinz RA. Parathyroidectomy outcomes according to operative approach. *American Journal of Surgery* 2007;**193**(3):368-72.

Calò 2017 *{published data only}*

Calò PG, Medas F, Loi G, Pisano G, Sorrenti S, Erdas E, et al. Parathyroidectomy for primary hyperparathyroidism in the elderly: experience of a single endocrine surgery center. *Aging Clinical and Experimental Research* 2017;**29**(1):15-21.

Carneiro 2000 *{published data only}*

Carneiro DM, Irvin GL 3rd. Late parathyroid function after successful parathyroidectomy guided by intraoperative hormone assay (QPTH) compared with the standard bilateral neck exploration. *Surgery* 2000;**128**(6):925-9.

Clerici 1992 *{published data only}*

Clerici T, Kull C, Sege D, Rittmann WW. Incidence of recurrence in long-term follow-up of 100 patients after surgery for primary hyperparathyroidism. *Helvetica Chirurgica Acta* 1992;**58**(6):767-70.

Del Rio 2013 *{published data only}*

Del Rio P, Vicente D, Maestroni U, Totaro A, Pattacini GM, Avital I. A comparison of minimally invasive video-assisted parathyroidectomy and traditional parathyroidectomy for parathyroid adenoma. *Journal of Cancer* 2013;**4**(6):458-63.

Ejlsmark-Svensson 2019 *{published data only}*

Ejlsmark-Svensson H, Sikjaer T, Webb SM, Rejnmark L, Rolighed L. Health-related quality of life improves 1 year after parathyroidectomy in primary hyperparathyroidism: a prospective cohort study. *Clinical Endocrinology* 2019;**90**(1):184-91.

Genc 2003 *{published data only}*

Genc H, Morita E, Perrier ND, Miura D, Ituarte P, Duh QY, et al. Differing histologic findings after bilateral and focused parathyroidectomy. *Journal of the American College of Surgeons* 2003;**196**(4):535-40.

Grant 2005 *{published data only}*

Grant CS, Thompson G, Farley D, van Heerden J. Primary hyperparathyroidism surgical management since the introduction of minimally invasive parathyroidectomy: Mayo Clinic experience. *Archives of Surgery* 2005;**140**(5):472-8.

Hansen 2012 *{published data only}*

Hansen S, Hauge EM, Rasmussen L, Jensen JE, Brixen K. Parathyroidectomy improves bone geometry and microarchitecture in female patients with primary

hyperparathyroidism: a one-year prospective controlled study using high-resolution peripheral quantitative computed tomography. *Journal of Bone and Mineral Research* 2012;**27**(5):1150-8.

Henry 2001 *{published data only}*

Henry JF, Raffaelli M, Iacobone M, Volot F. Video-assisted parathyroidectomy via the lateral approach vs conventional surgery in the treatment of sporadic primary hyperparathyroidism: results of a case-control study. *Surgical Endoscopy* 2001;**15**(10):1116-9.

Hessman 2010 *{published data only}*

Hessman O, Westerdahl J, Al-Suliman N, Christiansen P, Hellman P, Bergenfelz A. Randomized clinical trial comparing open with video-assisted minimally invasive parathyroid surgery for primary hyperparathyroidism. *British Journal of Surgery* 2010;**97**(2):177-84.

Lumachi 2003 *{published data only}*

Lumachi F, Ermani M, Basso SM, Camozzi V, Nardi A, Favia G, et al. Short- and long-term changes in bone mineral density of the lumbar spine after parathyroidectomy in patients with primary hyperparathyroidism. *Calcified Tissue International* 2003;**73**(1):44-8.

McGill 2008 *{published data only}*

McGill J, Sturgeon C, Kaplan SP, Chiu B, Kaplan EL, Angelos P. How does the operative strategy for primary hyperparathyroidism impact the findings and cure rate? A comparison of 800 parathyroidectomies. *Journal of the American College of Surgeons* 2008;**207**(2):246-9.

Mekel 2014 *{published data only}*

Mekel M, Gilshtein H, Chapchay K, Bishara B, Krausz MM, Freund HR, et al. Parathyroid surgery in the elderly: should minimally invasive surgery be abandoned? *Annals of Surgical Oncology* 2014;**21**(4):1369-73.

NCT00877981 *{unpublished data only}*

NCT00877981. Open compared to video-assisted minimal-invasive parathyroid surgery in primary hyperparathyroidism. clinicaltrials.gov/ct2/show/NCT00877981 (first received 8 April 2009).

NCT01776502 *{unpublished data only}*

NCT01776502. A prospective, non randomized, open, multicenter trial to evaluate non specific symptoms and quality of life (SF 36) before and after surgery for mild primary hyperparathyroidism. clinicaltrials.gov/ct2/show/NCT01776502 (first received 28 January 2013).

Nehs 2013 *{published data only}*

Nehs MA, Ruan DT, Gawande AA, Moore FD Jr, Cho NL. Bilateral neck exploration decreases operative time compared to minimally invasive parathyroidectomy in patients with discordant imaging. *World Journal of Surgery* 2013;**37**(7):1614-7. [DOI: 10.1007/s00268-013-2007-8]

Neychev 2016 *{published data only}*

Neychev VK, Ghanem M, Blackwood SL, Aragon Han P, Fazeli R, Schneider E, et al. Parathyroid surgery can be safely performed in a community hospital by experienced parathyroid surgeons: a retrospective cohort study. *International Journal of Surgery* 2016;**27**:72-6.

Nilsson 2017 *{published data only}*

Nilsson IL, Norenstedt S, Zedenius J, Pernow Y, Bränström R. Primary hyperparathyroidism, hypercalciuria, and bone recovery after parathyroidectomy. *Surgery* 2017;**162**(2):429-36.

Norenstedt 2009 *{published data only}*

Norenstedt S, Ekbom A, Brandt L, Zedenius J, Nilsson IL. Postoperative mortality in parathyroid surgery in Sweden during five decades: improved outcome despite older patients. *European Journal of Endocrinology* 2009;**160**(2):295-9.

Norman 2015 *{published data only}*

Norman J, Politz D, Lopez J, Boone D, Stojadinovic A. Surgical cure of primary hyperparathyroidism ameliorates gastroesophageal reflux symptoms. *World Journal of Surgery* 2015;**39**(3):706-12.

Pepe 2013 *{published data only}*

Pepe J, Curione M, Morelli S, Varrenti M, Cammarota C, Cilli M, et al. Parathyroidectomy eliminates arrhythmic risk in primary hyperparathyroidism, as evaluated by exercise test. *European Journal of Endocrinology* 2013;**169**(2):255-61.

Russel 2006 *{published data only}*

Russel CF, Dolan SJ, Laird JD. Randomized clinical trial comparing scan-directed unilateral versus bilateral exploration for parathyroidectomy due to solitary adenoma. *British Journal of Surgery* 2006;**93**(4):418-21.

Schneider 2012 *{published data only}*

Schneider DF, Mazeh H, Sippel RS, Chen H. Is minimally invasive parathyroidectomy associated with greater recurrence compared to bilateral exploration? analysis of more than 1,000 cases. *Surgery* 2012;**152**(6):1008-15.

Simonella 2005 *{published data only}*

Simonella G, Massaccesi E, De Marzi C, Staffolani P, Falco A, Morosini P. Minimally invasive surgery versus bilateral neck exploration for primary hyperparathyroidism: controlled prospective study. Role of intraoperative rapid parathyroid hormone assay and radiological preoperative detection of adenomas. *Recenti Progressi in Medicina* 2005;**96**(10):483-7.

Sozio 2005 *{published data only}*

Sozio A, Schietroma M, Franchi L, Mazzotta C, Cappelli S, Amicucci G. Parathyroidectomy: bilateral exploration of the neck vs minimally invasive radioguided treatment. *Minerva Chirurgica* 2005;**60**(2):83-9.

Tolley 2016 *{published data only}*

Tolley N, Garas G, Palazzo F, Prichard A, Chaidas K, Cox J, et al. Long-term prospective evaluation comparing robotic parathyroidectomy with minimally invasive open

parathyroidectomy for primary hyperparathyroidism. *Head & Neck* 2016;**38**(Suppl 1):E300-6.

Udelsman 2002 *{published data only}*

Udelsman R. Six hundred fifty-six consecutive explorations for primary hyperparathyroidism. *Annals of Surgery* 2002;**235**(5):665-70.

Udelsman 2011 *{published data only}*

Udelsman R, Lin Z, Donovan P. The superiority of minimally invasive parathyroidectomy based on 1650 consecutive patients with primary hyperparathyroidism. *Annals of Surgery* 2011;**253**(3):585-91. [DOI: 10.1097/SLA.0b013e318208fed9]

Usta 2015 *{published data only}*

Usta A, Alhan E, Cinel A, Turkyılmaz S, Erem C. A 20-year study on 190 patients with primary hyperparathyroidism in a developing country: Turkey experience. *International Surgery* 2015;**100**:648-55.

References to studies awaiting assessment

Fan 2013 *{published data only}*

Fan Y. Minimally invasive video-assisted parathyroidectomy compared with the conventional open operation for primary hyperparathyroidism: a randomized controlled trial. 83rd Annual Meeting of the American Thyroid Association; 2013 Oct 16-20; San Juan, Puerto Rico; A65.

References to ongoing studies

NCT00006329 *{unpublished data only}*

NCT00006329. Comparison of two methods of parathyroidectomy for primary hyperparathyroidism. clinicaltrials.gov/ct2/show/NCT00006329 (first received 5 October 2000).

Additional references

Agrawal 2014

Agrawal R, Mayilvaganan S. Primary hyperparathyroidism: an analysis of failure of parathyroidectomy. *World Journal of Surgery* 2014;**38**:2481-2.

Akl 2013

Akl EA, Johnston BC, Alonso-Coello P, Neumann I, Ebrahim, Briel M, et al. Addressing dichotomous data for participants excluded from trial analysis: a guide for systematic reviewers. *PloS One* 2013;**8**(2):1-7. [DOI: 10.1371/journal.pone.0057132]

Akl 2015

Akl EA, Kahale LA, Agoritsas T, Brignardello-Petersen R, Busse JW, Carrasco-Labra A, et al. Handling trial participants with missing outcome data when conducting a meta-analysis: a systematic survey of proposed approaches. *Systematic Reviews* 2015;**23**(4):98. [DOI: 10.1186/s13643-015-0083-6]

Alesina 2010

Alesina PF, Singaporewalla RM, Walz MK. Video-assisted bilateral neck exploration in patients with primary hyperparathyroidism and failed localization studies. *World Journal of Surgery* 2010;**34**(10):2344-9. [DOI: 10.1007/s00268-010-0700-4]

Allendorf 2007

Allendorf J, DiGorgi M, Spanknebel K, Inabnet W, Chabot J, Logerfo P. 1112 consecutive bilateral neck explorations for primary hyperparathyroidism. *World Journal of Surgery* 2007;**31**(11):2075-80.

Altman 2003

Altman DG, Bland JM. Interaction revisited: the difference between two estimates. *BMJ* 2003;**326**(7382):219. [PMID: 12543843]

Arici 2001

Arici C, Cheah WK, Ituarte PH, Morita E, Lynch TC, Siperstein AE, et al. Can localization studies be used to direct focused parathyroid operations? *Surgery* 2001;**129**(6):720-9.

Babinska 2012

Babinska D, Barczynski M, Stefaniak T, Oseka T, Babinska A, Babinski D, et al. Evaluation of selected cognitive functions before and after surgery for primary hyperparathyroidism. *Langenbeck's Archives of Surgery* 2012;**397**(5):825-31. [DOI: 10.1007/s00423-011-0885-5]

Baim 2008

Baim S, Binkley N, Bilezikian JP, Kendler DL, Hans DB, Lewiecki EM, et al. Official positions of the international society for clinical densitometry and executive summary of the 2007 ISCD position development conference. *Journal of Clinical Densitometry* 2008;**11**(1):75-91. [DOI: 10.1016/j.jocd.2007.12.007]

Barczynski 2009

Barczynski M, Konturek A, Hubalewska-Dydejczyk A, Cichon S, Nowak W. Evaluation of Halle, Miami, Rome, and Vienna intraoperative i PTH assay criteria in guiding minimally invasive parathyroidectomy. *Langenbeck's Archives of Surgery* 2009;**394**(5):843-9.

Bergenfelz 2007

Bergenfelz A, Jansson S, Mårtensson H, Reihnér E, Wallin G, Kristoffersson A, et al. Scandinavian quality register for thyroid and parathyroid surgery: audit of surgery for primary hyperparathyroidism. *Langenbeck's Archives of Surgery* 2007;**392**(4):445-51.

Bilezikian 2000

Bilezikian JP, Meng X, Shi Y, Silverberg SJ. Primary hyperparathyroidism in women: a tale of two cities-New York and Beijing. *International Journal of Fertility and Women's Medicine* 2000;**45**(2):158-65.

Bilezikian 2002

Bilezikian JP, Potts JT Jr, Fuleihan GE, Kleerekoper M, Neer R, Peacock M, et al. Summary statement from a workshop on

asymptomatic primary hyperparathyroidism: a perspective for the 21st century. *Journal of Clinical Endocrinology and Metabolism* 2002;**87**(12):5353-61.

Bilezikian 2014

Bilezikian JP, Brandi ML, Eastell R, Silverberg SJ, Udelsman R, Marcocci C, et al. Guidelines for the management of asymptomatic primary hyperparathyroidism: summary statement from the fourth international workshop. *Journal of Clinical Endocrinology and Metabolism* 2014;**99**(10):3561-9. [DOI: 10.1210/jc.2014-1413]

Boggs 1999

Boggs JE, Irvin GL 3rd, Carneiro DM, Molinari AS. The evolution of parathyroidectomy failures. *Surgery* 1999;**126**(6):998-1002.

Borenstein 2017a

Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods* 2017;**8**(1):5-18.

Borenstein 2017b

Borenstein M. Prediction intervals. www.meta-analysis.com/prediction (accessed 3 July 2017).

Boutron 2014

Boutron I, Altman DG, Hopewell S, Vera-Badillo F, Tannock I, Ravaud P. Impact of spin in the abstracts of articles reporting results of randomized controlled trials in the field of cancer: the SPIIN randomized controlled trial. *Journal of Clinical Oncology* 2014;**32**(36):4120-6. [DOI: 10.1200/JCO.2014.56.7503]

Burkey 2001

Burkey SH, van Heerden JA, Thompson GB, Grant CS, Schleck CD, Farley DR. Re-exploration for symptomatic hematomas after cervical exploration. *Surgery* 2001;**130**(6):914-20.

Carneiro 2003

Carneiro DM, Solorzano CC, Nader MC, Ramirez M, Irvin GL 3rd. Comparison of intraoperative i PTH assay (QPTH) criteria in guiding parathyroidectomy: which criterion is the most accurate? *Surgery* 2003;**134**(6):973-9.

Carty 1997

Carty SE, Worsey J, Virji MA, Brown ML, Watson CG. Concise parathyroidectomy: the impact of preoperative SPECT 99mTc sestamibi scanning and intraoperative quick parathormone assay. *Surgery* 1997;**122**(6):1107-14.

Carty 2002

Carty SE, Roberts MM, Virji MA, Haywood L, Yim JH. Elevated serum parathormone level after "concise parathyroidectomy" for primary sporadic hyperparathyroidism. *Surgery* 2002;**132**(6):1086-92.

Carty 2004

Carty SE. Prevention and management of complications in parathyroid surgery. *Otolaryngologic Clinics of North America* 2004;**37**(4):897-907.

Chen 1999

Chen H, Sokoll LJ, Udelsman R. Outpatient minimally invasive parathyroidectomy: a combination of sestamibi-SPECT localization, cervical block anesthesia, and intraoperative parathyroid hormone assay. *Surgery* 1999;**126**(6):1016-21.

Chen 2003

Chen Q, Kaji H, Lu M-F, Nomura R, Sowa H, Yamauchi M, et al. Effects of an excess and a deficiency of endogenous parathyroid hormone on volumetric bone mineral density and bone geometry determined by peripheral quantitative computed tomography in female subjects. *Journal of Clinical Endocrinology and Metabolism* 2003;**88**(10):4655-8.

Coker 2005

Coker LH, Rorie K, Cantley L, Kirkland K, Stump D, Burbank N, et al. Primary hyperparathyroidism, cognition, and health-related quality of life. *Annals of Surgery* 2005;**242**(5):642-50.

Corbett 2014

Corbett MS, Higgins JP, Woolacott NF. Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Research Synthesis Methods* 2014;**5**(1):79-85. [DOI: 10.1002/jrsm.1090]

Costa-Guda 2014

Costa-Guda J, Arnold A. Genetic and epigenetic changes in sporadic endocrine tumors: parathyroid tumors. *Molecular Cell Endocrinology* 2014;**386(1-2)**:46-54. [DOI: 10.1016/j.mce.2013.09.005]

Deeks 2017

Deeks JJ, Higgins JP, Altman DG, editor(s), on behalf of the Cochrane Statistical Methods Group. Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, editor(s). Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Dijkstra 2002

Dijkstra B, Healy C, Kelly LM, McDermott EW, Hill AD, O'Higgins N. Parathyroid localisation – current practice. *Journal of the Royal College of Surgeons of Edinburgh* 2002;**47**(4):599-607.

Eastell 2014

Eastell R, Brandi ML, Costa AG, D'Amour P, Shoback DM, Thakker RV. Diagnosis of asymptomatic primary hyperparathyroidism: proceedings of the fourth international workshop. *Journal of Clinical Endocrinology and Metabolism* 2014;**99**(10):3570-9.

Ebrahim 2013

Ebrahim S, Akl EA, Mustafa RA, Sun X, Walter SD, Heels-Ansdell D. Addressing continuous data for participants excluded from trial analysis: a guide for systematic reviewers. *Journal of Clinical Epidemiology* 2013;**66**(9):1014-21. [DOI: 10.1016/j.jclinepi.2013.03.014]

Fitzpatrick 2008

Fitzpatrick LA, Bilezikian JP, Silverberg SJ. Parathyroid hormone and the cardiovascular system. *Current Osteoporosis Reports* 2008;**6**(2):77-83.

Gracie 2012

Gracie D, Hussain SS. Use of minimally invasive parathyroidectomy techniques in sporadic primary hyperparathyroidism: systematic review. *Journal of Laryngology and Otology* 2012;**126**(3):221-7.

GRADEpro GDT 2014 [Computer program]

GRADE Working Group, McMaster University GRADEpro GDT. Version accessed 23 February 2017. Hamilton (ON): GRADE Working Group, McMaster University, 2014. Available at gradepro.org.

Greene 2009

Greene AB, Butler RS, McIntyre S, Barbosa GF, Mitchell J, Berber E, et al. National trends in parathyroid surgery from 1998 to 2008: a decade of change. *Journal of American College of Surgeons* 2009;**209**(3):332-43. [DOI: 10.1016/j.jamcollsurg.2009.05.029]

Guyatt 2011

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J. GRADE guidelines: 1. Introduction – GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology* 2011;**64**(4):383-94.

Henry 1999

Henry JF, Defechereux T, Gramatica L, De Boissezon C. Minimally invasive videoscopic parathyroidectomy by lateral approach. *Langenbeck's Archives of Surgery* 1999;**384**(3):298-301.

Henry 2008

Henry JF, Sebag F, Cherenko M, Ippolito G, Taieb D, Vaillant J. Endoscopic parathyroidectomy: why and when? *World Journal of Surgery* 2008;**32**(11):2509-15. [DOI: 10.1007/s00268-008-9709-3]

Higgins 2002

Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**(11):1539-58.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analysis. *BMJ (Clinical Research Ed.)* 2003;**327**:557-60.

Higgins 2009

Higgins JP, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 2009;**172**(1):137-59.

Higgins 2011a

Higgins JP, 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**:d5928.

Higgins 2011b

Higgins JP, Deeks JJ, Altman DG. Chapter 16: Special topics in statistics. In: Higgins JP, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org.

Higgins 2017

Higgins JP, Altman DG, Sterne JA. Chapter 8: Assessing risk of bias in included studies. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, editor(s), Cochrane Handbook for Systematic Reviews of Interventions Version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Hindié 2015

Hindié E, Zanotti-Fregonara P, Tabarin A, Rubello D, Morelec I, Wagner T, et al. The role of radionuclide imaging in the surgical management of primary hyperparathyroidism. *Journal of Nuclear Medicine* 2015;**56**(5):737-44.

Hozo 2005

Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology* 2005;**5**:13. [DOI: 10.1186/1471-2288-5-13]

Hróbjartsson 2013

Hróbjartsson A, Thomsen AS, Emanuelsson F, Tendal B, Hilden J, Boutron I. Observer bias in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both blinded and nonblinded assessors. *Canadian Medical Association Journal* 2013;**185**(4):E201-11.

Irvin 1991

Irvin GL 3rd, Dembrow VD, Prudhomme DL. Operative monitoring of parathyroid gland hyperfunction. *American Journal of Surgery* 1991;**162**(4):299-302.

Irvin 1993

Irvin GL 3rd, Dembrow VD, Prudhomme DL. Clinical usefulness of an intraoperative "quick parathyroid hormone" assay. *Surgery* 1993;**114**(6):1013-9.

Irvin 1994

Irvin GL 3rd, Prudhomme DL, Deriso GT, Sfakianakis G, Chandarlapaty SK. A new approach to parathyroidectomy. *Annals of Surgery* 1994;**219**(5):574-9.

Irvin 2004

Irvin GL 3rd, Solorzano CC, Carneiro DM. Quick intraoperative parathyroid hormone assay: surgical adjunct to allow limited parathyroidectomy, improve success rate, and predict outcome. *World Journal of Surgery* 2004;**28**(12):1287-92.

Joborn 1988

Joborn C, Hetta J, Johansson H, Rastad J, Agren H, Akerström G, Ljunghall S. Psychiatric morbidity in primary hyperparathyroidism. *World Journal of Surgery* 1988;**12**(4):476.

Jones 2015

Jones CW, Keil LG, Holland WC, Caughey MC, Platts-Mills TF. Comparison of registered and published outcomes in randomized controlled trials: a systematic review. *BMC Medicine* 2015;**13**:282. [DOI: 10.1186/s12916-015-0520-3]

Kaplan 1992

Kaplan EL, Yashiro T, Salti G. Primary hyperparathyroidism in the 1990s. Choice of surgical procedures for this disease. *Annals of Surgery* 1992;**215**(4):300-17.

Khosla 2002

Khosla S, Melton J. Fracture risk in primary hyperparathyroidism. *Journal of Bone and Mineral Research* 2002;**17 Suppl 2**:N103-7.

Kirkham 2010

Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ (Clinical Research Ed.)* 2010;**340**:c365. [DOI: 10.1136/bmj.c365]

Kunstman 2012

Kunstman JW, Udelsman R. Superiority of minimally invasive parathyroidectomy. *Advances in Surgery* 2012;**46**(1):171-89.

Kuntsman 2013

Kuntsman JW, Kirsch JD, Mahajan A. Clinical review: parathyroid localization and implications for clinical management. *Journal of Clinical Endocrinology and Metabolism* 2013;**98**:902-12.

Liberati 2009

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP. The PRISMA statement for reporting systematic and meta-analyses of studies that evaluate interventions: explanation and elaboration. *PLoS Medicine* 2009;**6**(7):1-28. [DOI: 10.1371/journal.pmed.1000100]

Lo 1999

Lo GP. Bilateral neck exploration for parathyroidectomy under local anesthesia: a viable technique for patients with coexisting thyroid disease with or without sestamibi scanning. *Surgery* 1999;**126**(6):1011-5.

Lowney 2000

Lowney JK, Weber B, Johnson S, Doherty GM. Minimal incision parathyroidectomy: cure, cosmesis, and cost. *World Journal of Surgery* 2000;**24**(11):1442-5.

Mathieu 2009

Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009;**302**:977-84.

McAllion 1989

McAllion SJ, Paterson CR. Psychiatric morbidity in primary hyperparathyroidism. *Postgraduate Medical Journal* 1989;**65**(767):628.

Meader 2014

Meader N, King K, Llewellyn A, Norman G, Brown J, Rodgers M. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Systematic Reviews* 2014;**3**:82.

Mehrabi 2012

Mehrabi BM, Jangjoo A, Afzal AM, Ebrahimzadeh M, Aliakbarian M. Transient and permanent hypoparathyroidism following thyroidectomy. *Minerva Chirurgica* 2012;**67**(5):433. [PMID: 23232482]

Miccoli 2011

Miccoli P, Materazzi G, Bagglani A, Miccoli M. Minimal invasive video-assisted surgery of the thyroid and parathyroid glands: a 2011 update. *Journal of Endocrinological Investigation* 2011;**34**:473-80.

Mishra 2001

Mishra SK, Agarwal G, Kar DK, Gupta SK, Mithal A, Rastad. Unique clinical characteristics of primary hyperparathyroidism in India. *British Journal of Surgery* 2001;**88**(5):708.

Ning 2009

Ning L, Sippel R, Schaefer S, Chen H. What is the clinical significance of an elevated parathyroid hormone level after curative surgery for primary hyperparathyroidism? *Annals of Surgery* 2009;**249**(3):469-72.

Norman 2012

Norman J, Lopez J, Politz D. Abandoning unilateral parathyroidectomy: why we reversed our position after 15,000 parathyroid operations. *Journal of the American College of Surgeons* 2012;**214**(3):260-9.

Parfitt 1991

Parfitt AM, Rao DS, Kleerekoper M. Asymptomatic primary hyperparathyroidism discovered by multichannel biochemical screening: clinical course and considerations bearing on the need for surgical intervention. *Journal of Bone and Mineral Research* 1991;**6 Suppl 2**:S97-101.

Pyram 2011

Pyram R, Mahajan G, Gliwa A. Primary hyperparathyroidism: skeletal and non-skeletal effects, diagnosis and management. *Maturitas* 2011;**70**(3):246-55.

Ramakant 2012

Ramakant P, Paul MJ, Paul TV, Rao SD, Abraham DT, Odumah OJ. Surgery versus surveillance for asymptomatic (mild) primary hyperparathyroidism in adults 50 years or older. *Cochrane Database of Systematic Reviews* 2012, Issue 9. Art. No: CD010093. [DOI: 10.1002/14651858.CD010093]

Review Manager 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Riley 2011

Riley RD, Higgins JP, Deeks JJ. Interpretation of random effects meta-analyses. *BMJ (Clinical Research Ed.)* 2011;**342**:d549.

Rubin 2005

Rubin MR, Maurer MS, McMahon DJ, Bilezikian JP, Silverberg SJ. Arterial stiffness in mild primary hyperparathyroidism. *Journal of Clinical Endocrinology and Metabolism* 2005;**85**:3515-9.

Rubin 2008

Rubin MR, Bilezikian J, McMahon DJ, Jacobs T, Shane E, Siris E, et al. The natural history of primary hyperparathyroidism with or without parathyroid surgery after 15 years. *Journal of Clinical Endocrinology and Metabolism* 2008;**93**(9):3462–70.

Ruda 2005

Ruda JM, Hollenbeak CS, Stack BC. A systematic review of the diagnosis and treatment of primary hyperparathyroidism from 1995 to 2003. A systematic review of the diagnosis and treatment of primary hyperparathyroidism from 1995 to 2003. *Otolaryngology and Head and Neck Surgery* 2005;**132**:359.

Rulli 2007

Rulli F, Galata G, Pompeo E, Farinon AM. A camera handler for Miccoli's minimally invasive video-assisted thyroidectomy and parathyroidectomy procedures. *Surgical Endoscopy and Other Interventional Techniques* 2007;**21**(6):1017-9.

Ryan 1997

Ryan JA, Eisenberg B, Pado KM, Lee FE. Efficacy of selective unilateral exploration in hyperparathyroidism based on localization tests. *Archives of Surgery* 1997;**132**(8):886-90.

Ryan 2004

Ryan JA Jr, Lee FT. Maximizing outcomes while minimizing exploration in hyperparathyroidism using localization tests. *Archives of Surgery* 2004;**139**(8):838-42.

Salti 1992

Salti GI, Fedorak I, Yashiro T, Fulton N, Hara H, Yousefzadeh D, et al. Continuing evolution in the operative management of primary hyperparathyroidism. *Archives of Surgery* 1992;**127**(7):831-7.

Schünemann 2017

Schünemann HJ, Oxman AD, Higgins JP, Vist GE, Glasziou P, Akl E, et al, on behalf of the Cochrane GRADEing Methods Group and the Cochrane Statistical Methods Group. Chapter 11: Completing 'Summary of findings' tables and grading the confidence in or quality of the evidence. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, editor(s), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Shea 2007

Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology* 2007;**7**:10. [DOI: 10.1186/1471-2288-7-10] [www.biomedcentral.com/1471-2288/7/10]

Silva 2013

Silva BC, Boutroy S, Zhang C, McMahon DJ, Zhou B, Wang J, et al. Trabecular bone score (TBS) – a novel method to evaluate bone microarchitectural texture in patients with primary hyperparathyroidism. *Journal of Clinical Endocrinology and Metabolism* 2013;**98**(5):1963-70.

Silverberg 2014

Silverberg SJ, Clarke BL, Peacock M, Bandeira F, Boutroy S, Cusano NE, et al. Current issues in the presentation of asymptomatic primary hyperparathyroidism: proceedings of the fourth international workshop. *Journal of Clinical Endocrinology and Metabolism* 2014;**99**(10):3580-94.

Starup-Linde 2012

Starup-Linde J, Waldhauer E, Rolighed L, Mosekilde L, Vestergaard P. Renal stones and calcifications in patients with primary hyperparathyroidism: associations with biochemical variables. *European Journal of Endocrinology* 2012;**166**(6):1093-100.

Stein 2013

Stein EM, Silva BC, Boutroy S, Zhou B, Wang J, Udesky J, et al. Primary hyperparathyroidism is associated with abnormal cortical and trabecular microstructure and reduced bone stiffness in postmenopausal women. *Journal of Bone and Mineral Research* 2013;**28**(5):1029-40.

Sterne 2011

Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ (Clinical Research Ed.)* 2011;**343**:d4002.

Sterne 2017

Sterne JA, Egger M, Moher D, Boutron I. Chapter 10: Addressing reporting biases. In: Higgins JP, Churchill R, Chandler J, Cumpston MS, (editor(s), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

Suh 2008

Suh JM, Cronan JJ, Monchik JM. Primary hyperparathyroidism: is there an increased prevalence of renal stone disease? *American Journal of Roentgenology* 2008;**191**(3):908-11.

Udelsman 2004

Udelsman R, Donovan PI. Open minimally invasive parathyroid surgery. *World Journal of Surgery* 2004;**28**(12):1224-6.

Udelsman 2014

Udelsman R, Donovan P, Shaw C. Cure predictability during parathyroidectomy. *World Journal of Surgery* 2014;**38**:525-33.

Vu 2013

Vu TD, Wang XF, Wang Q, Cusano NE, Irani D, Silva BC, et al. New insights into the effects of primary hyperparathyroidism

on the cortical and trabecular compartments of bone. *Bone* 2013;**55**(1):57-63.

Walker 2012

Walker MD, Rundek T, Homma S, Ditullio M, Iwata S, Lee JA, et al. Effect of parathyroidectomy on subclinical cardiovascular disease in mild primary hyperparathyroidism. *European Journal of Endocrinology* 2012;**167**(2):277-85.

Welk 1987

Welk RA, Alix DR. A community hospital experience with total parathyroidectomy and autotransplantation for renal hyperparathyroidism. *American Journal of Surgery* 1987;**53**(11):622-7.

Welsh 1984

Welsh CL, Taylor GW, Cattell WR, Baker LR. Parathyroid surgery in chronic renal failure: subtotal parathyroidectomy or autotransplantation? *British Journal of Surgery* 1984;**71**(8):591-2.

Wermers 2006

Wermers RA, Khosla S, Atkinson EJ, Achenbach SJ, Oberg AL, Grant CS, et al. Incidence of primary hyperparathyroidism in Rochester, Minnesota, 1993–2001: an update on the changing epidemiology of the disease. *Journal of Bone and Mineral Research* 2006;**21**:171-7.

Westin 2009

Westin G, Björklund P, Akerström G. Molecular genetics of parathyroid disease. *World Journal of Surgery* 2009;**33**(11):2224-33.

White 1986

White KS, Wilkinson CA, Nixon WP Jr, Almkuist RD. Total parathyroidectomy and forearm autotransplantation for chronic renal failure: a useful tool for the general surgeon in private practice. *Southern Medical Journal* 1986;**79**(7):844-6.

Wood 2008

Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ (Clinical Research Ed.)* 2008;**336**(7644):601-5.

References to other published versions of this review

Kreidieh 2013

Kreidieh OI, Ahmadieh H, Akl EA, El-Hajj Fuleihan G. Minimally invasive parathyroidectomy guided by intraoperative parathyroid hormone monitoring (IOPTH) and preoperative imaging versus bilateral neck exploration for primary hyperparathyroidism in adults. *Cochrane Database of Systematic Reviews* 2013, Issue 10. Art. No: CD010787. [DOI: 10.1002/14651858.CD010787]

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Bergenfelz 2002

Study characteristics	
Methods	**Study design:** parallel-group randomised controlled study **Setting:** hospital **Country where study was performed:** Sweden **Number of study centres:** 1 (Lund University Hospital, Sweden) 91 participants who presented for their first-time surgery for PHPT were enrolled in the study after written and oral information was given. Preoperative symptoms and signs were recorded at first and then randomised to either MIP or BNE group.
Participants	**Inclusion criteria:** people with the biochemical diagnosis of PHPT **Exclusion criteria:** family history of PHPT (MEN 1, MEN 2, hereditary PHPT), previous neck surgery, other planned operations during the surgical procedure (including thyroid surgery), pregnancy and breast-feeding, emergency operation due to hypercalcaemic crisis and inability to understand information or to comply with scheduled follow-up. **Diagnostic criteria:** preoperative biochemical diagnosis of PHPT was based on a serum calcium level > 2.60 mmol/L (10.4 mg/dL) and serum PTH level > 3.5 pmol/L (32.9 pg/mL), with serum creatinine level < 200 mol/L (2.26 mg/dL).

Bergenfelz 2002 *(Continued)*

	Age group: adults **Gender distribution:** women and men
Interventions	**Intervention:** after the randomisation process, participants in the unilateral group underwent preoperative sestamibi subtraction scintigraphy; no localisation procedure was performed in the bilateral group. In the MIP group, surgery was started on the same side as indicated by the preoperative scintigram. In cases when no enlarged parathyroid gland was visualised on the scintigram, then the left side was explored first. Then upon finding the first enlarged parathyroid gland, blood samples were taken for intraoperative measurement of PTH before and at 5 and 15 minutes after gland excision. Surgery was only terminated when the PTH levels declined > 50% within 5 minutes or > 60% after 15 minutes. In the BNE group, surgery was always started on the left side and comprehensive BNE was performed. There was an attempt to visualise the 4 parathyroid glands. The enlarged parathyroid glands were then removed and taken for frozen section. The decision to terminate surgery was based on the gross morphology of excised parathyroid glands in addition to the results of the frozen section. Normal parathyroid glands were not biopsied and intraoperative PTH levels were not monitored. **Comparator:** BNE **Duration of surgery:** reported as operating time **Length of hospital stay:** not reported **Duration of follow-up:** Bergenfelz study was followed up by Westerdahl and Bergenfelz (2007) after 5 years. Follow-up was performed after 6 weeks, 1 year and 5 years postoperatively. **Run-in period:** follow-up was days 1–3 postoperatively, and 6 weeks after surgery **Extension period:** follow-up study was published reporting 1-year and 5-year follow-up data
Outcomes	**Reported outcome(s) in full text of publication** Quote: "The primary end-point was the use of postoperative medication for hypocalcaemia symptoms." "The secondary outcome measures were symptomatic hypocalcaemia, defined as serum calcium < 2.00 mmol/L, persistent HPT, complications, operative time, and cost." Serum levels of calcium (severe hypocalcaemia defined as serum calcium 2.00 mmol/L (8.00 mg/dL): at 6 weeks, 1 and 5 years Postoperative medication for hypocalcaemia during the first 4 postoperative days: participants informed to medicate with calcium up to 3 g/24 hours whenever symptomatic. Medication intake was recorded. Operative time: not clearly defined Cost: calculated from official in-hospital charges for services performed by different departments Biochemistries, including serum levels of calcium, alkaline phosphatase, phosphate, creatinine, 25-hydroxycholecalciferol and PTH, were analysed at 6 weeks, 1 year and 5 years.
Study details	**Trial identifier:** not reported The study was not terminated early
Publication details	**Language of Publication:** English **Funding:** no sources of funding stated **Publication status:** peer-reviewed journal
Stated aim for study	**Quote:** "We focused on the impact of surgical strategy on early postoperative hypocalcaemia."

Bergenfelz 2002 *(Continued)*

Notes	Follow-up study published by Westerdahl and Bergenfelz (2007) reported data after 1 year and 5 years postoperatively, with biochemistry and with registration of complications.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	**Comment:** author replied by email and confirmed randomisation through block randomisation by a computer software.
Allocation concealment (selection bias)	Low risk	**Comment:** author replied by email and confirmed use of sequentially numbered, concealed, opaque envelopes.
Blinding of participants and personnel (performance bias) all-cause mortality	Low risk	**Comment:** author replied by email indicating that participants were blinded to study group. Data on mortality were provided in the follow-up study at 5 years. The risk of bias was judged as low, since blinding of participants was judged to have no effect on the mortality rate that was reported at 5 years.
Blinding of participants and personnel (performance bias) conversion rate	High risk	**Comment:** although participants were blinded, as per email reply from the author, conversion rate is dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) duration of surgery	High risk	**Comment:** although participants were blinded, as per email reply from the author, surgery duration was in part dependent on conversion rate, and therefore, surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) length of hospital stay	Unclear risk	**Quote:** "After the first postoperative day, the patients were kept in the patient hotel of the hospital and made individual visits on postoperative days 2 to 4."
Blinding of participants and personnel (performance bias) postoperative increase in PTH with eucalcaemia	Low risk	**Comment:** author replied by email indicating that participants were blinded to study group, and therefore, could not have intervened with any supplementation that may affect PTH elevations. Therefore, we assessed the risk of bias as low.
Blinding of participants and personnel (performance bias) socioeconomic effects	High risk	**Comment:** although participants were blinded, as per email reply from the author, socioeconomic cost may be affected by the number of tests requested by the surgeon, who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) success rate	High risk	**Comment:** although participants were blinded, as per email reply from the author, success rate was dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) specific adverse events	Low risk	**Comment:** participants were blinded as per author email reply. Since this outcome was dependent on participants, therefore, we assessed the risk of bias as low.
Blinding of participants and personnel (performance bias)	Low risk	**Comment:** participants were blinded as per author email reply. Since this outcome was dependent on participants, therefore, we assessed the risk of bias as low.

Bergenfelz 2002 *(Continued)*

perioperative adverse events		
Blinding of outcome assessment (detection bias) all-cause mortality	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) conversion rate	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) duration of surgery	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) length of hospital stay	Unclear risk	**Quote:** "After the first postoperative day, the patients were kept in the patient hotel of the hospital and made individual visits on postoperative days 2 to 4."
Blinding of outcome assessment (detection bias) postoperative increase in PTH with eucalcaemia	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) socioeconomic effects	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) success rate	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) specific adverse events	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) perioperative adverse events	Low risk	**Comment:** author replied by email indicating that outcome assessors were blinded to study group. Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) all-cause mortality	Unclear risk	**Comment:** 38/47 and 33/44 randomised participants with missing outcome data at 5-year follow-up; there were 16 deaths and 4 dropouts, but the author did not specify numbers by assigned group.
Incomplete outcome data (attrition bias) conversion rate	Low risk	**Comment:** conversion rate is assessed at the time of surgery. There were no missing outcome data for conversion rate. Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) duration of surgery	Unclear risk	**Comment:** although the authors provided mean time for operation for both unilateral and bilateral groups, it is unclear whether there were any missing data for that outcome. Therefore, we assessed the risk of bias as unclear.
Incomplete outcome data (attrition bias) postoperative increase in PTH with eucalcaemia	Unclear risk	**Comment:** this outcome was not clearly assessed by treatment arm. Therefore, we assessed the risk of bias as unclear.

Bergenfelz 2002 *(Continued)*

Incomplete outcome data (attrition bias) socioeconomic effects	Unclear risk	**Comment:** although the authors provided data for costs for both unilateral and bilateral groups, it was unclear whether there were any missing data for that outcome. Therefore, we assessed the risk of bias as unclear.
Incomplete outcome data (attrition bias) success rate	Unclear risk	**Comment:** there were 47 participants in MIP group and 44 in BNE group at study entry, 45 in MIP group and 43 in BNE group at 1 year, and 38 in MIP group and 33 in BNE group at 5 years. The authors provided insufficient information regarding missing data and did not clarify how missing data were handled. The authors did not clearly define success rate, neither did they directly compare success rate between the treatment groups at 6 weeks (Bergenfelz 2002), and at 1 and 5 years (Westerdahl 2007). At 6 weeks, the authors reported that 2 participants had persistent HPT in the unilateral group and 1 in bilateral group (Bergenfelz 2002). This outcome was not clearly assessed by treatment arm at 1 and 5 years (Westerdahl 2007). At 5 years, 6 participants had persistent (3) or recurrent (3) HPT; 4 participants in the unilateral group (3 of these were bilaterally explored) and 2 in the bilateral group. Therefore, we assessed the risk of bias as unclear.
Incomplete outcome data (attrition bias) specific adverse events	Unclear risk	**Comment:** there were 47 participants in MIP group and 44 in BNE group at study entry, 45 in MIP group and 43 in BNE group at 1 year, and 38 in MIP group and 33 in BNE group at 5 years. Hypocalcaemia symptoms at 6 weeks were reported in 19/39 participants (49%) in BNE group and 12/43 in MIP group (28%). The proportion of participants with missing data were 5/43 (11.6%) in BNE group and 1/44 (2%) in MIP group, with clear disparate attrition rates between the groups and the authors did not clarify how missing data were handled. Follow-up rates were 97% (88/91) at 1 year and 78% (71/91) at 5 years; 38/47 in MIP group and 33/44 in BNE group. No data were provided on adverse events at 1 year. Of the 20 participants unavailable at 5 years, 16 had died and 4 refused further investigation. Adverse events were not broken down by treatment group. Therefore, we assessed the risk of bias as unclear.
Incomplete outcome data (attrition bias) perioperative adverse events	Unclear risk	**Comment:** there were 47 participants in MIP group and 44 in BNE group at study entry, but on the first postoperative day, hypocalcaemia symptoms were reported in 19/39 participants in bilateral group (49%) and 12/43 (28%) in unilateral group. The proportion of participants with missing data were 4/43 (11.6%) in BNE group and 1/44 (2%) in MIP group, with clear disparate attrition rates between the groups and the authors did not clarify how missing data were handled. Therefore, we assessed the risk of bias as unclear.
Selective reporting (reporting bias)	High risk	**Comment:** protocol was not identified. For the secondary outcomes, severe hypocalcaemia and persistent HPT, the study stated that outcome was analysed but reported no results. See also ORBIT (Appendix 8).
Other bias	Unclear risk	**Comment:** surgeon familiarity and experience with each surgical technique was not detailed. There was no mentioning of the sources of funding either. The presence of conflict of interest could not be excluded as a result of these 2 factors. Therefore, we assessed the risk of bias as unclear.

Bergenfelz 2005

Study characteristics	
Methods	**Study design:** parallel-group randomised controlled study **Setting:** hospital

Bergenfelz 2005 *(Continued)*

	Country where study was performed: Germany **Number of study centres:** 1
Participants	Participants presenting for their first-time parathyroid surgery at the Faculty of Medicine, Philipps-University Marburgin were consented for the possibility of randomisation. **Inclusion criteria:** people with biochemically confirmed PHPT who had not undergone previous surgery and who had solitary parathyroid adenoma on imaging. **Exclusion criteria:** hereditary PHPT (MEN 1 and 2, non-MEN-related familial HPT), suspicion of involvement of multiple parathyroid glands on sestamibi scanning, previous neck exploration for thyroid disorders, anticipated or planned simultaneous thyroid operations, allergy to drugs used for local anaesthesia, people who could not fully comprehend the information given or who rejected confirmation to participate, aged < 18 years, people with hypercalcaemic crisis and high-risk people (American Society of Anesthesiologists grade IV). **Diagnostic criteria:** biochemically confirmed PHPT participants were included for randomisation. **Age group:** adults **Gender distribution:** women and men
Interventions	**Intervention:** the MIP procedure was an open targeted operation with the aim of parathyroid adenoma excision. Surgery was started on the same side as indicated by the preoperative scintigram. Then the parathyroid adenoma was localised, dissected and sent for frozen-section analysis. A decrease in the PTH level of > 50% from baseline values after 5 minutes, or of > 60% after 15 minutes, led to procedure termination. The BNE procedure was started with a short Kocher incision. Surgery was always started on the left side and comprehensive BNE was performed. The enlarged parathyroid glands were then removed and taken for frozen section. The wound was closed, but the participant remained under anaesthesia until the results of the frozen section examination had been received. Intraoperative PTH levels were not monitored. **Comparator:** BNE **Duration of surgery:** reported as operating time **Length of hospital stay:** not reported **Duration of follow-up:** first 4 days after surgery, 1 and 6 months after surgery **Run-in period:** after randomisation, participants underwent sestamibi scintigraphy imaging for localisation of parathyroid adenomas, and only participants who were found to have a single enlarged parathyroid gland were eligible for inclusion in the study. These participants were again asked to reaffirm participation. Follow-up was carried out at 1 and 6 months. **Extension period:** there was no extension period after the decided follow-up date.
Outcomes	**Reported outcome(s) in full text of publication** Serum levels of total calcium during the first 4 days after surgery served as the primary endpoint. Secondary outcomes were number of conversions from MIP to BNE, operating time (skin–skin), complications (recurrent nerve palsy, wound infection, haematoma) and serum level of calcium, oral calcium and vitamin D supplementation recorded at 1 and 6 months after surgery
Study details	**Trial identifier:** not reported The study was not terminated early
Publication details	**Language of Publication:** English

Bergenfelz 2005 *(Continued)*

Funding: no sources of funding were stated

Publication status: peer-reviewed journal

Stated aim for study	**Quote:** "The aim of the study was to document whether or not MIP performed under local anaesthesia would reduce the frequency and severity of postoperative hypocalcaemia in comparison with the standard approach."
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	**Comment:** author replied by email and confirmed randomisation was through block randomisation by computer software. However, this randomisation was not balanced. Therefore, we assessed the risk of bias as high.
Allocation concealment (selection bias)	Low risk	**Comment:** author replied by email and confirmed the use of sequentially numbered, concealed, opaque envelopes.
Blinding of participants and personnel (performance bias) conversion rate	High risk	**Comment:** author replied by email confirming that participants and personnel were not blinded to the treatment group due to the use of local anaesthesia. This outcome was investigator assessed. The study protocol defined conversion from MIP to BNE in the following situations: intraoperative demonstration of 2 normal parathyroid glands on the side where the scan had suggested the adenoma; inadequate decrease in PTH concentration after adenoma excision; no confirmation of parathyroid tissue by frozen-section analysis; and intraoperative suspicion of multiple gland disease. Conversion was also allowed for safety reasons and the participant's well-being, for example when there was a technical problem or the participant felt uncomfortable during the procedure. **Comment:** conversion rate is dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) cosmetic satisfaction	High risk	**Comment:** author replied by email confirming that participants and personnel were not blinded to the treatment group due to the use of local anaesthesia. As this self-reported outcome may have been influenced by cointervention and considering the lack of blinding. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) duration of surgery	High risk	**Comment:** author replied by email confirming that participants and personnel were not blinded to the treatment group due to the use of local anaesthesia. Surgery duration is in part dependent on conversion rate, and therefore, surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) health-related quality of life	High risk	**Comment:** author replied by email confirming that participants were not blinded to the treatment group due to the use of local anaesthesia. As this self-reported outcome may have been influenced by cointervention and considering the lack of blinding. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) length of hospital stay	Unclear risk	**Comment:** the length of hospital stay was fixed to 4 days for both groups, but it was not defined as outcome. Therefore, this outcome could not be assessed and therefore, we assessed the risk of bias as unclear.

Bergenfelz 2005 *(Continued)*

Blinding of participants and personnel (performance bias) success rate	Unclear risk	**Comment:** author replied by email confirming that participants and personnel were not blinded to the treatment group due to the use of local anaesthesia. Success rate was not defined and therefore, we assessed the risk of bias as unclear.
Blinding of participants and personnel (performance bias) specific adverse events	High risk	**Comment:** author replied by email confirming that participants and personnel were not blinded to the treatment group due to the use of local anaesthesia. Included events that were self-reported and others that were reported by the investigators. Therefore, this outcome may have been influenced by cointervention and considering the lack of blinding. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) perioperative adverse events	High risk	**Comment:** author replied by email confirming that participants and personnel were not blinded to the treatment group due to the use of local anaesthesia. Included events that were self-reported and others that were reported by the investigators. Therefore, this outcome may have been influenced by cointervention and considering the lack of blinding. Therefore, we assessed the risk of bias as high.
Blinding of outcome assessment (detection bias) conversion rate	High risk	**Comment:** author replied by email indicating that outcome assessors were not blinded to study group. Therefore, we assessed the risk of bias as high.
Blinding of outcome assessment (detection bias) cosmetic satisfaction	High risk	**Comment:** author replied by email indicating that outcome assessors were not blinded to study group. Therefore, we assessed the risk of bias as high.
Blinding of outcome assessment (detection bias) duration of surgery	High risk	**Comment:** the paper did not specifically state who assessed the duration of surgery. Author replied by email indicating that outcome assessors were not blinded to study group. Therefore, we assessed the risk of bias as high.
Blinding of outcome assessment (detection bias) health-related quality of life	High risk	**Comment:** author replied by email indicating that outcome assessors were not blinded to study group. Therefore, we assessed the risk of bias as high.
Blinding of outcome assessment (detection bias) length of hospital stay	Unclear risk	**Comment:** the length of hospital stay was fixed to 4 days for both groups, but it was not defined as outcome. Therefore, this outcome could not be assessed and we assessed this risk of bias as unclear.
Blinding of outcome assessment (detection bias) success rate	High risk	**Comment:** author replied by email indicating that outcome assessors were not blinded to study group and success rate was not clearly defined. Therefore, we assessed the risk of bias as high.
Blinding of outcome assessment (detection bias) specific adverse events	High risk	**Comment:** author replied by email indicating that outcome (detection bias) assessors were not blinded to study group. Therefore, we assessed the risk of bias as high.
Blinding of outcome assessment (detection bias) perioperative adverse events	High risk	**Comment:** author replied by email indicating that outcome (detection bias) assessors were not blinded to study group. Therefore, we assessed the risk of bias as high.
Incomplete outcome data (attrition bias) conversion rate	Low risk	**Comment:** there was no loss to follow-up and this outcome was reported on all participants. Therefore, we assessed this risk of bias as low.
Incomplete outcome data (attrition bias)	Low risk	**Comment:** there was no loss to follow-up and this outcome was reported on all participants. Therefore, we assessed this risk of bias as low.

Bergenfelz 2005 *(Continued)*

cosmetic satisfaction		
Incomplete outcome data (attrition bias) duration of surgery	Low risk	**Comment:** there was no loss to follow-up and this outcome was reported on all participants. Therefore, we assessed this risk of bias as low.
Incomplete outcome data (attrition bias) health-related quality of life	Low risk	**Comment:** there was no loss to follow-up and this outcome was reported on all participants. Therefore, we assessed this risk of bias as low.
Incomplete outcome data (attrition bias) success rate	Low risk	**Comment:** there was no loss to follow-up and this outcome was reported on all participants. Therefore, we assessed this risk of bias as low.
Incomplete outcome data (attrition bias) specific adverse events	Low risk	**Comment:** there was no loss to follow-up and this outcome was reported on all participants. Therefore, we assessed this risk of bias as low.
Incomplete outcome data (attrition bias) perioperative adverse events	Low risk	**Comment:** there was no loss to follow-up and this outcome was reported on all participants. Therefore, we assessed this risk of bias as low.
Selective reporting (reporting bias)	High risk	**Comment:** some of the secondary outcomes were either analysed but no results were reported, or they were measured but not analysed. Therefore, we assessed this risk of bias as high. See ORBIT Appendix 8.
Other bias	Unclear risk	**Comment:** we considered surgical expertise, funding and conflict of interest as other potential sources of bias. Bergenfelz and colleagues specified that all surgeons performing operations were experienced, but did not dwell on the level of experience and familiarity with each technique used in particular. It did not mention sources of funding either. The presence of conflict of interest could not be excluded as a result of those 2 factors. Therefore, we assessed this risk of bias as unclear.

Miccoli 1999

Study characteristics	
Methods	**Study design:** parallel-group randomised controlled study **Setting:** hospital **Country where study was performed:** Italy **Number of study centres:** 1 (Università Degli Studi di Pisa, Pisa, Italy)
Participants	**Inclusion criteria:** sporadic form of PHPT, no prior neck surgery, absence of thyroid nodules and preoperative ultrasonography suggestive for solitary parathyroid adenoma **Exclusion criteria:** prior neck surgery, thyroid nodules and preoperative ultrasonography not suggestive for solitary parathyroid adenoma **Diagnostic criteria:** no clear diagnostic criteria provided **Age group:** adults

Miccoli 1999 *(Continued)*

	Gender distribution: women and men
Interventions	Participants with PHPT were referred for parathyroidectomy. Despite the preoperative localisation studies that may have been ordered earlier by their referring physician, an ultrasound examination of the neck performed by an expert radiologist. All procedures were performed by the same surgeon and participants were randomly allocated to MIP or BNE group. **Intervention:** the VAP procedure was started with a 15-mm incision at the notch level, and under endoscopic vision, carbon dioxide insufflation (12 mmHg) was done and then a 30-degree 5-mm endoscope allowed optimal visualisation of the operative field. Needle-scopic instruments (2 mm) were used in order to identify the parathyroid adenoma. The completeness of the surgical resection was only confirmed when there was a ≥ 50% decrease in intact PTH values, with respect to the highest pre-excision level; and then measurements were obtained 5 and 10 minutes after the adenoma removal. The BNE was done under endotracheal general anaesthesia. Frozen section was used during the procedure and no biopsy specimens were obtained from the normal parathyroid glands. Intraoperative PTH levels were not measured. **Comparator:** conventional cervicotomy with BNE **Duration of surgery:** reported as operative time **Length of hospital stay:** not reported **Duration of follow-up:** 12, 24 and 48 hours; 1, 3 and 6 months after surgery **Run-in period:** participants were followed up at 12, 24 and 48 hours after the operation and 1, 3 and 6 months after surgery **Extension period:** no extension period was done after the decided follow-up date
Outcomes	**Reported outcome(s) in full text of publication** Success rate, operative time, pain, cost analysis, fever, symptomatic hypocalcaemia, wound infection, vocal cord disorders, inactivity period and the time required to return to normal activities. A personal opinion at 1, 3 and 6 months after surgery was also reported.
Study details	**Trial identifier:** not reported The study was not terminated early.
Publication details	**Language of publication:** English **Funding:** no sources of funding were stated. **Publication status:** peer-reviewed journal
Stated aim for study	**Quote:** "Aiming to compare a minimally invasive operation with a conventional operation for PHRPT, we analysed, in a prospective randomised study, costs and results of 2 surgical procedures that are currently performed in our unit: the video assisted parathyroidectomy (VAP) as previously described and the classical bilateral neck exploration."
Notes	Authors did not include or analyse participants with multiglandular disease.

Risk of bias

Bias	Authors' judgement	Support for judgement

Miccoli 1999 *(Continued)*

Random sequence generation (selection bias)	Low risk	**Quote:** "Thirty-eight participants were considered eligible for VAP [video-assisted parathyroidectomy]; they were enrolled in the study and randomly divided into 2 groups by flipping a coin."
Allocation concealment (selection bias)	Unclear risk	**Quote:** "Flipping a coin." Unclear if coin flipping was done prior to patient presentation or after inclusion which would determine whether or not there was allocation concealment. **Quote from *Cochrane Handbook for Systematic Reviews of Interventions*:** "This is usually the case if the method of concealment is not described or not described in sufficient detail to allow a definite judgement."
Blinding of participants and personnel (performance bias) conversion rate	High risk	**Comment:** there was insufficient information about blinding of participants and study personnel but conversion rate is dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) cosmetic satisfaction	Unclear risk	**Comment:** there was insufficient information about blinding of participants and study personnel. Therefore, we assessed this risk of bias as unclear.
Blinding of participants and personnel (performance bias) duration of surgery	High risk	**Comment:** blinding was unclear but the outcome, namely duration of surgery, was likely to be influenced by lack of blinding of the surgeon. Therefore, we assessed this risk of bias as high.
Blinding of participants and personnel (performance bias) socioeconomic effects	High risk	**Comment:** there was insufficient information about blinding of participants and study personnel, but socioeconomic cost may be affected by the number of tests requested by the surgeon, who in this instance was not blinded. Therefore, we assessed this risk of bias as high.
Blinding of participants and personnel (performance bias) success rate	Unclear risk	**Comment:** there was insufficient information about blinding of participants and study personnel, and success rate was not clearly defined as an outcome. It was only mentioned in the Results section "All patients were normocalcaemic 6 months after surgery" and therefore, we assessed this risk of bias as unclear.
Blinding of participants and personnel (performance bias) specific adverse events	High risk	**Comment:** there was insufficient information about blinding of participants and study personnel, and the outcome was likely to have been influenced by lack of blinding, and therefore, we assessed this risk of bias as high.
Blinding of participants and personnel (performance bias) perioperative adverse events	High risk	**Comment:** there was insufficient information about blinding of participants and study personnel, and the outcome was likely to have been influenced by lack of blinding, and therefore, we assessed this risk of bias as high.
Blinding of outcome assessment (detection bias) absence from work	High risk	**Comment:** there was insufficient information about blinding of participants and study personnel, and the outcome was likely to have been influenced by lack of blinding, and therefore, we assessed this risk of bias as high.
Blinding of outcome assessment (detection bias) cosmetic satisfaction	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors, and therefore, we assessed this risk of bias as unclear.
Blinding of outcome assessment (detection bias)	High risk	**Quote:** "The patient's eligibility for VAP was considered on the basis of both clinical history and ultrasound findings: sporadic form of PHPT."

Miccoli 1999 *(Continued)*

duration of surgery		**Comment:** although blinding of outcome assessors was not mentioned, we judged the risk of bias as high as this outcome may have been affected by the fact that randomisations were done only after finding a positive finding in preoperative ultrasound finding.
Blinding of outcome assessment (detection bias) length of hospital stay	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors, and therefore, we assessed this risk of bias as unclear.
Blinding of outcome assessment (detection bias) socioeconomic effects	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors, and therefore, we assessed this risk of bias as unclear.
Blinding of outcome assessment (detection bias) success rate	High risk	**Comment:** insufficient information about the blinding of outcome assessors. Although blinding of outcome assessors was not mentioned, we judged the risk of bias as high, as this outcome may have been affected by the fact that randomisation was done only after finding a positive finding in preoperative ultrasound finding.
Blinding of outcome assessment (detection bias) specific adverse events	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors, and therefore, we assessed this risk of bias as unclear.
Blinding of outcome assessment (detection bias) perioperative adverse events	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors, and therefore, we assessed this risk of bias as unclear.
Incomplete outcome data (attrition bias) absence from work	Unclear risk	**Quote:** "Two patients had multiglandular disease that was discovered during surgery and were excluded from the study." **Comment:** we judged the risk of bias as unclear, as there was insufficient information to assess whether missing data (omitting 2/38 participants with multiglandular disease from study results) in combination with the method used to handle missing data were likely to induce bias.
Incomplete outcome data (attrition bias) cosmetic satisfaction	Unclear risk	**Quote:** "Two patients had multiglandular disease that was discovered during surgery and were excluded from the study." **Comment:** we judged the risk of bias as unclear, as there was insufficient information to assess whether missing data (omitting 2/38 participants with multiglandular disease from study results) in combination with the method used to handle missing data were likely to induce bias.
Incomplete outcome data (attrition bias) duration of surgery	Unclear risk	**Quote:** "Two patients had multiglandular disease that was discovered during surgery and were excluded from the study." **Comment:** we judged the risk of bias as unclear, as there was insufficient information to assess whether missing data (omitting 2/38 participants with multiglandular disease from study results) in combination with the method used to handle missing data were likely to induce bias.
Incomplete outcome data (attrition bias) socioeconomic effects	Unclear risk	**Quote:** "Two patients had multiglandular disease that was discovered during surgery and were excluded from the study." **Comment:** we judged risk of bias as unclear, as there was insufficient information to assess whether missing data (omitting 2/38 participants with multiglandular disease from study results) in combination with the method used to handle missing data were likely to induce bias.

Miccoli 1999 *(Continued)*

Incomplete outcome data (attrition bias) success rate	Unclear risk	**Quote:** "Two patients had multiglandular disease that was discovered during surgery and were excluded from the study." **Comment:** we judged the risk of bias as unclear, as there was insufficient information to assess whether missing data (omitting 2/38 participants with multiglandular disease from study results) in combination with the method used to handle missing data were likely to induce bias.
Incomplete outcome data (attrition bias) specific adverse events	Unclear risk	**Quote:** "Two patients had multiglandular disease that was discovered during surgery and were excluded from the study'" **Comment:** we judged the risk of bias as unclear, as there was insufficient information to assess whether missing data (omitting 2/38 participants with multiglandular disease from study results) in combination with the method used to handle missing data were likely to induce bias.
Incomplete outcome data (attrition bias) perioperative adverse events	Unclear risk	**Quote:** "Two patients had multiglandular disease that was discovered during surgery and were excluded from the study." **Comment:** we judged the risk of bias as unclear, as there was insufficient information to assess whether missing data (omitting 2/38 participants with multiglandular disease from study results) in combination with the method used to handle missing data were likely to induce bias.
Selective reporting (reporting bias)	High risk	**Comment:** protocol was not identified. Study did not specify primary vs secondary outcomes in the Methods and Results section of the paper. Moreover, some of the outcomes were reported in the Results section (normocalcaemic rate at 6 months and postoperative specific adverse events) but these outcomes were not mentioned in the Methods section. In addition, postoperative pain evaluation and cosmetic satisfaction results were only provided in a figure, and not reported adequately. Therefore, we assessed this risk of bias as high. See ORBIT Appendix 8.
Other bias	Unclear risk	**Comment:** surgeon familiarity and experience with each surgical technique was not detailed. There was no mention of any sources of funding either. The presence of conflict of interest could not be excluded as a result of those 2 factors.

Miccoli 2008

Study characteristics	
Methods	**Study design:** parallel-group randomised controlled study **Setting:** hospital **Country where study was performed:** Italy **Number of study centres:** 1 (University of Pisa, Italy)
Participants	**Inclusion/exclusion/diagnostic criteria:** people undergoing surgery for PHPT. No clear inclusion, exclusion or diagnostic criteria identified. **Age group:** adults **Gender distribution:** women and men

Miccoli 2008 *(Continued)*

Interventions	**Intervention:** eligible participants were randomly allocated to either parathyroidectomy using the MIVAP technique plus intraoperative PTH, or MIVAP plus a bilateral endoscopic neck exploration. All participants underwent preoperative localisation studies with ultrasounds and sestamibi-99Tc scans. The MIVAP technique was carried through a central neck incision of 15–20 mm, 2 cm above the sternal notch, using an endoscope without gas insufflation. The surgical procedure was ended when a decrease of ≥ 50% with respect to the highest pre-excision level. The BNE was performed via the same central neck access for a MIVAP procedure, with the procedure ended after visualisation of the 4 parathyroid glands and removal of the macroscopically enlarged glands. Intraoperative PTH levels were not monitored. **Comparator:** endoscopic parathyroidectomy plus BNE **Duration of surgery:** reported as mean operative time **Length of hospital stay:** not reported **Duration of follow-up:** 48 hours, 1 and 6 months after surgery **Run-in period:** participants were followed up until 6 months postoperatively. **Extension period:** no extension period was done after the decided follow-up date.
Outcomes	**Reported outcome(s) in full text of publication** Mean operative time and outcome of the surgical procedure (PTH and calcium normalisation at 1 and 6 months postoperatively)
Study details	**Trial identifier:** not reported The study was not terminated early.
Publication details	**Language of publication:** English **Funding:** no sources of funding were stated **Publication status:** peer-reviewed journal
Stated aim for study	**Quote:** "A comparison of bilateral video-assisted neck exploration after removal of enlarged gland and focused MIVAP plus quick iPTHa was conducted in order to evaluate the effectiveness of the two techniques in the treatment of patients with PHPT, their relative outcomes, and operative time with an attempt to also examine their relative costs."
Notes	No clear information was available of whether only participants with uniglandular disease were included after the preoperative localisation procedure was done.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	**Quote:** "Patients were randomly allocated into one of two groups." **Comment:** insufficient information about the sequence generation process.
Allocation concealment (selection bias)	Unclear risk	**Comment:** it was not adequately described so there was insufficient detail to allow a definite judgement.
Blinding of participants and personnel (performance bias) conversion rate	High risk	**Quote:** "All patients were blind to the technique used as well as the pathologist who examined the specimens."

Miccoli 2008 *(Continued)*

		Comment: author replied by email indicating that participants were blinded to study group, but conversion rate is dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) duration of surgery	High risk	**Quote:** "All patients were blind to the technique used as well as the pathologist who examined the specimens." **Comment:** although participants were blinded, surgery duration is in part dependent on conversion rate, and therefore, surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) socioeconomic effects	High risk	**Comment:** although participants were blinded, socioeconomic cost may be affected by the number of tests requested by the surgeon, who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) success rate	High risk	**Comment:** although participants were blinded, as per email reply from the author, success rate is dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) specific adverse events	Low risk	**Comment:** the authors mentioned that participants were blinded, and, since this outcome is not dependent on the surgeon, we assessed the risk of bias as low.
Blinding of participants and personnel (performance bias) perioperative adverse events	Low risk	**Comment:** the authors mentioned that participants were blinded, and since this outcome is not dependent on the surgeon, we assessed the risk of bias as low.
Blinding of outcome assessment (detection bias) conversion rate	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) duration of surgery	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) socioeconomic effects	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) success rate	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) specific adverse events	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) perioperative adverse events	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Incomplete outcome data (attrition bias)	Low risk	**Comment:** no loss of follow-up as data were provided for all the participants included. Therefore, we assessed the risk of bias as low.

Miccoli 2008 *(Continued)*

conversion rate		
Incomplete outcome data (attrition bias) duration of surgery	Low risk	**Comment:** no loss of follow-up as data were provided for all the participants included. Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) socioeconomic effects	Low risk	**Comment:** no loss of follow-up as data were provided for all the participants included. Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) success rate	Low risk	**Comment:** no loss of follow-up as data were provided for all the participants included. Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) specific adverse events	Low risk	**Comment:** no loss of follow-up as data were provided for all the participants included. Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) perioperative adverse events	Low risk	**Comment:** no loss of follow-up as data were provided for all the participants included. Therefore, we assessed the risk of bias as low.
Selective reporting (reporting bias)	High risk	**Comment:** protocol was not identified. Study did not have any information on outcomes of interest in the Method section, and therefore the risk of reporting bias was considered as high risk. See ORBIT Appendix 8.
Other bias	Unclear risk	**Comment:** surgeon familiarity and experience with each surgical technique was not detailed. There was no mentioning of any sources of funding either. The presence of conflict of interest could not be excluded as a result of those 2 factors.

Slepavicius 2008

Study characteristics	
Methods	**Study design:** parallel-group randomised controlled study **Setting:** hospital **Country where study was performed:** Lithuania **Number of centres:** 2 (department of abdominal and endocrine surgery of Klaipeda University Hospital and second department of abdominal surgery of Vilnius University Hospital "Santariskiu Klinikos", Vilnius, Lithuania).
Participants	**Inclusion criteria:** people with sporadic PHPT **Exclusion criteria:** family history of PHPT, relapse of PHPT, previous neck surgery, people with indications for partial or complete removal of thyroid gland, severe concomitant pathology, making surgical treatment impossible, people who due to psychological disorders could not evaluate adequately their health status, pregnancy and breastfeeding, people with symptoms of hypercalcaemic crisis, people refusing to participate during the study **Diagnostic criteria:** PHPT determined clinically and with laboratory tests **Age group:** adults

Slepavicius 2008 *(Continued)*

	Gender distribution: not specified
Interventions	**Intervention:** the MIP participants had preoperative localisation studies before operation. Then a 2- to 2.5-cm transverse incision done. The incision for presumed inferior gland was placed 2 cm above the clavicle, whereas that 1 for presumed superior gland was placed higher. No attempts were made to visualise normal parathyroid glands. The adequacy of resection was assessed with a decrease of ≥ 50% of intraoperative rapid PTH measurements from the baseline at 15 minutes after gland resection as indicative of successful parathyroidectomy. The BNE participants were those for which parathyroidectomy was performed with traditional Kocher incision and revision of all parathyroid glands. No localisation preoperative examination carried out. Intraoperative PTH levels were not monitored in those participants as well. **Comparator:** conventional surgery group (BNE) **Duration of surgery:** reported as operative time **Length of hospital stay:** not reported **Duration of follow-up:** 4, 8, 16, 24, 36 and 48 hours after surgery, then at 1 and 6 months and 1 year after surgery **Run-in period:** participants were followed up until 1 year postoperatively. **Extension period:** there was no extension period after the decided follow-up date.
Outcomes	**Reported outcome(s) in full text of publication** **Primary outcome:** cure rate: normocalcaemia or hypocalcaemia at 6 months were considered evidence for cure **Secondary endpoints:** postoperative pain intensity, analgesics consumption, time of surgery, cosmetic satisfaction, quality of life, cost effectiveness, vocal cord function, hypocalcaemia and eucalcaemic HPT
Study details	**Trial identifier:** not reported The study was not terminated early.
Publication details	**Language of publication:** English **Funding:** no sources of funding were stated. **Publication status:** peer-reviewed journal
Stated aim for study	**Quote:** "The study is aimed to compare BNE and focused parathyroidectomy (FP) in a prospective, randomised, blind trial."
Notes	It appeared that surgeon completed skin incision and then confirmed PTH results. The authors reported time from skin to skin to be 36 minutes, but they also stated that PTH results lengthened surgery by 30 minutes, which would render a total surgical time of 36 minutes unlikely.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	**Quote:** "Forty seven patients, which, according to double random principle, were subdivided into two groups according to the surgery mode (traditional surgery or focused operation) were included into the study."

Slepavicius 2008 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	**Quote:** "Help of envelopes." **Comment:** no information about opaqueness of envelopes.
Blinding of participants and personnel (performance bias) conversion rate	High risk	**Comment:** insufficient information about the blinding of participants and personnel, but conversion rate is dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) cosmetic satisfaction	Unclear risk	**Comment:** insufficient information about the blinding of participants and personnel. Therefore, we assessed the risk of bias as unclear.
Blinding of participants and personnel (performance bias) duration of surgery	High risk	**Comment:** insufficient information about the blinding of participants. Surgery duration is in part dependent on conversion rate, and therefore, surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) health-related quality of life	Unclear risk	**Comment:** insufficient information about the blinding of participants and personnel. Therefore, we assessed the risk of bias as unclear.
Blinding of participants and personnel (performance bias) postoperative increase in PTH with eucalcaemia	High risk	**Comment:** insufficient information about the blinding of participants. Postoperative increase in PTH with eucalcaemia is in part dependent on conversion rate, and therefore on surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) socioeconomic effects	High risk	**Comment:** insufficient information about the blinding of participants, but socioeconomic cost may be affected by the number of tests requested by the surgeon, who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) success rate	High risk	**Comment:** insufficient information about the blinding of participants, but success rate is dependent on the surgeon who in this instance was not blinded. Therefore, we assessed the risk of bias as high.
Blinding of participants and personnel (performance bias) specific adverse events	Unclear risk	**Comment:** insufficient information about the blinding of participants and personnel. Therefore, we assessed the risk of bias as unclear.
Blinding of participants and personnel (performance bias) perioperative adverse events	Unclear risk	**Comment:** insufficient information about the blinding of participants and personnel. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) conversion rate	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias)	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.

Slepavicius 2008 *(Continued)*

cosmetic satisfaction		
Blinding of outcome assessment (detection bias) duration of surgery	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) health-related quality of life	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) postoperative increase in PTH with eucalcaemia	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) socioeconomic effects	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) success rate	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) specific adverse events	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Blinding of outcome assessment (detection bias) perioperative adverse events	Unclear risk	**Comment:** insufficient information about the blinding of outcome assessors. Therefore, we assessed the risk of bias as unclear.
Incomplete outcome data (attrition bias) conversion rate	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) cosmetic satisfaction	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) duration of surgery	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) health-related quality of life	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) length of hospital stay	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.

Slepavicius 2008 *(Continued)*

Incomplete outcome data (attrition bias) postoperative increase in PTH with eucalcaemia	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) socioeconomic effects	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) success rate	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) specific adverse events	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Incomplete outcome data (attrition bias) perioperative adverse events	Low risk	**Comment:** missing outcome data were balanced in numbers across intervention groups (2/23 in BNE group and 3/24 in MIP group), with similar reasons for missing data in both groups (namely hyperplasia found intraoperatively). Therefore, we assessed the risk of bias as low.
Selective reporting (reporting bias)	High risk	**Comment:** protocol was not identified. In addition, 1 of the outcomes was measured and analysed; study report stated that outcome was analysed but reported only that result was not significant. Therefore, we assessed the risk of bias as high. See ORBIT Appendix 8.
Other bias	Unclear risk	**Comment:** surgeon familiarity and experience with each surgical technique was not detailed. There was no mentioning of any sources of funding either. The presence of conflict of interest could not be excluded as a result of those 2 factors.

Note: where the judgement is 'Unclear' and the description is blank, the study did not report that particular outcome.
BNE: bilateral neck exploration; HPT: hyperparathyroidism; MEN: multiple endocrine neoplasia; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; PHPT: primary hyperparathyroidism; PTH: parathyroid hormone; VAP: video-assisted parathyroidectomy.

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Aarum 2007	No use of intraoperative PTH.
Adler 2008	Not a randomised controlled study
Baliski 2008	Not a randomised controlled study
Barczyński 2006	No comparison between MIP and BNE
Beyer 2007	Not a randomised controlled study
Calò 2017	Not a randomised controlled study

Study	Reason for exclusion
Carneiro 2000	Not a randomised controlled study
Clerici 1992	Not a randomised controlled study
Del Rio 2013	Not a randomised controlled study
Ejlsmark-Svensson 2019	Not a randomised controlled study
Genc 2003	Not a randomised controlled study
Grant 2005	Not a randomised controlled study
Hansen 2012	Not a randomised controlled study
Henry 2001	Not a randomised controlled study
Hessman 2010	No comparison between MIP and BNE
Lumachi 2003	Not a randomised controlled study
McGill 2008	Not a randomised controlled study
Mekel 2014	Not a randomised controlled study
NCT00877981	No comparison between MIP and BNE
NCT01776502	Not a randomised controlled study
Nehs 2013	Not a randomised controlled study
Neychev 2016	Not a randomised controlled study
Nilsson 2017	Not a randomised controlled study
Norenstedt 2009	Not a randomised controlled study
Norman 2015	Not a randomised controlled study
Pepe 2013	Not a randomised controlled study
Russel 2006	No use of intraoperative PTH
Schneider 2012	Not a randomised controlled study
Simonella 2005	Not a randomised controlled study
Sozio 2005	No use of intraoperative PTH
Tolley 2016	Not a randomised controlled study
Udelsman 2002	Not a randomised controlled study
Udelsman 2011	Not a randomised controlled study
Usta 2015	Not a randomised controlled study

BNE: bilateral neck exploration; MIP: minimally invasive parathyroidectomy; PTH: parathyroid hormone.

Characteristics of studies awaiting classification *[ordered by study ID]*

Fan 2013

Methods	**Study design:** randomised controlled study
Participants	**Inclusion criteria:** people with PHPT **Exclusion criteria:** not available **Diagnostic criteria:** not available **Setting:** Medical Center of Thyroid Disease, Affiliated Sixth People's Hospital, Shanghai Jiaotong University, Shanghai, China; 2 Surgery, Shanghai Sixth People's Hospital, Shanghai, China **Age group:** adults **Country where study was performed**: China
Interventions	**Intervention:** preoperative localisation and intraoperative PTH detection were performed in the minimally invasive video-assisted parathyroidectomy group. **Comparator:** only tumour size observation and rapid frozen pathology detection were used in the BNE group. **Duration of surgery:** operative time **Length of hospital stay:** not reported **Duration of follow-up:** follow-up was performed after 6 weeks, 6 months and 1 year postoperatively **Run-in period:** not reported **Extension period:** not available **Number of study centres:** 2 (Medical Center of Thyroid Disease, Affiliated Sixth People's Hospital, Shanghai Jiaotong University, Shanghai, China; 2 Surgery, Shanghai Sixth People's Hospital, Shanghai, China) 108 participants (52 in BNE group and 56 in MIVAP group) with PHP were enrolled in the study.
Outcomes	Cure rate, persistent or recurrent PHPT, early severe hypocalcaemia, cosmetic satisfaction rate, operative time, postoperative pain and hospital stay
Study details	**Trial identifier:** not reported Unclear in the poster whether the study was terminated early.
Publication details	Only presented as a poster in the 83rd Annual Meeting of American Thyroid Association.
Stated aim of study	Quote: "The purpose of this study is to perform a randomised controlled trial to compare the therapeutic effects of BNE and MIVAP."
Notes	This was a poster presented in the 83rd Annual Meeting of American Thyroid Association, but there was no clear evidence of a journal publication.

BNE: bilateral neck exploration; MIVAP: minimally invasive video-assisted parathyroidectomy; PHPT: primary hyperparathyroidism; PTH: parathyroid hormone.

Characteristics of ongoing studies *[ordered by study ID]*

NCT00006329

Study name	Comparison of two methods of parathyroidectomy for primary hyperparathyroidism
Methods	**Study design:** randomised controlled study
Participants	**Inclusion criteria:** participants with PHPT with elevated serum level of calcium on ≥ 2 occasions. Furthermore, it was stated that if the level is < 11.0 mg/dL, then this must have been present for ≥ 6 months. Furthermore, if the serum albumin concentration is not within normal then the ionised calcium level is to be measured and needs to be elevated for inclusion of participants aged ≥ 12 years. Participants need to have nephrolithiasis or documented bone mineral density < 2.5 standard deviations below age-matched means and elevated or non-suppressed serum intact parathyroid hormone level. **Exclusion criteria:** not available **Diagnostic criteria:** elevated serum level of calcium on ≥ 2 occasions, and if the level is < 11.0 mg/dL, then this must have been present for ≥ 6 months. This is along with the presence of an elevated or non-suppressed serum intact parathyroid hormone level for the participants to be included into the study. **Setting:** University of Michigan Hospital, Department of Surgery **Age group:** adolescents and adults aged ≥ 12 years **Country where study is being performed:** US **Number of study centres:** 1 (University of Michigan Hospital, Department of Surgery) Number of participants was not stated.
Interventions	**Intervention:** preoperative localisation with a sestamibi nuclear medicine scan and intra-operative parathyroid hormone detection performed in the minimally invasive video-assisted parathyroidectomy group **Comparator:** BNE **Duration of surgery:** not stated **Length of hospital stay:** not stated **Duration of follow-up:** not stated **Extension period:** not available
Outcomes	Total cost of the care, testing for complications such as vocal cord dysfunction, assessment of the general health status, pain levels, patient satisfaction with the operation
Starting date	5 October 2000
Contact information	Study research staff using the contact information provided by the sponsor (National Center for Research Resources)
Study identifier	NCT00006329
Official title	Comparison of two methods of parathyroidectomy for primary hyperparathyroidism
Stated purpose of study	Quote: "This study serves to directly compare the costs, the effectiveness, and the safety associated with each type of operation."

NCT00006329 *(Continued)*

Notes	Last update posted was 24 June 2005 with no information available about the publication of the results of this trial.

BNE: bilateral neck exploration; PHPT: primary hyperparathyroidism.

DATA AND ANALYSES

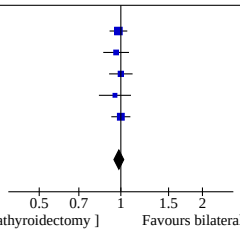
Comparison 1. Minimally invasive parathyroidectomy versus bilateral neck exploration

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Success rate up to 6 months	5	266	Risk Ratio (M-H, Random, 95% CI)	0.98 [0.94, 1.03]
1.2 Success rate at 5 years	1		Risk Ratio (M-H, Fixed, 95% CI)	Totals not selected
1.3 Total incidence of perioperative adverse events	5	266	Risk Ratio (M-H, Random, 95% CI)	0.50 [0.33, 0.76]
1.4 Bleeding	2		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.5 Symptomatic hypocalcaemia	4	202	Risk Ratio (M-H, Random, 95% CI)	0.54 [0.32, 0.92]
1.6 Permanent hypocalcaemia	5		Risk Ratio (M-H, Fixed, 95% CI)	Subtotals only
1.7 Postoperative pain score	2	133	Std. Mean Difference (IV, Random, 95% CI)	-0.70 [-1.69, 0.28]
1.8 Vocal cord paralysis	5	261	Risk Ratio (M-H, Random, 95% CI)	1.87 [0.47, 7.51]
1.9 Postoperative increase in parathyroid hormone with eucalcaemia	2	133	Risk Ratio (M-H, Random, 95% CI)	0.81 [0.43, 1.53]
1.10 Duration of surgery	3	171	Mean Difference (IV, Random, 95% CI)	-18.33 [-30.71, -5.95]

Analysis 1.1. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 1: Success rate up to 6 months

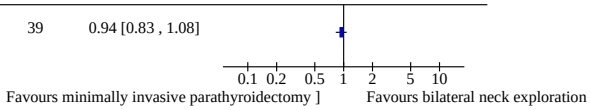
Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Bergenfelz 2002	45	47	43	44	31.1%	0.98 [0.91 , 1.06]
Bergenfelz 2005	24	25	25	25	14.7%	0.96 [0.86 , 1.07]
Miccoli 1999	20	20	18	18	17.7%	1.00 [0.91 , 1.10]
Miccoli 2008	19	20	20	20	9.5%	0.95 [0.83 , 1.09]
Slepavicius 2008	24	24	23	23	26.9%	1.00 [0.92 , 1.08]
Total (95% CI)		**136**		**130**	**100.0%**	**0.98 [0.94 , 1.03]**
Total events:	132		129			

Heterogeneity: Tau² = 0.00; Chi² = 0.74, df = 4 (P = 0.95); I² = 0%
Test for overall effect: Z = 0.80 (P = 0.43)
Test for subgroup differences: Not applicable



Analysis 1.2. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 2: Success rate at 5 years

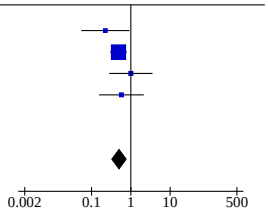
Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Risk Ratio M-H, Fixed, 95% CI
Bergenfelz 2002	34	38	37	39	0.94 [0.83 , 1.08]



Analysis 1.3. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 3: Total incidence of perioperative adverse events

Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Miccoli 1999	2	20	8	18	8.7%	0.23 [0.05 , 0.92]
Bergenfelz 2002	14	47	27	44	70.4%	0.49 [0.30 , 0.80]
Bergenfelz 2005	4	25	4	25	10.8%	1.00 [0.28 , 3.56]
Slepavicius 2008	3	24	5	23	10.1%	0.57 [0.15 , 2.14]
Miccoli 2008	0	20	0	20		Not estimable
Total (95% CI)		**136**		**130**	**100.0%**	**0.50 [0.33 , 0.76]**
Total events:	23		44			

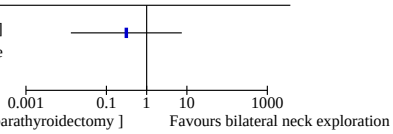
Heterogeneity: Tau² = 0.00; Chi² = 2.43, df = 3 (P = 0.49); I² = 0%
Test for overall effect: Z = 3.26 (P = 0.001)
Test for subgroup differences: Not applicable



Analysis 1.4. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 4: Bleeding

Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Risk Ratio M-H, Fixed, 95% CI
Bergenfelz 2002	0	47	1	44	0.31 [0.01 , 7.47]
Miccoli 2008	0	20	0	20	Not estimable

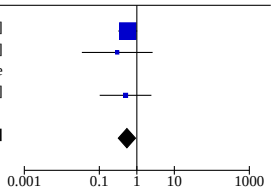
Test for subgroup differences: Not applicable



Analysis 1.5. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 5: Symptomatic hypocalcaemia

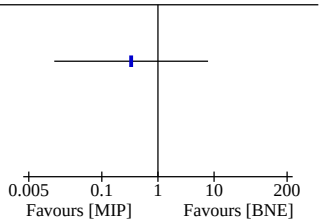
Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Bergenfelz 2002	12	43	19	39	83.1%	0.57 [0.32 , 1.02]
Miccoli 1999	1	20	3	18	5.9%	0.30 [0.03 , 2.63]
Miccoli 2008	0	20	0	20		Not estimable
Slepavicius 2008	2	21	4	21	11.0%	0.50 [0.10 , 2.44]
Total (95% CI)		**104**		**98**	**100.0%**	**0.54 [0.32 , 0.92]**
Total events:	15		26			

Heterogeneity: Tau² = 0.00; Chi² = 0.34, df = 2 (P = 0.84); I² = 0%
Test for overall effect: Z = 2.27 (P = 0.02)
Test for subgroup differences: Not applicable



Analysis 1.6. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 6: Permanent hypocalcaemia

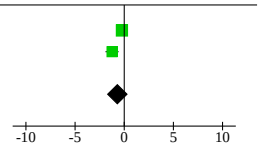
Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Risk Ratio M-H, Fixed, 95% CI
Bergenfelz 2002	0	47	0	44	Not estimable
Bergenfelz 2005	0	25	1	25	0.33 [0.01 , 7.81]
Miccoli 1999	0	20	0	18	Not estimable
Miccoli 2008	0	20	0	20	Not estimable
Slepavicius 2008	0	24	0	23	Not estimable



Analysis 1.7. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 7: Postoperative pain score

Study or Subgroup	Minimally invasive parathyroidectomy Mean	SD	Total	Bilateral neck exploration Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Bergenfelz 2002 (1)	15.9	16.25	47	20	18.4	44	53.5%	-0.23 [-0.65 , 0.18]
Slepavicius 2008	0	5	21	10	10	21	46.5%	-1.24 [-1.91 , -0.58]
Total (95% CI)			**68**			**65**	**100.0%**	**-0.70 [-1.69 , 0.28]**

Heterogeneity: Tau² = 0.43; Chi² = 6.34, df = 1 (P = 0.01); I² = 84%
Test for overall effect: Z = 1.40 (P = 0.16)
Test for subgroup differences: Not applicable

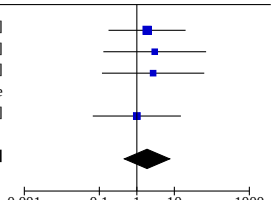


Footnotes
(1) Data estimated from figure 2 in the publication

Analysis 1.8. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 8: Vocal cord paralysis

Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Bergenfelz 2002	2	47	1	44	34.6%	1.87 [0.18 , 19.93]
Bergenfelz 2005	1	25	0	25	19.4%	3.00 [0.13 , 70.30]
Miccoli 1999	1	20	0	18	19.6%	2.71 [0.12 , 62.70]
Miccoli 2008	0	20	0	20		Not estimable
Slepavicius 2008	1	21	1	21	26.4%	1.00 [0.07 , 14.95]
Total (95% CI)		**133**		**128**	**100.0%**	**1.87 [0.47 , 7.51]**
Total events:	5		2			

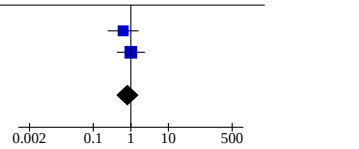
Heterogeneity: Tau² = 0.00; Chi² = 0.35, df = 3 (P = 0.95); I² = 0%
Test for overall effect: Z = 0.88 (P = 0.38)
Test for subgroup differences: Not applicable



Analysis 1.9. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 9: Postoperative increase in parathyroid hormone with eucalcaemia

Study or Subgroup	Minimally invasive parathyroidectomy Events	Total	Bilateral neck exploration Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Bergenfelz 2002	6	47	9	44	44.9%	0.62 [0.24 , 1.61]
Slepavicius 2008	7	21	7	21	55.1%	1.00 [0.43 , 2.35]
Total (95% CI)		**68**		**65**	**100.0%**	**0.81 [0.43 , 1.53]**
Total events:	13		16			

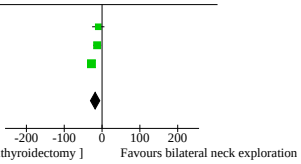
Heterogeneity: Tau² = 0.00; Chi² = 0.53, df = 1 (P = 0.47); I² = 0%
Test for overall effect: Z = 0.65 (P = 0.51)
Test for subgroup differences: Not applicable



Analysis 1.10. Comparison 1: Minimally invasive parathyroidectomy versus bilateral neck exploration, Outcome 10: Duration of surgery

Study or Subgroup	Minimally invasive parathyroidectomy Mean [minutes]	SD [minutes]	Total	Bilateral neck exploration Mean [minutes]	SD [minutes]	Total	Weight	Mean Difference IV, Random, 95% CI [minutes]
Miccoli 1999	72	42	47	82	37	44	25.4%	-10.00 [-26.24 , 6.24]
Bergenfelz 2002	57	15	20	70	18	18	34.0%	-13.00 [-23.60 , -2.40]
Slepavicius 2008	36	3.8	21	64	14	21	40.6%	-28.00 [-34.20 , -21.80]
Total (95% CI)			**88**			**83**	**100.0%**	**-18.33 [-30.71 , -5.95]**

Heterogeneity: Tau² = 88.19; Chi² = 8.37, df = 2 (P = 0.02); I² = 76%
Test for overall effect: Z = 2.90 (P = 0.004)
Test for subgroup differences: Not applicable



ADDITIONAL TABLES

Table 1. Overview of study populations

Study (design)	Intervention(s) and comparator(s)	Description of power and sample size calculation	Screened/ eligible (n)	Randomised (n)	Analysed (n)	Finishing study (n)	Randomised finishing study (%)	Follow-up
Bergenfelz 2005 (parallel RCT)	I: MIP guided by intraoperative PTH monitoring and preoperative imaging	—	233	25	25	25	100	First 4 days after surgery, 1 and 6 months
	C: BNE			25	25	25	100	
	Total:		**233**	**50**	**50**	**50**	**100**	
Bergenfelz 2002 (parallel RCT)	I: MIP guided by intraoperative PTH monitoring and preoperative imaging	—	47	47	47	38	80	First 4 postoperative days, 6 weeks, 1 and 5 years
	C: BNE		44	44	44	33	75	
	Total:		**91**	**91**	**91**	**71**	**78**	
Miccoli 2008 (parallel RCT)	I: focused endoscopic parathyroidectomy (MIVAP) – MIP guided by intraoperative PTH monitoring and preoperative imaging	—	20	20	20	20	100	First postoperative day; 1 and 6 months
	C: BNE		20	20	20	20	100	
	Total:		**40**	**40**	**40**	**40**	**100**	
Miccoli 1999 (parallel RCT)	I: MIVAP guided by intraoperative PTH monitoring and preoperative imaging	—	38	20	20	20	100	12, 24, and 48 hours after surgery; 1, 3 and 6 months
	C: BNE			18	18	18	100	
	Total:		**38**	**38**	**38**	**38**		
Slepavicius 2008 (parallel RCT)	I: MIP guided by intraoperative PTH monitoring and preoperative imaging		47	24	24	24	100	48 hours after surgery, 4 weeks, 1 and 6 months, 1 year
	C: BNE			23	23	23	100	

Table 1. Overview of study populations *(Continued)*

	Total:		**47**	**47**	**47**	**47**	**100**
Overall total	**All interventions**	—		**136**		**127**	—
	All comparators			**130**		**119**	
	All interventions and comparators			**266**		**246**	

—: denotes not clearly described in the publication
BNE: bilateral neck exploration; C: comparator; I: intervention; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; n: number of participants; PTH: parathyroid hormone; RCT: randomised controlled trial.

Table 2. Sensitivity analyses

Protocol item	Symptomatic hypocalcaemia	Duration of surgery
Restricting to published	NA	NA
Restricting the analysis taking into account risk of bias[a]	The sensitivity analysis based on restricting to studies with low risk of bias showed that the pooled effect estimate lost statistical significance (RR 0.5, 95% CI 0.1 to 2.44).	The sensitivity analysis based on restricting to studies with low risk of bias showed that the pooled effect estimate did not lose statistical significance: MD was –28 minutes (95% CI –34 to –22).
Making plausible assumptions about the outcome of participants with missing data	The sensitivity analysis based on different assumptions for the outcomes of participants with missing data found that the pooled effect estimate lost statistical significance for about half of these analyses (Akl 2013; Akl 2015; Ebrahim 2013).	The sensitivity analysis based on different assumptions for the outcomes of participants with missing data found that the pooled effect estimate did not lose statistical significance for any of these (Akl 2013; Akl 2015; Ebrahim 2013).
Restricting the analysis to very long or large studies	NA	NA
Restricting by diagnostic criteria	NA	NA
Language of publication	NA	NA
Source of funding	NA	NA
Country of origin	NA	NA

[a]Studies judged at high risk of bias were: Bergenfelz 2005 (selection bias and reporting bias), Miccoli 1999 and Miccoli 2008 (reporting bias), Slepavicius 2008 (reporting bias) as assessed using the risk of bias graph (see Figure 2; Figure 3).
CI: confidence interval; MD: mean difference; NA: not applicable; RR: risk ratio.

APPENDICES

Appendix 1. Search strategies

Cochrane Central Register of Controlled Trials (Cochrane Register of Studies Online)
[Primary hyperparathyroidism]
1. MESH DESCRIPTOR Hyperparathyroidism, Primary
2. MESH DESCRIPTOR Parathyroid Neoplasms
3. (primary hyperparathyroidis*):TI,AB,KY
4. (parathyroid ADJ2 (adenoma* or neoplasm*)):TI,AB,KY
5. #1 OR #2 OR #3 OR #4
[Parathyroidectomy/surgery/neck exploration]
6. MESH DESCRIPTOR Parathyroidectomy
7. MESH DESCRIPTOR Minimally Invasive Surgical Procedures

(Continued)

8. MESH DESCRIPTOR Video-Assisted Surgery

9. MESH DESCRIPTOR Monitoring, Intraoperative

10. (intraoperative adj3 monitoring):TI,AB,KY

11. parathyroidectom*:TI,AB,KY

12. ((bilateral or neck or focus*) ADJ2 exploration):TI,AB,KY

13. (surg* ADJ4 (bilateral or unilateral or minimal* or conventional* or tradition*)):TI,AB,KY

14. (resection* ADJ4 (bilateral or unilateral or minimal* or conventional* or tradition*)):TI,AB,KY

15. #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14

16. #5 AND #15

MEDLINE (via OvidSP)

[Primary hyperparathyroidism]

1. Hyperparathyroidism, Primary/

2. Parathyroid Neoplasms/

3. (primary hyperparathyroidis*).tw.

4. (parathyroid adj2 (adenoma* or neoplasm*)).tw.

5. or/1-4

[Parathyroidectomy/surgery/neck exploration]

6. Parathyroidectomy/

7. Minimally Invasive Surgical Procedures/

8. Video-Assisted Surgery/

9. Monitoring, Intraoperative/

10. (intraoperative adj3 monitoring).tw.

11. parathyroidectom*.tw.

12. ((bilateral or neck or focus*) adj2 exploration).tw.

13. (surg* adj4 (bilateral or unilateral or minimal* or conventional* or tradition*)).tw.

14. (resection* adj4 (bilateral or unilateral or minimal* or conventional* or tradition*)).tw.

15. or/6-14

16. 5 and 15

[17-26: Cochrane Handbook 2008 RCT filter - sensitivity maximizing version]

17. randomised controlled trial.pt.

18. controlled clinical trial.pt.

19. randomi?ed.ab.

20. placebo.ab.

21. drug therapy.fs.

(Continued)

22. randomly.ab.

23. trial.ab.

24. groups.ab.

25. or/17-24

26. exp animals/ not humans/

27. 25 not 26

28. 16 and 27

ClinicalTrials.gov (advanced search)

Conditions: “primary hyperparathyroidism” OR “parathyroid neoplasm” OR “parathyroid neoplasms” OR “parathyroid adenoma” OR “parathyroid adenomas”

Interventions: parathyroidectomy OR minimally OR video OR intraoperative OR intraoperatively OR exploration OR surgery OR surgical OR resection

WHO ICTRP (standard search)

primary* AND hyperparathyroid* AND parathyroidectom* OR

primary* AND hyperparathyroid* AND minimally* OR

primary* AND hyperparathyroid* AND video* OR

primary*y AND hyperparathyroid* AND intraoperativ* OR

primary* AND hyperparathyroid* AND exploration* OR

primary* AND hyperparathyroid* AND surg* OR

primary* AND hyperparathyroid* AND resection* OR

parathyroid* AND adenoma* AND parathyroid*ectom* OR

parathyroid* AND adenoma* AND minimally* OR

parathyroid* AND adenoma* AND video* OR

parathyroid* AND adenoma* AND intraoperativ* OR

parathyroid* AND adenoma* AND exploration* OR

parathyroid* AND adenoma* AND surg* OR

parathyroid* AND adenoma* AND resection* OR

parathyroid* AND neoplasm* AND parathyroid*ectom* OR

parathyroid* AND neoplasm* AND minimally* OR

parathyroid* AND neoplasm* AND video* OR

parathyroid* AND neoplasm* AND intraoperativ* OR

parathyroid* AND neoplasm* AND exploration* OR

parathyroid* AND neoplasm* AND surg* OR

parathyroid* AND neoplasm* AND resection*

Appendix 2. 'Risk of bias' assessment

'Risk of bias' domains

Random sequence generation (selection bias due to inadequate generation of a randomised sequence)

For each included study, we described the method used to generate the allocation sequence in sufficient detail to allow an assessment of whether it should produce comparable groups.

- Low risk of bias: study authors achieved sequence generation using computer-generated random numbers or a random numbers table. Drawing of lots, tossing a coin, shuffling cards or envelopes, and throwing dice are adequate if an independent person performed this who was not otherwise involved in the study. We considered the use of the minimisation technique as equivalent to being random.
- Unclear risk of bias: insufficient information about the sequence generation process.
- High risk of bias: the sequence generation method was non-random or quasi-random (e.g. sequence generated by odd or even date of birth; sequence generated by some rule based on date (or day) of admission; sequence generated by some rule based on hospital or clinic record number; allocation by judgement of the clinician; allocation by preference of the participant; allocation based on the results of a laboratory test or a series of tests; or allocation by availability of the intervention).

Allocation concealment (selection bias due to inadequate concealment of allocation prior to assignment)

We described for each included study the method used to conceal allocation to interventions prior to assignment and we assessed whether intervention allocation could have been foreseen in advance of or during recruitment or changed after assignment.

- Low risk of bias: central allocation (including telephone, interactive voice-recorder, Internet-based and pharmacy-controlled randomisation); sequentially numbered drug containers of identical appearance; sequentially numbered, opaque, sealed envelopes.
- Unclear risk of bias: insufficient information about the allocation concealment.
- High risk of bias: used an open random allocation schedule (e.g. a list of random numbers); assignment envelopes used without appropriate safeguards; alternation or rotation; date of birth; case record number; any other explicitly unconcealed procedure.

We also evaluated study baseline data to incorporate assessment of baseline imbalance into the 'Risk of bias' judgement for selection bias (Corbett 2014).

Chance imbalances may also affect judgements on the risk of attrition bias. In the case of unadjusted analyses, we distinguished between studies that we rated at low risk of bias on the basis of both randomisation methods and baseline similarity, and studies that we judged at low risk of bias on the basis of baseline similarity alone (Corbett 2014). We reclassified judgements of unclear, low or high risk of selection bias as specified in Appendix 3.

Blinding of participants and study personnel (performance bias due to knowledge of the allocated interventions by participants and personnel during the study)

We evaluated the risk of detection bias separately for each outcome (Hróbjartsson 2013). We noted whether endpoints were self-reported, investigator-assessed or adjudicated outcome measures (see below).

- Low risk of bias: blinding of participants and key study personnel was ensured, and it was unlikely that the blinding could have been broken; no blinding or incomplete blinding, but we judged that the outcome was unlikely to have been influenced by lack of blinding.
- Unclear risk of bias: insufficient information about the blinding of participants and study personnel; the study did not address this outcome.
- High risk of bias: no blinding or incomplete blinding, and the outcome was likely to have been influenced by lack of blinding; blinding of study participants and key personnel attempted, but likely that the blinding could have been broken, and the outcome was likely to be influenced by lack of blinding.

Blinding of outcome assessment (detection bias due to knowledge of the allocated interventions by outcome assessment)

We evaluated the risk of detection bias separately for each outcome (Hróbjartsson 2013). We noted whether endpoints were self-reported, investigator-assessed or adjudicated outcome measures (see below).

- Low risk of bias: blinding of outcome assessment was ensured, and it was unlikely that the blinding could have been broken; no blinding of outcome assessment, but we judged that the outcome measurement was unlikely to have been influenced by lack of blinding.

(Continued)

- Unclear risk of bias: insufficient information about the blinding of outcome assessors; the study did not address this outcome.
- High risk of bias: no blinding of outcome assessment, and the outcome measurement was likely to have been influenced by lack of blinding; blinding of outcome assessment, but likely that the blinding could have been broken, and the outcome measurement was likely to be influenced by lack of blinding.

Incomplete outcome data (attrition bias due to quantity, nature or handling of incomplete outcome data)

For each included study or each outcome, or both, we described the completeness of data, including attrition and exclusions from the analyses. We stated whether the study reported attrition and exclusions, and we reported the number of participants included in the analysis at each stage (compared with the number of randomised participants per intervention/comparator groups). We also noted if the study reported the reasons for attrition or exclusion, and whether missing data were balanced across groups or were related to outcomes. We considered the implications of missing outcome data per outcome such as high dropout rates (e.g. above 15%) or disparate attrition rates (e.g. difference of 10% or more between study arms).

- Low risk of bias: no missing outcome data; reasons for missing outcome data unlikely to be related to true outcome (for survival data, censoring unlikely to introduce bias); missing outcome data balanced in numbers across intervention groups, with similar reasons for missing data across groups; for dichotomous outcome data, the proportion of missing outcomes compared with observed event risk was not enough to have a clinically relevant impact on the intervention effect estimate; for continuous outcome data, plausible effect size (mean difference or standardised mean difference) among missing outcomes was not enough to have a clinically relevant impact on observed effect size; appropriate methods, such as multiple imputation, were used to handle missing data.
- Unclear risk of bias: insufficient information to assess whether missing data in combination with the method used to handle missing data were likely to induce bias; the study did not address this outcome.
- High risk of bias: reason for missing outcome data was likely to be related to true outcome, with either imbalance in numbers or reasons for missing data across intervention groups; for dichotomous outcome data, the proportion of missing outcomes compared with observed event risk enough to induce clinically relevant bias in the intervention effect estimate; for continuous outcome data, plausible effect size (mean difference or standardised mean difference) among missing outcomes enough to induce clinically relevant bias in observed effect size; 'as-treated' or similar analysis done with substantial departure of the intervention received from that assigned at randomisation; potentially inappropriate application of simple imputation.

Selective reporting (reporting bias due to selective outcome reporting)

We assessed outcome reporting bias by integrating the results of the appendix 'Matrix of study endpoints (publications and trial documents)' (Boutron 2014; Jones 2015; Mathieu 2009), with those of the appendix 'High risk of outcome reporting bias according to the Outcome Reporting Bias In Trials (ORBIT) classification' (Kirkham 2010). This analysis formed the basis for the judgement of selective reporting.

- Low risk of bias: the study protocol was available and all the study's prespecified (primary and secondary) outcomes that were of interest to this review were reported in the prespecified way; the study protocol was unavailable, but it was clear that the published reports included all expected outcomes (ORBIT classification).
- Unclear risk of bias: insufficient information about selective reporting.
- High risk of bias: not all the study's prespecified primary outcomes were reported; one or more primary outcomes were reported using measurements, analysis method, or subsets of the data (e.g. subscales) that were not prespecified; one or more reported primary outcomes were not prespecified (unless clear justification for their reporting was provided, such as an unexpected adverse effect); one or more outcomes of interest in the Cochrane Review were reported incompletely so that we could not enter them into a meta-analysis; the study report failed to include results for a key outcome that we would expect to have been reported for such a study (ORBIT classification).

Other bias

- Low risk of bias: the study appeared free from other sources of bias.
- Unclear risk of bias: information was insufficient to assess whether an important risk of bias existed; insufficient rationale or evidence that an identified problem introduced bias.
- High risk of bias: the study had a potential source of bias related to the specific study design used; the study was claimed to be fraudulent or the study had some other serious problem.

Appendix 3. Selection bias decisions

Selection bias decisions for studies that reported unadjusted analyses: comparison of results obtained using method details alone versus results obtained using method details and study baseline information[a]

Reported randomisation and allocation concealment methods	Risk of bias judgement using methods reporting	Information gained from study characteristics data	Risk of bias using baseline information and methods reporting
Unclear methods	Unclear risk	Baseline imbalances present for important prognostic variable(s)	**High risk**
		Groups appear similar at baseline for all important prognostic variables	**Low risk**
		Limited or no baseline details	Unclear risk
Would generate a truly random sample, with robust allocation concealment	Low risk	Baseline imbalances present for important prognostic variable(s)	**Unclear risk** [b]
		Groups appeared similar at baseline for all important prognostic variables	Low risk
		Limited baseline details, showing balance in some important prognostic variables[c]	Low risk
		No baseline details	**Unclear risk**
Sequence was not truly randomised or allocation concealment was inadequate	High risk	Baseline imbalances present for important prognostic variable(s)	High risk
		Groups appeared similar at baseline for all important prognostic variables	**Low risk**
		Limited baseline details, showing balance in some important prognostic variables[c]	**Unclear risk**
		No baseline details	High risk

[a]Taken from Corbett 2014; judgements highlighted in bold indicate situations in which the addition of baseline assessments would change the judgement about risk of selection bias compared with using methods reporting alone.
[b]Imbalance was identified that appeared likely to be due to chance.
[c]Details for the remaining important prognostic variables are not reported.

Appendix 4. Description of interventions

Study ID	Intervention(s) and comparator(s) (route, frequency, total dose/day)
Bergenfelz 2002	**I:** unilateral neck exploration (MIP) Surgery was started on the same side as indicated by the preoperative scintigram. In case no enlarged parathyroid gland was visualised on scintigram, then the left side was automatically explored. The completeness of the surgical resection was only confirmed when the PTH levels de-

(Continued)

	creased by > 50% within 5 minutes or > 60% after 15 minutes with the PTH levels being obtained 5 and 15 minutes after adenoma removal.
	C: BNE Surgery was always started on the left side and comprehensive surgery was performed. Then an attempt was made to identify the 4 parathyroid glands, and any macroscopically enlarged parathyroid glands were removed and taken for frozen section. The decision to terminate surgery was based on the gross morphology of excised parathyroid glands in addition to the results of the frozen section. Intraoperative PTH levels were not measured.
Bergenfelz 2005	**I:** MIP Surgery was started on same side as indicated by the preoperative scintigram. Then the parathyroid adenoma was localised, dissected and sent for frozen-section analysis. The completeness of the surgical resection was only confirmed when the PTH levels decreased by > 50% within 5 minutes or > 60% after 15 minutes with the PTH levels being obtained 5 and 15 minutes after adenoma removal.
	C: BNE Participants had a short Kocher incision from the beginning. Surgery was always started on the left side and comprehensive BNE was done. Then an attempt was made to identify the 4 parathyroid glands, and any macroscopically enlarged parathyroid glands was removed and taken for frozen section. The participant remained under anaesthesia until the results of the frozen-section examination had been received. Intraoperative PTH levels were not measured.
Miccoli 1999	**I:** video-assisted parathyroidectomy (MIP) All surgeries were performed by the same surgeon. MIP surgery started with a 15-mm incision at the level of the notch, and this was followed with carbon dioxide insufflation (12 mmHg) under endoscopic vision. This technique enabled optimal visualisation of the surgical field. The parathyroid adenomas were identified with the help of needle-scopic instruments (2 mm). The completeness of the surgical resection was only confirmed when a decrease of ≥ 50% in intact PTH values from the highest pre-excision level was noted with the PTH levels being obtained 5 and 10 minutes after adenoma removal.
	C: conventional cervicotomy with BNE All surgeries were performed by the same surgeon. BNE was done under endotracheal general anaesthesia. Surgery was started with a traditional cervicotomy, which exposed the thyrotracheal groove, and then the laryngeal recurrent nerve was identified. Afterwards an attempt was made to identify the 4 parathyroid glands, and any macroscopically enlarged parathyroid glands were removed. Frozen section was used and no biopsy specimens were obtained from the normal parathyroid glands. Intraoperative PTH levels were not measured.
Miccoli 2008	**I:** focused endoscopic parathyroidectomy (MIVAP technique) (MIP) The MIP technique was carried through a central neck incision of 15–20 mm, 2 cm above the sternal notch, using an endoscope without gas insufflation. The completeness of the surgical resection was only confirmed when the PTH levels declined by ≥ 50% with respect to highest pre-excision level. The quick PTH assay was performed at baseline (before incision of the skin), at 5 and 10 minutes after the removal of the pathological gland.
	C: endoscopic parathyroidectomy plus BNE The BNE was performed via the same central neck access for a MIVAP procedure, with the procedure ended after visualisation of the 4 parathyroid glands and removal of the macroscopically enlarged glands. Intraoperative PTH levels were not measured.
Slepavicius 2008	**I:** focused parathyroidectomy (MIP)

(Continued)

	A 2- to 2.5-cm transverse incision was made. The incision for presumed inferior gland was placed 2 cm above the clavicle, whereas the one for presumed superior gland was placed higher. No attempts were made to visualise normal parathyroid glands. The completeness of the surgical resection was only confirmed when there was a decrease of ≥ 50% of intraoperative rapid PTH measurements from the baseline at 15 minutes, with the PTH levels being obtained 5 and 15 minutes after adenoma removal.
	C: conventional surgery group (BNE) Parathyroidectomy was performed with traditional Kocher incision and identification of all parathyroid glands was performed. No localisation preoperative examination was carried out. Intraoperative PTH levels were not measured.

BNE: bilateral neck exploration; C: comparator; I: intervention; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; PTH: parathyroid hormone.

Appendix 5. Baseline characteristics (I)

Study ID	Intervention(s) and comparator(s)	Duration of follow-up	Description of participants	Study period	Country	Setting	Ethnic groups (%)	Duration of primary hyperparathyroidism (mean years (SD))
Bergenfelz 2002	I: unilateral neck exploration (MIP)	6 weeks; 1 year and 5 years (extended follow-up after 5 years)	Participants with primary hyperparathyroidism	September 1996 to March 2001 (extended follow-up after 5 years	Sweden	Department of surgery	—	—
	C: BNE						—	—
Bergenfelz 2005	I: MIP	4 days; 1 and 6 months	Participants with primary hyperparathyroidism	February 1999 to September 2002	Germany	Department of surgery	—	—
	C: conventional bilateral cervical exploration (BNE)						—	—
Miccoli 1999	I: video-assisted parathyroidectomy (MIP)	12, 24 and 48 hours; 1, 3 and 6 months	Participants with primary hyperparathyroidism	March to November 1998	Italy	Department of surgery	—	—
	C: conventional cervicotomy with BNE					Department of endocrinology	—	—
Miccoli 2008	I: focused endoscopic parathyroidectomy (MIVAP technique) (MIP)	48 hours; 1 and 6 months	Participants with primary hyperparathyroidism	October 2005 to February 2006	Italy	Department of surgery	—	—
	C: endoscopic parathyroidectomy plus BNE						—	—
Slepavicius 2008	I: focused parathyroidectomy (MIP)	4, 8, 16, 24, 36 and 48 hours; 1 and 6 months; 1 year	Participants with primary hyperparathyroidism	February 2005 to February 2008	Lithuania	Department of abdominal and endocrine surgery	—	—
	C: BNE						—	—

—: denotes not reported.

(Continued)

BNE: bilateral neck exploration; C: comparator; I: intervention; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; SD: standard deviation.

Appendix 6. Baseline characteristics (II)

Study ID	Intervention(s) and comparator(s)	Sex (% women)	Age (mean/ range years (SD))	Preoperative serum calcium (mean mg/dL or mmol/ L / range (SD))	Preoperative serum PTH concentration (mean pg/mL or pmol/L / range (SD))	Preoperative serum vitamin D concentration (ng/mL)	Preoperative imaging results (% single gland disease)	Comedications/cointerventions	Comorbidities (% of participants)
Bergenfelz 2002	I: unilateral neck exploration (MIP)	83	66 (SD 14)	2.79 mmol/L (11.16 mg/ dL)	9.7 (SD 4.4) pmol/L 92.3 (SD 41.9) pg/mL	21.2	89	After randomisation participants in the unilateral group underwent preoperative sestamibi subtraction scintigraphy and preoperatively, all participants underwent indirect laryngoscopy.	—
	C: BNE	77	67 (15)	2.75 mmol/L (11 mg/ dL)	10.5 (SD 6.0) pmol/L 100 (SD 57.1) pg/mL	20	91	After randomisation participants in the unilateral group underwent preoperative sestamibi subtraction scintigraphy. No localisation procedure was performed in the bilateral group. Preoperatively, all participants underwent indirect laryngoscopy.	—
Bergenfelz 2005	I: MIP	80	57 (SD 15)	2.98 mmol/L (11.92 mg/ dL)	Intact PTH (relative value where 1 represents the upper normal limit of the respective assay used in different laboratories): 2.5	—	58	All participants underwent Sestambi scan prior to randomisation. Vocal cord function was assessed before surgery by indirect laryngoscopy. The completeness of the surgical resection was confirmed by a ≥ 50% decrease in intact PTH values, with respect to the highest pre-excision level; measurements were obtained 5 and 15 minutes after the removal of the adenoma.	—
	C: conventional bilateral cervical exploration (BNE)	76	62 (SD 12)	2.91 mmol/L (11.64 mg/ dL)	Intact PTH (relative value where 1 represents the upper normal limit	—	58	All participants underwent Sestambi scan prior to randomisation. Vocal cord function was assessed before surgery by indirect laryngoscopy.	—

(Continued)

					of the respective assay): 2.5				
Miccoli 1999	I: video-assisted parathyroidectomy (MIP)	65	48 (24 to 66) (SD 13)	11.1 mg/dL (10.1–13.4)	221 (90–379)	—	81	Either general endotracheal anaesthesia (18 participants) was used or bilateral superficial cervical block in association with laryngeal mask (2 participants). The completeness of the surgical resection was confirmed by a ≥ 50% decrease in intact PTH values, with respect to the highest pre-excision level; measurements were obtained 5 and 10 minutes after the removal of the adenoma	—
	C: conventional cervicotomy with BNE	61	60 (22–80) (SD 14)	10.8 mg/dL (10–12.9)	195 (70–320)	—	81	Participants underwent a bilateral exploration of the neck under endotracheal general anaesthesia. Intraoperative quick PTH assay was not measured.	—
Miccoli 2008	I: focused endoscopic parathyroidectomy (MIVAP) technique (MIP)	80	57	11.02 mg/dL	308.1	—	58	The quick PTH assay was performed at baseline (before incision of the skin), at 5 and 10 minutes after the removal of the pathological gland. The surgical procedure was ended when a decrease ≥ 50% of the highest preoperative value was reported by quick PTH assay.	—
	C: endoscopic parathyroidectomy plus BNE	85	60	11.34 mg/dL	320.5	—	—	Procedure ended after visualisation of the 4 parathyroid glands and removal of the macroscopically enlarged glands.	—
Slepavicius 2008	I: focused parathyroidectomy (MIP)	—	18–90	2.92 mmol/L (SD 0.17) 11.68 mg/dL (SD 0.68)	264.4 (SD 161.8)	—	—	Preoperative localisation studies before operation as well as intraoperative PTH monitoring were performed. Parathyroid scintigraphy was performed with 99m Tc99m-sestamibi for preoperative dual-phase sestamibi parathyroid scan of the neck and chest with planar images. All participants were consulted by ENT spe-	—

(Continued)

							cialist concerning evaluation of vocal folds function.	
C: BNE	—	18–90	2.98 mmol/L (SD 0.22) 11.92 mg/dL (SD 0.88)	236.9 (SD 90.5)	—	—	Localisation examination before surgery was not carried out. Intraoperative PTH monitoring was not performed. All participants were consulted by ENT specialist concerning evaluation of vocal folds function.	—

—: denotes not reported.

BNE: bilateral neck exploration; C: comparator; ENT: ear, nose and throat; I: intervention; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; PTH: parathyroid hormone; SD: standard deviation; VAP: video-assisted parathyroidectomy.

Appendix 7. Matrix of study endpoints (publications and trial documents)

Bergenfelz 2002	**Endpoints quoted in trial document(s)** (ClinicalTrials.gov, FDA/EMA document, manufacturer's website, published design paper)[a,b]
	Source: NT
	Endpoints quoted in publication(s)[b,c]
	Primary outcome measure: postoperative medication for hypocalcaemia during the first 4 postoperative days **Secondary outcome measures**: symptomatic hypocalcaemia, serum levels of calcium (severe hypocalcaemia defined as serum calcium < 2.00 mmol/L), persistent hyperparathyroidism, complications, operative time, cost **Other outcome measures**: —
	Endpoints quoted in abstract of publication(s)[b,c]
	Primary outcome measure: use of postoperative medication for hypocalcaemic symptoms **Secondary outcome measures**: — **Other outcome measures**: —
Bergenfelz 2005	**Endpoints quoted in trial document(s)** (ClinicalTrials.gov, FDA/EMA document, manufacturer's website, published design paper)[a,b]
	Source: NT
	Endpoints quoted in publication(s)[b,c]
	Primary outcome measure: serum levels of total calcium during the first 4 days after surgery **Secondary outcome measures**: number of conversions from MIP to BCE, operating time (skin–skin), complications (recurrent nerve palsy, wound infection, haematoma), serum level of calcium, as well as oral calcium and vitamin D supplementation recorded at 1 and 6 months after surgery **Other outcome measures**: —
	Endpoints quoted in abstract of publication(s)[b,c]
	Primary outcome measure: postoperative hypocalcaemia **Secondary outcome measures**: operating time, complications, postoperative analgesia, recurrent disease **Other outcome measures**: —
Miccoli 1999	**Endpoints quoted in trial document(s)** (ClinicalTrials.gov, FDA/EMA document, manufacturer's website, published design paper)[a,b]
	Source: NT
	Endpoints quoted in publication(s)[b,c]
	Primary outcome measure: —

(Continued)

	Secondary outcome measure: —
	Other outcome measures: operative time; postoperative pain evaluation by a visual analogue scale, 12, 24 and 48 hours postsurgery; fever > 38 °C; symptomatic hypocalcaemia; vocal cord disorders; time required to return to normal activities; cosmetic satisfaction (a personal opinion on the aesthetics of the scar (with a score ranging from 1 (poor) to 10 (excellent) assessed at 1, 3 and 6 months after surgery); cost-identification analysis and participants who were normocalcaemic 6 months after surgery
	Endpoints quoted in abstract of publication(s)[b,c]
	Primary outcome measure: —
	Secondary outcome measures: —
	Other outcome measures: operative time, postoperative pain, fever, hypocalcaemia, cosmetic result, costs
Miccoli 2008	**Endpoints quoted in trial document(s)** (ClinicalTrials.gov, FDA/EMA document, manufacturer's website, published design paper)[a,b]
	Source: NT
	Endpoints quoted in publication(s)[b,c]
	Primary outcome measure: —
	Secondary outcome measures: —
	Other outcome measures: mean operative time, postoperative complications (haemorrhage, laryngeal nerve palsy, hypocalcaemia), normal PTH and serum calcium levels at 1- and 6-month follow-up
	Endpoints quoted in abstract of publication(s)[b,c]
	Primary outcome measure: —
	Secondary outcome measures: —
	Other outcome measures: mean operative time, conversion to BNE, postoperative complications and eucalcaemic hyperparathyroidism
Slepavicius 2008	**Endpoints quoted in trial document(s)** (ClinicalTrials.gov, FDA/EMA document, manufacturer's website, published design paper)[a,b]
	Source: NT
	Endpoints quoted in publication(s)[b,c]
	Primary outcome measure: cure rate
	Secondary outcome measures: postoperative pain intensity, analgesics consumption, time of surgery, cosmetic satisfaction, quality of life, cost-effectiveness
	Other outcome measures: —
	Endpoints quoted in abstract of publication(s)[b,c]
	Primary outcome measure: —

(Continued)

Secondary outcome measures: —

Other outcome measures: operative time, pain intensity, consumption of analgesics, scar length, cosmetic satisfaction, cost

—: denotes not reported.
[a]Trial document(s) refers to all available information from published design papers and sources other than regular publications (e.g. FDA/EMA documents, manufacturer's websites, trial registers).
[b]Primary and secondary outcomes refer to verbatim specifications in publication/records. Unspecified outcome measures refer to all outcomes not described as primary or secondary outcome measures.
[c]Publication(s) refers to study information published in scientific journals (primary reference, duplicate publications, companion documents or multiple reports of a primary study).

BCE: bilateral cervical exploration; EMA: European Medicines Agency; FDA: Food and Drug Administration (US); MIP: minimally invasive parathyroidectomy; NT: no trial document available; PTH: parathyroid hormone.

Appendix 8. High risk of outcome reporting bias according to Outcome Reporting Bias In Trials (ORBIT) classification

Study ID	Outcome	High risk of bias (category A)[a]	High risk of bias (category D)[b]	High risk of bias (category E)[c]	High risk of bias (category G)[d]
Bergenfelz 2002	Postoperative medication for hypocalcaemia during the first 4 postoperative days	ND			
	Symptomatic hypocalcaemia	ND			
	Complications	ND			
	Operative time	ND			
	Socioeconomic effects	ND			
	Severe hypocalcaemia	No	**Yes**	No	No
	Persistent hyperparathyroidism	No	**Yes**	No	No
Bergenfelz 2005	Serum levels of total calcium during the first 4 days after surgery	ND			
	Number of conversions from MIP to BCE	NA			
	Serum level of calcium	ND			
	Oral calcium and vitamin D supplementation recorded at 1 and 6 months after surgery	No	**Yes**	No	No
	Complications (recurrent nerve palsy, wound infection, haematoma)	No	No	**Yes**	No
Miccoli 1999	Authors did not specify outcomes in general, neither primary nor secondary outcomes	No	No	No	**Yes**

(Continued)

Miccoli 2008	Authors did not specify outcomes in general, neither primary nor secondary outcomes	No	No	No	**Yes**
Slepavicius 2008	Primary outcome: cure rate	ND			
	Postoperative pain intensity	ND			
	Analgesics consumption	ND			
	Time of surgery	ND			
	Cosmetic satisfaction	ND			
	Cost effectiveness	ND			
	Health-related quality of life	**Yes**	No	No	No

[a]Clear that outcome was measured and analysed; study report stated that outcome was analysed but reported only that result was not significant (Classification 'A', table 2, Kirkham 2010).
[b]Clear that outcome was measured and analysed; study report stated that outcome was analysed but reported no results (Classification 'D', table 2, Kirkham 2010).
[c]Clear that outcome was measured but was not necessarily analysed; judgement said likely to have been analysed but not reported due to non-significant results (Classification 'E', table 2, Kirkham 2010).
[d]Unclear whether outcome was measured; not mentioned, but clinical judgement said likely to have been measured and analysed but not reported on the basis of non-significant results (Classification 'G', table 2, Kirkham 2010).

BCE: bilateral cervical exploration; MIP: minimally invasive parathyroidectomy; ND: none detected.

Appendix 9. Definition of endpoint measurement[a]

Study ID	Endpoints	Definition
Bergenfelz 2002	**Success rate**	ND. Comment: success rate was not clearly defined but the term cure rate was used and serum levels of calcium and phosphorous were taken on days 1–4 and at 6 weeks postsurgery (IO).
	Total incidence of perioperative adverse events	ND
	Bone fracture rate	NR
	Nephrolithiasis rate	NR
	Absence from work	NR
	Postoperative increase in PTH with eucalcaemia	Increased PTH in combination with normal calcium levels (IO).
	Duration of surgery	Duration of surgery was defined as mean operative time (IO).
	Length of hospital stay	NR
	Cosmetic satisfaction	NR

(Continued)

	Conversion rate from minimally invasive to open procedure	ND
	Health-related quality of life	NR
	Bleeding events	ND. Comment: it was not clearly defined, however, in the study 1 participant had to be reoperated for rebleeding.
	Infection	NR
	Hypocalcaemia	Symptomatic hypocalcaemia was defined as having symptoms of tingling in the lips, fingers and toes, and severe hypocalcaemia was defined as serum calcium < 2.00 mmol/L.
	Postoperative pain	Pain was evaluated using the visual analogue scale and was evaluated during the first 4 postoperative days (SO).
	Vocal cord paralysis	Vocal cord disorders were assessed using indirect laryngoscopy (AO).
	Anaesthetic related complications	NR
	All-cause mortality	Mortality was defined as death and was assessed in the follow-up study at 5-year period but the cause of death was not explicitly stated.
	Absence from work	NR
	Socioeconomic effects	Socioeconomic costs were defined as the costs calculated from official in-hospital charges for services performed by different departments. The cost for sestamibi scintigraphy was USD 134, intraoperative PTH USD 126, frozen section USD 155 and time for anaesthesia USD 12/minute.
	Severe/serious adverse events	Serious adverse events were presented in a table. Serious adverse events were reported in 4 participants in the BNE group: in 1 participant vitamin D substitution therapy was needed at 1 year after surgery. In 1 participant with rheumatic arthritis, postoperative acute tracheal oedema was observed with no response to conservative treatment and there was a need for an emergency operation with tracheostomy. The same participant had temporary paresis of the recurrent laryngeal nerve and temporary vitamin D substitution therapy due to hypocalcaemia. In 1 participant, tracheal oedema occurred on the first postoperative day. 1 participant was operated due to rebleeding. Temporary paresis of the recurrent laryngeal nerve in 2 participants. 1 participant needed temporary vitamin D substitution therapy due to hypocalcaemia. 2 participants who were assigned to the MIP group had serious adverse events, both of which were temporary paresis of the recurrent laryngeal nerve (IO).
Bergenfelz 2005	**Success rate**	ND. Comment: success rate was not clearly defined but serum levels of total and ionised calcium, intact PTH, creatinine, alkaline phosphatase and osteocalcin were measured at 4 days, 1 month and 6 months after surgery (IO).
	Total incidence of perioperative adverse events	ND. Comment: perioperative adverse events was not clearly defined but complications such as recurrent nerve palsy, wound infection and haematoma were recorded at 1 and 6 months after surgery (AO).

(Continued)

	Bone fracture rate	NR
	Nephrolithiasis rate	NR
	Absence from work	NR
	Postoperative increase in PTH with eucalcaemia	NR
	Duration of surgery	Duration of surgery was defined as the total time required for the operation (skin to skin), the time required to finish the parathyroid exploration, the time required to dissect and excise the parathyroid adenoma following its identification by parathyroid exploration, and the time until the results of the frozen-section analysis or quick IOPTH assay (or both) were available. In the MIP group, it was defined as the time needed for local anaesthesia in addition to the total operating time (IO)
	Length of hospital stay	NR
	Cosmetic satisfaction	Cosmetic satisfaction was defined as participant's satisfaction regarding the cosmetic result of the surgery (SO).
	Conversion rate from minimally invasive to open procedure	ND. Comment: conversion rate was not defined but conversion from MIP to BNE was allowed in certain situations.
	Health-related quality of life	Quality of life was assessed using the 36-item Short Form questionnaire (SO).
	Bleeding events	ND. Comment: haematoma was not clearly defined but recorded at 1 and 6 months after surgery.
	Infection	ND. Comment: infection was not clearly defined but recorded at 1 and 6 months after surgery.
	Hypocalcaemia	ND. Comment: hypocalcaemia was not clearly defined but the study mentioned that oral calcium was administered freely whenever a participant reported symptoms of hypocalcaemia.
	Postoperative pain	Postoperative pain was defined as pain requiring analgesia (SO).
	Vocal cord paralysis	Vocal cord disorders were assessed using indirect laryngoscopy (AO).
	Anaesthetic-related complications	NR
	All-cause mortality	NR
	Absence from work	NR
	Socioeconomic effects	NR
	Severe/serious adverse events	NR
Miccoli 1999	**Success rate**	ND. Comment: success rate was not clearly defined but normocalcaemic participants were reported at 6 months (IO).

(Continued)

Total incidence of perioperative adverse events	ND
Bone fracture rate	NR
Nephrolithiasis rate	NR
Absence from work	ND. Comment: postoperative inactivity period was mentioned but it was not specifically mentioned as time needed to return to work (IO).
Postoperative increase in PTH with eucalcaemia	NR
Duration of surgery	Comment: in the BNE group, surgery was not considered done up until the time when the pathologist reported the result of the frozen section and 4 parathyroid glands were identified. In the MIP group, it was the time until the laboratory technician reported the result of the IPTH (IO).
Length of hospital stay	NR
Cosmetic satisfaction	Cosmetic satisfaction was defined as participant's satisfaction regarding the cosmetic result with a score provided ranging from 1 (poor) to 10 (excellent) and evaluated at 1, 3 and 6 months after surgery were investigated (SO).
Conversion rate from minimally invasive to open procedure	NR
Health-related quality of life	NR
Bleeding events	NR
Infection	ND. Comment: infection per se was not defined, however, wound infection was mentioned in the study.
Hypocalcaemia	ND
Postoperative pain	Pain was evaluated using the visual analogue scale with a score ranging from 1 (no pain) to 10 (worst pain ever) and was evaluated at 12, 24 and 48 hours after the operation (SO).
Vocal cord paralysis	ND. Comment: vocal cord disorders were not defined but recorded 48 hours and 6 months after surgery (AO).
Anaesthetic-related complications	NR
All-cause mortality	NR
Absence from work	NR
Socioeconomic effects	Socioeconomic costs were defined as the costs linked with the adoption of new technologies; which included the use of laparoscopic equipment, the necessity of preoperative localisation studies done in cases of minimally invasive approaches mainly (IO).

(Continued)

	Severe/serious adverse events	NR
Miccoli 2008	**Success rate**	ND. Comment: success rate was not clearly defined but serum levels of PTH and calcium were measured at 48 hours, 1 month and 6 months after surgery (IO).
	Total incidence of perioperative adverse events	ND. Comment: perioperative adverse events was not clearly defined but haemorrhage, laryngeal nerve palsy and hypocalcaemia were reported up until 48 hours after surgery (SO).
	Bone fracture rate	NR
	Nephrolithiasis rate	NR
	Absence from work	NR
	Postoperative increase in PTH with eucalcaemia	NR
	Duration of surgery	Duration of surgery was defined as the mean operative time (IO).
	Length of hospital stay	NR
	Cosmetic satisfaction	NR
	Conversion rate from minimally invasive to open procedure	The term conversion to cervicotomy was used.
	Health-related quality of life	ND
	Bleeding events	Bleeding was defined as haemorrhage.
	Infection	NR
	Hypocalcaemia	ND. Comment: hypocalcaemia symptoms were not clearly defined but were mentioned.
	Postoperative pain	NR
	Vocal cord paralysis	ND. Comment: the term laryngeal nerve palsy was used.
	Anaesthetic-related complications	NR
	All-cause mortality	NR
	Absence from work	NR
	Socioeconomic effects	ND. Comment: the word costs was mentioned but not defined.
	Severe/serious adverse events	NR

(Continued)

Slepavicius 2008	**Success rate**	ND. Comment: success rate was not clearly defined but serum levels of PTH and calcium were measured on the second day postsurgery, and after 1 and 6 months after surgery (IO).
	Total incidence of perioperative adverse events	Perioperative adverse events were defined as intrasurgery complications.
	Bone fracture rate	NR
	Nephrolithiasis rate	NR
	Absence from work	NR
	Postoperative increase in PTH with eucalcaemia	Increased PTH in combination with normal calcium levels.
	Duration of surgery	Duration of surgery was defined as the time elapsed from adenoma detection, adenoma excision time after its detection, and the surgery time from skin to skin, and the time until the results of the frozen section.
	Length of hospital stay	NR
	Cosmetic satisfaction	Cosmetic satisfaction was evaluated by numeric modified Hollander scale, which ranges from 0 to 7, where 0 = optimal result and 1–7 = suboptimal result (SO).
	Conversion rate from minimally invasive to open procedure	ND
	Health-related quality of life	Quality of life was assessed using the 36-item Short Form questionnaire (SO).
	Bleeding events	NR
	Infection	NR
	Hypocalcaemia	ND. Comment: hypocalcaemia symptoms were not clearly defined but it was mentioned that calcium and vitamin D preparations were given in case of occurrence of symptoms of hypocalcaemia.
	Postoperative pain	Pain was evaluated using visual analogue scale. The intensity of pain was evaluated from 0 (pain is absent) to 100 (unbearable pain) points (SO).
	Vocal cord paralysis	Postoperative vocal cord dysfunction was assessed by ear, nose and throat specialist at 1 and 6 months after surgery (AO).
	Anaesthetic-related complications	NR
	All-cause mortality	NR
	Absence from work	NR
	Socioeconomic effects	Costs were provided as costs needed for ultrasonography, Sestamibi scintigraphy, preoperative IIPTH from internal jugular vein, pre-incision and postexcision IIPTH, analgesics and hypocalcaemia treatment.

(Continued)

Severe/serious adverse events	NR

[a]In addition to definition of endpoint measurement, description who measured the outcome (**AO**: adjudicated outcome measurement; **IO**: investigator-assessed outcome measurement; **SO**: self-reported outcome measurement).

BNE: bilateral neck exploration; IIPTH: intraoperative intact parathyroid hormone; IOPTH: intraoperative parathyroid hormone; MIP: minimally invasive parathyroidectomy; ND: not defined; NR: not reported; PTH: parathyroid hormone.

Appendix 10. Adverse events (I)

Study ID	Intervention(s) and comparator(s)	Participants included in analysis (n)	Deaths (n)	Deaths (% of participants)	Participants with ≥ 1 adverse event (n)	Participants with ≥ 1 adverse event (%)	Participants with ≥ 1 severe/serious adverse event (n)	Participants with ≥ 1 severe/serious adverse event (%)
Bergenfelz 2002	I: unilateral neck exploration (MIP)	47	16[a]	17.5	14	29.8	2	4.3
	C: BNE	44			27	61.4	5	11.3
Bergenfelz 2005	I: MIP	25	0	0	4	16	0	0
	C: conventional bilateral cervical exploration (BNE)	25	0	0	4	16	0	0
Miccoli 1999	I: video-assisted parathyroidectomy (MIP)	20	0	0	3	15	1	5
	C: conventional cervicotomy with bilateral exploration (BNE)	18	0	0	8	44.4	0	0
Miccoli 2008	I: focused endoscopic parathyroidectomy (MIVAP) technique plus quick PTH assay (MIP)	20	0	0	0	0	0	0
	C: endoscopic parathyroidectomy plus bilateral exploration (BNE)	20	0	0	0	0	0	0
Slepavicius 2008	I: focused parathyroidectomy (MIP)	24	0	0	3	12.5	0	0
	C: BNE	23	0	0	5	21.7	0	0

[a]At 5 years' follow-up.

BNE: bilateral neck exploration; C: comparator; I: intervention; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; n: number of participants.

Appendix 11. Adverse events (II)

Study ID	Intervention(s) and comparator(s)	Participants included in analysis (n)	Participants discontinuing study due to an adverse event (n)	Participants discontinuing study due to an adverse event (%)	Severe hypocalcaemia (n)	Severe hypocalcaemia (%)	Mild to moderate hypocalcaemia (n)	Mild to moderate hypocalcaemia (%)
Bergenfelz 2002	I: unilateral neck exploration (MIP)	47	0	0	3	7.3	—	—
	C: BNE	44	0	0	10	25	—	—
Bergenfelz 2005	I: MIP	25	0	0	0	0	17 (requested calcium)	68
	C: conventional bilateral cervical exploration (BNE)	25	0	0	0	0	15 (requested calcium)	60
Miccoli 1999	I: video-assisted parathyroidectomy (MIP)	20	0	0	0	0	1	5
	C: conventional cervicotomy with bilateral exploration (BNE)	18	0	0	0	0	3	16.6
Miccoli 2008	I: focused endoscopic parathyroidectomy (MIVAP) technique plus quick PTH assay (MIP)	20	0	0	0	0	0	0
	C: endoscopic parathyroidectomy plus bilateral exploration (BNE)	20	0	0	0	0	0	0
Slepavicius 2008	I: focused parathyroidectomy (MIP)	24	0	0	0	0	2	8.3
	C: BNE	23	0	0	0	0	4	17.4

—: denotes not reported

BNE: bilateral neck exploration; C: comparator; I: intervention; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; n: number of participants.

Appendix 12. Adverse events (III)

Study ID	Intervention(s) and comparator(s)	Participants included in analysis (n)	Laryngeal nerve injury (n)	Laryngeal nerve injury (%)	Haematoma (n)	Haematoma (%)	Infection (n)	Infection (%)
Bergenfelz 2002	I: unilateral neck exploration (MIP)	47	2	4.3	0	0	—	—
	C: BNE	44	1	2.3	1	2.3	—	—
Bergenfelz 2005	I: MIP	25	1	4	—	—	0	0
	C: conventional bilateral cervical exploration (BNE)	25	0	0	—	—	1	4
Miccoli 1999	I: video-assisted parathyroidectomy (MIP)	20	1	5	—	—	0	5
	C: conventional cervicotomy with bilateral exploration (BNE)	18	0	0	—	—	1	5.5
Miccoli 2008	I: focused endoscopic parathyroidectomy (MIVAP) technique plus quick PTH assay (MIP)	20	—	—	0	0	—	—
	C: endoscopic parathyroidectomy plus bilateral exploration (BNE)	20	—	—	0	0	—	—
Slepavicius 2008	I: focused parathyroidectomy (MIP)	24	1	4.1	—	—	—	—
	C: BNE	23	1	4.3	—	—	—	—

—: denotes not reported.

BNE: bilateral neck exploration; C: comparator; I: intervention; MIP: minimally invasive parathyroidectomy; MIVAP: minimally invasive video-assisted parathyroidectomy; n: number of participants.

Appendix 13. Checklist to aid consistency and reproducibility of GRADE assessments

Items		(1) Success rate	(2) Total incidence of perioperative adverse events	(3) All-cause mortality	(4) Health-related quality of life	(5) Cosmetic satisfaction	(6) Duration of surgery	(7) Length of hospital stay
Trial limitations (risk of bias)[a]	Was random sequence generation used (i.e. no potential for selection bias)?	Yes	Yes	No study explicitly reported on the occurrence of perioperative mortality. In 1 study with 5 years' follow-up, there were 16 deaths, no data per intervention group were reported.	2 studies reported that there were no clear differences between intervention groups; however, data were not presented	Yes	Yes	Not reported
	Was allocation concealment used (i.e. no potential for selection bias)?	Unclear	Unclear			Unclear	Unclear	
	Was there blinding of participants and personnel (i.e. no potential for performance bias) or outcome not likely to be influenced by lack of blinding?	No (↓)	No (↓)			Unclear	No (↓)	
	Was there blinding of outcome assessment (i.e. no potential for detection bias) or was outcome measurement not likely to be influenced by lack of blinding?	No (↓)	Unclear			Unclear	Unclear	
	Was an objective outcome used?	Yes	Yes			No (↓)	Yes	
	Were > 80% of participants enrolled in studies included in the analysis (i.e. no potential reporting bias)?[b]	Yes	Yes			Yes	Unclear	
	Were data reported consistently for the outcome of interest (i.e. no potential selective reporting)?	No (↓)	No (↓)			No (↓)	No (↓)	
	No other biases reported (i.e. no potential of other bias)?	Unclear	Unclear			Unclear	Unclear	
	Did the studies end up as scheduled (i.e. not stopped early)?	Yes	Yes			Yes	Yes	

(Continued)

Inconsistency[c]	Point estimates did not vary widely?	Yes	Yes	NA	Yes
	To what extent did confidence intervals overlap (substantial: all confidence intervals overlap ≥ 1 of the included studies point estimate; some: confidence intervals overlap but not all overlap ≥ 1 point estimate; no: ≥ 1 outlier: where the confidence interval of some of the studies do not overlap with those of most included studies)?	Substantial	Substantial	NA	Substantial
	Was the direction of effect consistent?	Yes	Yes	NA	Yes
	What was the magnitude of statistical heterogeneity (as measured by I^2): low (I^2 < 40%), moderate (I^2 40–60%), high I^2 > 60%)?	Low	Low	NA	High (↓)
	Was the test for heterogeneity statistically significant (P < 0.1)?	Not statistically significant	Not statistically significant	NA	Statistically significant (↓)
Indirectness	Were the populations in included studies applicable to the decision context?	Highly applicable in 2 studies, applicable in 3 studies (the concerns were: participants with multiglandular disease were excluded in 2 studies and preoperative localisation was used prior to randomisation in 3 studies. Multiglandular disease represents up to 20% of the pathol-	Highly applicable in 2 studies, applicable in 3 studies (the concerns were: participants with multiglandular disease were excluded in 2 studies and preoperative localisation was used prior to randomisation in 3 studies. Multiglandular disease represents up to 20% of the pathol-	NA	Highly applicable

(Continued)

		ogy in participants with hyperparathyroidism, and participants receiving BNE do not routinely undergo preoperative imaging. Participants with multiglandular hyperparathyroidism but with concomitant false imaging results showing single gland disease will likely undergo MIP in real life and their exclusion from studies is not reflective of real-life outcomes).	ogy in participants with hyperparathyroidism, and participants receiving BNE do not routinely undergo preoperative imaging. Participants with multiglandular hyperparathyroidism but with concomitant false imaging results showing single gland disease will likely undergo MIP in real life and their exclusion from studies is not reflective of real-life outcomes).		
	Were the interventions in the included studies applicable to the decision context?	Highly applicable in 2 studies, applicable in 3 studies (preoperative localisation was used prior to randomisation in 3 studies)	Highly applicable in 2 studies, applicable in 3 studies (preoperative localisation was used prior to randomisation in 3 studies)	NA	Highly applicable
	Was the included outcome not a surrogate outcome?	No (↓)	Yes	No (↓)	Yes
	Was the outcome timeframe sufficient?	Yes	Yes	Yes	Yes
	Were the conclusions based on direct comparisons?	Yes	Yes	Yes	Yes

(Continued)

Imprecision[d]	Was the confidence interval for the pooled estimate not consistent with benefit and harm?	No (↓)	Yes	NA	Yes
	What is the magnitude of the median sample size (high: > 300 participants, intermediate: 100–300 participants, low: < 100 participants)?[b]	Low (↓)	Low (↓)	Low (↓)	Low (↓)
	What was the magnitude of the number of included studies (large: > 10 studies, moderate: 5–10 studies, small: < 5 studies)?[b]	Moderate	Moderate	Small (↓)	Small (↓)
	Was the outcome a common event (e.g. occurs > 1/100)?	Yes	Yes	NA	Yes
Publication bias[e]	Was a comprehensive search conducted?	Yes	Yes	Yes	Yes
	Was grey literature searched?	Yes	Yes	Yes	Yes
	Were no restrictions applied to study selection on the basis of language?	Yes	Yes	Yes	Yes
	There was no industry influence on studies included in the review?	Yes	Yes	Yes	Yes
	There was no evidence of funnel plot asymmetry?	NA	NA	Unclear	Unclear
	There was no discrepancy in findings between published and unpublished studies?	Unclear	Unclear	Unclear	Unclear

[a]Questions on risk of bias are answered in relation to the majority of the aggregated evidence in the meta-analysis rather than to individual studies.
[b]Depends on the context of the systematic review area.
[c]Questions on inconsistency are primarily based on visual assessment of forest plots and the statistical quantification of heterogeneity based on the I^2 statistic.
[d]When judging the width of the confidence interval it is recommended to use a clinical decision threshold to assess whether the imprecision is clinically meaningful.
[e]Questions address comprehensiveness of the search strategy, industry influence, funnel plot asymmetry and discrepancies between published and unpublished studies.

(Continued)

(↓): key item for potential downgrading the certainty of the evidence (GRADE) as shown in the footnotes of the 'Summary of finding' table(s); BNE: bilateral neck exploration; MIP: minimally invasive parathyroidectomy; NA: not applicable.

Appendix 14. Health-related quality of life: instruments

	Instrument	Dimensions (subscales) (number of items)	Validated instrument	Answer options	Scores	Score range	Weighting of scores	Direction of scales	Minimal important difference
	SF-36 (G) Employed in: Bergenfelz 2005; Slepavicius 2008	SF-36 is a 36-item participant-reported survey of participant health consisting of an 8-point score	Yes	SF-36 has 8 sections: vitality, physical functioning, bodily pain, general health perceptions, physical role conditioning, emotional role conditioning, social role conditioning and mental health	SF-36 is an 8-point score, which are the weighted sums of questions in their section. Each scale is directly transformed into 0–100 scale on the assumption that each question carries equal weight. Lower score means more the disability, higher score means less disability	0–100	Yes The overall score consists of the weighted sums of questions in each of the sections	Higher values reflect better health	Minimal clinically important difference varies depending on the outcome of interest, and can be determined using various methods such as distribution-based, anchor-based or with the Delphi method. We were unable to identify a validated cut-off for a minimally clinically important difference in people undergoing parathyroidectomy. Furthermore, neither of the 2 studies evaluating SF-36 found any substantial or marked differences between the 2 groups

G: generic; SF-36: 36-item Short Form health survey.

Appendix 15. Survey of study investigators providing information on included studies

Study ID	Date study author contacted	Date study author replied	Summary of request for additional information	Reply
Bergenfelz 2002	11 June 2013	30 January 2014	1. The paper reported that 2 participants died within 1 year and 16 within 5 years of the study. Which treatment groups were they assigned to? 2. Were there any available data about the following outcomes: health-related quality of life, cosmetic satisfaction, bone fracture rate, BMD loss, symptomatic nephrolithiasis, absenteeism from work, length of hospitalisation, infections, and postoperative pain score? 3. Was any specific method of randomisation used? 4. Were participants or outcome assessors blinded? 5. Why were hypocalcaemia results only reported for 43 MIP participants and 39 BNE participants even though 47 and 44 were initially randomised to each group respectively?	1. Data were not available. 2. Answer not provided. 3. Author replied by email and confirmed randomisation through block randomisation by a computer software. Used sequentially numbered, concealed, opaque envelopes. 4. Author replied by email indicating that participants and outcome assessors were blinded to study group. 5. Answer not provided
Bergenfelz 2005	11 June 2013	30 January 2014	1. Were mean and standard deviation for the time of surgery from skin to skin if available? 2. Were there any available data about the following outcomes: economic cost, health-related quality of life, bone fracture rate, BMD loss, symptomatic nephrolithiasis, absence from work, length of hospitalisation, number of participants who had mild and severe types of hypocalcaemia and postoperative pain score, postoperative eucalcaemia hyperparathyroidism, cosmetic satisfaction by a scale, and long-term success rate? 3. What was the specific timing for the first postoperative laryngoscopy? 4. What randomisation process and, if concealed, opaque envelopes were used? 5. Were participants or outcome assessors blinded? 6. Were all surgeons performing operations experienced?	1. Answer not provided. 2. Answer not provided. 3. Answer not provided. 4. Author replied by email and confirmed randomisation through block randomisation by computer software. Author confirmed use of sequentially numbered, concealed, opaque envelopes. 5. Author replied by email confirming that participants and outcome assessors were not blinded to the treatment group due to the use of local anaesthesia.

(Continued)

				6. Authors confirmed that all surgeons performing operations were experienced, but did not provide the level of experience and familiarity with each technique
Miccoli 1999	11 June 2013	No reply	1. Were there any available data about the following outcomes: conversion rate from minimally invasive to open parathyroidectomy, postoperative increase in PTH with eucalcaemia, all-cause mortality, health-related quality of life, bone fracture rate, BMD loss, nephrolithiasis rate and length of hospitalisation? 2. Were participants or outcome assessors blinded? 3. Were all surgeons performing operations experienced?	NA
Miccoli 2008	11 June 2013	No reply	1. Were there any available data about the following outcomes: health-related quality of life, cosmetic satisfaction, bone fracture rate, BMD loss, nephrolithiasis rate, absence from work, length of hospitalisation, and postoperative pain score? 2. Were outcome assessors blinded? 3. Were all surgeons performing operations experienced?	NA
Slepavicius 2008	11 June 2013	No reply	1. Were there any available data about the following outcomes: conversion rate from MIP to BNE, postoperative increase in PTH with eucalcaemia, bone fracture rate, BMD loss, nephrolithiasis rate, absence from work and length of hospital stay? 2. Were participants or outcome assessors blinded? 3. Were all surgeons performing operations experienced?	NA

BMD: bone mineral densitometry; BNE: bilateral neck exploration; MIP: minimally invasive parathyroidectomy; NA: not applicable; PTH: parathyroid hormone.

HISTORY

Protocol first published: Issue 10, 2013
Review first published: Issue 10, 2020

CONTRIBUTIONS OF AUTHORS

Hala Ahmadieh (HA): study selection, data extraction, data analysis, data interpretation, review draft, finalisation of submitted manuscript and future review update.

Omar Kreidieh (OK): protocol draft, search strategy development, acquiring study reports, study selection, data extraction, data analysis and data interpretation.

Elie A Akl (EAA): methodology input, protocol draft, search strategy development, data analysis, data interpretation, review draft and future review update.

Ghada El-Hajj Fuleihan (GEHF): content expert, protocol draft, search strategy development, oversight of study selection, data extraction and data analysis, data interpretation, review draft, finalisation of submitted manuscript and future review update.

DECLARATIONS OF INTEREST

HA: none.

OK: none.

EAA: none.

GEHF: has received funding as primary investigator from the American University of Beirut to conduct an investigator-initiated vitamin D trial in pregnancy. She is also a co-primary investigator on an investigator-initiated protocol to investigate vitamin D supplementation in patients post-bariatric surgery. Has received royalties as an UpToDate contributor and payment for development of educational presentations (MTP on hyperparathyroidism).

SOURCES OF SUPPORT

Internal sources

- American University of Beirut, Faculty of Medicine, Lebanon

  Medical Resource Package Plan

External sources

- No sources of support supplied

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Because of the long time period between publication of the protocol (Kreidieh 2013) and review, all major parts of the review had to be adapted to the newest Cochrane and Cochrane Metabolic and Endocrine Disorders Group standards.

NOTES

Portions of the background and methods sections, the appendices, additional tables and figures 1 to 3 of this Cochrane Review are based on a standard template established by the Cochrane Metabolic and Endocrine Disorders Group.

INDEX TERMS

Medical Subject Headings (MeSH)

Bias; Hyperparathyroidism, Primary [blood] [*diagnostic imaging] [*surgery]; Hypocalcemia [epidemiology]; Minimally Invasive Surgical Procedures [adverse effects] [methods]; Monitoring, Intraoperative [*methods]; Neck [surgery]; Neck Dissection [methods]; Operative Time; Parathyroid Hormone [*blood]; Parathyroidectomy [adverse effects] [*methods]; Postoperative Complications [epidemiology]; Quality of Life; Vocal Cord Paralysis [epidemiology]

MeSH check words

Adult; Humans