National Institute for Health and Care Excellence

Draft for consultation

# Hyperparathyroidism (primary): diagnosis, assessment and initial management

## [C] Evidence review for Indications for surgery

*NICE guideline*

*Intervention evidence review*

*November 2018*



*Draft for consultation*

*This evidence review was developed by the National Guideline Centre*

**Disclaimer**

The recommendations in this guideline represent the view of NICE, arrived at after careful consideration of the evidence available. When exercising their judgement, professionals are expected to take this guideline fully into account, alongside the individual needs, preferences and values of their patients or service users. The recommendations in this guideline are not mandatory and the guideline does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and, where appropriate, their carer or guardian.

Local commissioners and providers have a responsibility to enable the guideline to be applied when individual health professionals and their patients or service users wish to use it. They should do so in the context of local and national priorities for funding and developing services, and in light of their duties to have due regard to the need to eliminate unlawful discrimination, to advance equality of opportunity and to reduce health inequalities. Nothing in this guideline should be interpreted in a way that would be inconsistent with compliance with those duties.

NICE guidelines cover health and care in England. Decisions on how they apply in other UK countries are made by ministers in the Welsh Government, Scottish Government, and Northern Ireland Executive. All NICE guidance is subject to regular review and may be updated or withdrawn.

# Contents

# 1 Indications for surgery

## 1.1 Review question

### 1.1.1 What is the clinical and cost effectiveness of surgery (parathyroidectomy) in people with primary hyperparathyroidism?

### 1.1.2 What are the indications for surgery (parathyroidectomy) in people with primary hyperparathyroidism?

## 1.2 Introduction

There is considerable variation in who is considered for surgical treatment of primary hyperparathyroidism (PHPT). Indications for surgery for symptomatic disease include the presence of end organ damage such as renal stones or reduced bone mineral density. There is much debate over whether surgery should be considered for people who are asymptomatic. In the UK, most practice adheres to the National Institute for Health consensus guidelines. They recommend surgery for the following indications:

- Serum calcium (>upper limit of normal): 1.0 mg/dL (0.25 mmol/L);
- BMD by DXA: T-score ≤2.5 at lumbar spine, total hip, femoral neck, or distal 1/3 radius;
- Vertebral fracture by x-ray, CT, MRI, or VFA;
- Creatinine clearance <60 cc/min; 24-h urine for calcium >400 mg/d (>10 mmol/d) and increased stone risk by biochemical stone risk analysis;
- Presence of nephrolithiasis or nephrocalcinosis by x-ray, ultrasound, or CT;
- <50 years

It is relevant to consider the evidence base underpinning these consensus-based US recommendations.

## 1.3 PICO table

For full details see the review protocol in appendix A.

**Table 1: PICO characteristics of review question**

| | |
|---|---|
| **Population** | Adults (18 years or over) with confirmed primary hyperparathyroidism<br><br>Strata:<br>• People with normocalcaemic PHPT<br>• Previous unsuccessful parathyroidectomy (reoperation)<br>• Pregnant women |
| **Intervention** | Parathyroid surgery |
| **Comparisons** | • No surgery (surveillance/conservative management)<br>• Calcimimetic treatment<br>• Bisphosphonate treatment<br>• Combination pharmacological treatment (calcimimetics and bisphosphonates) |
| **Outcomes** | Health related quality of life (HRQOL); mortality; preservation of end organ function [deterioration in renal function; fractures (vertebral or long bone); occurrence of kidney stones; BMD of the distal radius or the lumbar spine]; persistent hypercalcaemia (dichotomous outcome); cardiovascular events; adverse events; cancer. |

| | |
|---|---|
| **Study design** | RCT and systematic review of RCTs<br>NRS to be included in the absence of RCT evidence for the critical outcomes. NRS must be adjusted for the key confounders. |

The aim of this review was to investigate the effectiveness of surgery (parathyroidectomy) in people with different 'severities' of PHPT. As there is no one tool to define severity of disease in PHPT, subgroup populations were included in the review protocol in order to investigate the subpopulations in which surgery is effective and should be recommended. The committee defined the subgroup populations using the same criteria as set out in the 4th International Guidelines for the Management of Asymptomatic PHPT, in order to determine in whom (the presence of which individual indications) surgery is effective and should be recommended. Therefore, evidence from this review informed review questions 1.1.1 and 1.1.2.

The committee did not define people with symptomatic and asymptomatic PHPT as separate strata or subgroups in the protocol, due to the difficulty in defining who is truly asymptomatic. Also, an absence of symptoms may not necessarily indicate milder disease, as end-organ effects can be present without symptoms. For these reasons, the committee wanted to move away from classifying people as symptomatic and asymptomatic.

As non-surgical options are available in people who do not have surgery, the comparators listed in the protocols also included non-surgical pharmacological options, in addition to conservative management (monitoring only).

# 1.4 Clinical evidence

## 1.4.1 Included studies

Eleven papers (reporting eight primary studies) were included in the review;[7, 13, 27, 34, 44, 50, 51, 64, 83, 87, 88, 90] these are summarised in Table 2 and Table 3 below. Evidence from these studies is summarised in the clinical evidence summary tables below (Table 4 and **Table 5**). See also the study selection flow chart in appendix C, study evidence tables in appendix D, forest plots in appendix E and GRADE tables in appendix F.

### 1.4.1.1 Included RCTs

Seven papers (reporting five studies) were RCTs included in the review. All studies compared surgery with conservative management.

For the comparison of surgery versus conservative management, all the available studies described the population as asymptomatic. As stated above, the committee defined subgroups in order to determine in whom (the presence of which indications) surgery is effective, with the aim of investigating the effectiveness of surgery in people with asymptomatic and biochemically mild PHPT. There were an insufficient number of studies to perform subgroup analysis for any of the protocol outcomes (to determine the effectiveness of surgery in people with or without the individual indications). However, the majority of the evidence was in people who overall do not meet the current criteria for surgery with the exception of one study[34] in which the protocol subgroup criteria were unclear except to say people were free of symptoms, and another study[7] which included a small number of people with osteoporosis (as it was based on the criteria for surgery prior to 2002); had the criteria of the 2002 Workshop on Asymptomatic PHPT been adopted, 29 of the 50 participants would have met these criteria for surgery. No studies were available in people with symptomatic disease or in people with asymptomatic disease who would be eligible for surgery under the current international consensus guidelines.

No RCT evidence was identified on the clinical effectiveness of surgery in any of the population strata listed in the protocol (people with normocalcaemic PHPT, people with previous unsuccessful parathyroidectomy or pregnant women).

For the comparison of surgery versus conservative management, the critical outcome of mortality was reported by one RCT, and the critical outcome of quality of life was reported in 4 of the 6 studies for this comparison. However, data from 3 of the studies reporting QOL could not be analysed in the meta-analysis as it was only reported as graphs or narrative statements in the studies. The final study did report QOL in a format that could be analysed, but each domain of the SF-36 was reported separately and the overall physical and mental components were not reported. This study also reported the SF-36 scores as estimated annual changes from the gradient of the slope, and did not report baseline to end of study change scores, or end of study final values. As there was insufficient evidence from RCTs for the critical outcome of quality of life for the comparison of surgery versus conservative management, NRSs meeting the study protocol were included. The outcome cardiovascular events was reported by one RCT for the comparison surgery versus conservative management, however a definition for this outcome was not provided in the study.

No RCT evidence was identified for the comparators of bisphosphonates, calcimimetics or combination treatment (calcimimetics and bisphosphonates).Therefore, NRSs meeting the study protocol were investigated to see if they reported outcomes for these comparisons.

#### 1.4.1.2 Included NRS

Four papers (reporting 3 studies) were NRSs included in the review. All of these studies compared surgery with conservative management. No NRSs were identified comparing surgery with bisphosphonates or any of the other comparators listed in the protocol. Only NRS that adjusted for confounding factors were included in the review, however none of the included studies adjusted for all the key confounders listed in our protocol.

For the comparison of surgery versus conservative management, the outcomes reported were fracture, mortality, kidney stones and cancer. No evidence was available for the critical outcome of QOL. Evidence for all of the reported outcomes was already available from RCT evidence, however the population represented by the NRSs is likely to be different to that represented by the RCTs. For the NRSs, details of the severity of PHPT or details to inform our protocol subgroups were not reported, but it is likely that these studies included a mixed population of people who would and would not be eligible for surgery according to the current guidelines (in contrast to the RCT evidence which was in people not currently eligible for surgery).

No evidence was identified for the outcome of persistent hypercalcaemia from either RCTs or NRSs.

### 1.4.2 Excluded studies

See the excluded studies list in appendix I.

### 1.4.3 Summary of clinical studies included in the evidence review

See appendix D for full evidence tables.

**Table 2: Summary of RCTs included in the evidence review**

| Study | Intervention and comparison | Population | Outcomes | Comments |
|---|---|---|---|---|
| Ambrogini 2007[7] | Parathyroidectomy<br>vs<br>Conservative management<br><br>Follow-up:<br>12 months | n=50<br><br>Patients with mild PHPT who did not meet any of the NIH criteria for surgery (based on guidelines prior to 2002[(a)] so does not exclude people with osteoporosis based on the T score but does exclude people with low BMD Z score <-2).<br><br>Protocol subgroups:<br>1. Adjusted serum calcium: <2.85mmol/L<br>2. Age: ≥50 years old<br>3. Creatinine clearance: ≥ 60 mL/min (study reports as not less than 30% age-matched value).<br>4. End-organ effects: mixed (people with kidney stones and fractures excluded, some people had osteoporosis but subgroups analysis done within study) | • QOL: SF-36 and SCL-90R (unable to analyse in meta-analysis)<br>• Fractures (clinical vertebral fragility fracture)<br>• Kidney stones<br>• Lumbar spine BMD (% change from baseline)<br>• Distal radius BMD (% change from baseline)<br>• Adverse events (study outcome surgical complications, such as laryngeal nerve dysfunction)<br>• Cancer | The QOL outcomes were not reported in a format able to put into meta-analysis – only reported as graphs or narrative statements about whether there were any significant differences between the two groups |
| Elvius 1995[34] | Parathyroidectomy<br>vs<br>Conservative management<br><br>Follow-up:<br>17 years | n=48<br><br>Females with hyperparathyroidism (no detail given on diagnosis, except for females with raised serum calcium concentrations who were free of symptoms of the disease).<br><br>Protocol subgroups: | • Distal radius BMD (study outcome: bone mineral content [g/cm$^2$])<br>• Kidney function | |

| | | | | |
|---|---|---|---|---|
| | | 1. Adjusted serum calcium: not stated<br>2. Age: not stated<br>3. Creatinine clearance: not stated<br>4. End-organ effects: not stated | | |
| Rao 2004[64] | Parathyroidectomy vs Conservative management<br><br>Follow-up:<br>24 months | n=53<br><br>Patients with mild asymptomatic PHPT<br><br>Protocol subgroups:<br>1. Adjusted serum calcium: <2.85mmol/L<br>2. Age: ≥50 years old<br>3. Creatinine clearance: ≥ 60 mL/min (study states serum creatinine <1.5mg/dL (<133umol/L)<br>4. End-organ effects: absent (excluded people with non-traumatic vertebral or hip fractures and nephrolithiasis. Forearm bone mineral density within 2 S.D. adjusted for age, sex and race [Z-scores]) | • QOL: SF-36 (unable to analyse in meta-analysis)<br>• Renal dysfunction<br>• Fractures (skeletal fractures: X-ray performed to assess vertebral fractures)<br>• Kidney stones<br>• Lumbar spine BMD (unable to analyse in meta-analysis)<br>• Distal radius BMD (unable to analyse in meta-analysis)<br>• Adverse events | The QOL outcomes were not reported in a format able to put into meta-analysis – only reported as graphs or narrative statements<br><br>The BMD outcomes were given as means in each group but without any measure of variance, therefore unable to analyse in meta-analysis. |
| Scandinavian Investigation on Primary Hyperparathyroidism (SIPH) trial: Bollerslev 2007[13] (Lundstam 2015[50, 51]) | Parathyroidectomy vs Conservative management[(b)]<br><br>Follow-up:<br>1, 2 and 5 years | n=191<br><br>Adults with mild asymptomatic PHPT.<br><br>Protocol subgroups:<br>1. Adjusted serum calcium: <2.85mmol/L<br>2. Age: ≥50 years old<br>3. Creatinine clearance: unclear (excluded impaired kidney function [creatinine level > 130umol/l]).<br>4. End-organ effects: absent (excluded people with kidney stones and hyperparathyroid bone disease) | • QOL: SF-36 (unable to analyse in meta-analysis; 1 & 2 years)<br>• Mortality (5 years)<br>• Fractures (vertebral fractures on radiograph; 5 years)<br>• Fractures (minor traumatic skeletal fractures; 5 years)<br>• Kidney stones (5 years)<br>• Lumbar spine BMD (Z score; 5 years)<br>• Radius 33% (BMD, g/cm$^2$ at 5 | The QOL outcomes were not reported in a format able to put into meta-analysis – only reported as graphs or narrative statements |

| | | | | |
|---|---|---|---|---|
| | | | years)<br>• Ultra-distal radius (BMD, g/cm$^2$ at 5 years)<br>• CV events (5 years)<br>• Cancer (study outcome: development of malignancies; 5 years) | |
| Talpos 2000[83] | Parathyroidectomy<br>vs<br>Conservative management<br><br>Follow-up:<br>2 years | n=53<br><br>Women at least 5 years after menopause with persistent albumin-adjusted serum calcium level 10.1 - 11.5 mg/dL (2.52 - 2.87mmol/L) from at least 3 measurements over a period of at least 3 months; intact parathyroid hormone level > 20pg/mL; no other cause for hypercalcaemia.<br><br>Protocol subgroups:<br>1. Adjusted serum calcium: <2.85mmol/L<br>2. Age: ≥50 years old<br>3. Creatinine clearance: ≥ 60 mL/min (study reports an exclusion criteria of having a creatinine clearance level < 70%).<br>4. End-organ effects: absent (excluded people with a forearm BMD >2 SD below the expected value, vertebral compression fractures, urolithiasis on kidneys, history of non-traumatic vertebral/hip fractures; nephrolithiasis in the past 2 years) | • QOL: SF-36 (all domains reported separately) | |

*(a) The study began before the 2002 Workshop on Asymptomatic PHPT, therefore, the older guidelines formed the basis for the inclusion criteria. Had the criteria of the 2002 Workshop on Asymptomatic PHPT been adopted, 29 of the 50 participants would have met these criteria for surgery*

*(b) In the medical observation group, 9 patients received oestrogens and 3 bisphosphonates*

1 **Table 3: Summary of NRSs included in the evidence review**

| Study | Intervention and comparison | Population | Outcomes | Comments |
|---|---|---|---|---|
| Clifton-Bligh 2015[27] | Parathyroidectomy<br>vs<br>Conservative management<br><br>Follow-up:<br>Not reported | n=561<br><br>Diagnosed with PHPT either because surgery restored eucalcaemia, full investigation failed to find another cause of hypercalcaemia or serum calcium and PTH were above the upper limits of the reference range<br><br>No details of severity of PHPT<br><br>Protocol subgroups:<br>1. Adjusted serum calcium: not stated<br>2. Age: not stated<br>3. Creatinine clearance: not stated<br>4. End-organ effects: not stated | • Mortality | Adjusted for age, sex and time of diagnosis.<br>Confounders in our protocol not adjusted for: serum calcium and end-organ effects.<br>Retrospective cohort study |
| Vanderwalde 2006[87] (Vanderwalde 2009[88])<br><br>(Results from second paper used: same study but second paper adjusted for BMD | Parathyroidectomy<br>vs<br>Conservative management<br><br>Follow-up:<br>7.4 years (range: 13 days to 10 years) | n=533 (n=1569 in original study but BMD data not available for all people for adjusted analysis)<br><br>People on the database defined as having PHPT if they had an intact parathyroid hormone (PTH) level greater than 65 pg/mL, a calcium level greater than 10.5 mg/dL (>2.6 mmol/L), and a creatinine level less than 2.5 mg/dL (<221.0 µmol/L).<br><br>No details of severity of PHPT<br><br>Protocol subgroups:<br>1. Adjusted serum calcium: not stated<br>2. Age: ≥50 years old (89% ≥ 50 years old) | • Fractures (hospitalised fractures) | Adjusted for age, sex, Charlson comorbidity index (CCI); levels of calcium, PTH, and creatinine; BMD (T score femur)<br>Confounders in our protocol not adjusted for: end-organ effects.<br><br>Retrospective cohort study<br><br>Outcome of fracture taken from records of hospitalised fractures (so would not pick up all vertebral fractures on radiograph or outpatient fractures of the extremities). |

| | | | | |
|---|---|---|---|---|
| | | 3. Creatinine clearance: not stated<br>4. End-organ effects: not stated (22% had osteoporosis at baseline; kidney stones or history of fragility fractures not reported) | | |
| Vestergaard 2003[90] | Parathyroidectomy<br>vs<br>Conservative management<br><br>Follow-up:<br>6.1 years | n=3213<br><br>First time diagnosis from national hospital discharge database<br><br>No details of severity of PHPT<br><br>Protocol subgroups:<br>1. Adjusted serum calcium: not stated<br>2. Age: not stated<br>3. Creatinine clearance: not stated<br>4. End-organ effects: not stated | • Mortality<br>• Fracture<br>• Kidney stones<br>• Cancer | Adjusted for age, sex and presence of the endpoint in question at baseline.<br>Confounders in our protocol not adjusted for: serum calcium and end-organ effects.<br><br>Retrospective cohort study<br><br>Outcomes are based on whether the person had a hospital contact for that outcome in the records. |

## 1.4.4 Clinical evidence

**Table 4: Clinical evidence summary: Surgery versus conservative management**

| Outcomes | No of Participants (studies) Follow up | Quality of the evidence (GRADE) | Relative effect (95% CI) | Anticipated absolute effects | |
|---|---|---|---|---|---|
| | | | | Risk with No surgery (in mild PHPT) | Risk difference with Surgery (95% CI) |
| QOL (SF-36 Physical functioning subscale) annual change estimate. Scale from: 0 to 100. | 53<br>(1 study)<br>2 years | VERY LOW[a,b]<br>due to risk of bias, imprecision | - | The mean QOL (SF-36 physical functioning subscale) in the control groups was<br>-0.552 annual change estimate | The mean QOL (SF-36 physical functioning subscale) in the intervention groups was<br>2.1 lower (5.43 lower to 1.23 higher) |
| QOL (SF-36 Social functioning subscale) | 53<br>(1 study) | VERY LOW[a,b]<br>due to risk of bias, | - | The mean QOL (SF-36 social functioning subscale) in the | The mean QOL (SF-36 social functioning subscale) in the |

| Outcomes | No of Participants (studies) Follow up | Quality of the evidence (GRADE) | Relative effect (95% CI) | Anticipated absolute effects | |
|---|---|---|---|---|---|
| | | | | Risk with No surgery (in mild PHPT) | Risk difference with Surgery (95% CI) |
| annual change estimate. Scale from: 0 to 100. | 2 years | imprecision | | control groups was -3.653 annual change estimate | intervention groups was 3.92 higher (1.19 to 6.64 higher) |
| QOL (SF-36 Physical role functioning subscale) annual change estimate. Scale from: 0 to 100. | 53 (1 study) 2 years | VERY LOW[a,b] due to risk of bias, imprecision | - | The mean QOL (SF-36 physical role functioning subscale) in the control groups was -4.47 annual change estimate | The mean QOL (SF-36 physical role functioning subscale) in the intervention groups was 0.39 higher (5.82 lower to 6.61 higher) |
| QOL (SF-36 Emotional role functioning subscale) annual change estimate. Scale from: 0 to 100. | 53 (1 study) 2 years | VERY LOW[a,b] due to risk of bias, imprecision | - | The mean QOL (SF-36 emotional role functioning subscale) in the control groups was -5.536 annual change estimate | The mean QOL (SF-36 emotional role functioning subscale) in the intervention groups was 5.96 higher (1.47 to 10.44 higher) |
| QOL (SF-36 mental health subscale) annual change estimate. Scale from: 0 to 100. | 50 (1 study) 2 years | LOW[a] due to risk of bias | - | The mean QOL (SF-36 mental health subscale) in the control groups was 0.17 annual change estimate | The mean QOL (SF-36 mental health subscale) in the intervention groups was 0.23 higher (1.58 lower to 2.03 higher) |
| QOL (SF-36 vitality subscale) annual change estimate. Scale from: 0 to 100. | 53 (1 study) 2 years | VERY LOW[a,b] due to risk of bias, imprecision | - | The mean QOL (SF-36 vitality subscale) in the control groups was -1.77 annual change estimate | The mean QOL (SF-36 vitality subscale) in the intervention groups was 0.97 higher (1.19 lower to 3.13 higher) |
| QOL (SF-36 Bodily pain subscale) annual change estimate. Scale from: 0 to 100. | 53 (1 study) 2 years | VERY LOW[a,b] due to risk of bias, imprecision | - | The mean QOL (SF-36 bodily pain subscale) in the control groups was -1.977 annual change estimate | The mean QOL (SF-36 bodily pain subscale) in the intervention groups was 0.65 higher (2.55 lower to 3.84 higher) |
| QOL (SF-36 General health subscale) annual | 53 (1 study) | VERY LOW[a,b] due to risk of bias, | - | The mean QOL (SF-36 general health subscale) in the control | The mean QOL (SF-36 general health subscale) in the intervention groups |

| Outcomes | No of Participants (studies) Follow up | Quality of the evidence (GRADE) | Relative effect (95% CI) | Anticipated absolute effects | |
|---|---|---|---|---|---|
| | | | | Risk with No surgery (in mild PHPT) | Risk difference with Surgery (95% CI) |
| change estimate. Scale from: 0 to 100. | 2 years | imprecision | | groups was -2.961 annual change estimate | was 1.81 higher (0.38 lower to 4.01 higher) |
| QOL (SF-36 Health transition) annual change estimate. Scale from: 0 to 100. | 53 (1 study) 2 years | VERY LOW[a,b, c] due to risk of bias, imprecision | - | The mean QOL (SF-36 health transition) in the control groups was -1.154 | The mean QOL (SF-36 health transition) in the intervention groups was 0.12 higher (3.1 lower to 3.33 higher) |
| Mortality | 191 (1 study) 5 years | VERY LOW[a,b] due to risk of bias, imprecision | RR 1.98 (0.18 to 21.46) | Moderate | |
| | | | | 11 per 1000 | 11 more per 1000 (from 9 fewer to 225 more) |
| Renal Dysfunction | 73 (2 studies) 2-17 years | LOW [a, e] due to risk of bias, imprecision | Not estimable | Moderate | |
| | | | | 0 per 1000 | 0 more per 1000 (from 180 fewer to 180 more)[d] |
| Vertebral fractures | 208 (3 studies) 1-5 years | LOW[a] due to risk of bias | OR 0.14 (0.03 to 0.69) | Moderate | |
| | | | | 40 per 1000 | 60 fewer per 1000 (from 110 fewer to 0 more)[d] |
| Peripheral skeletal fractures | 106 (1 study) 5 years | VERY LOW[a,b] due to risk of bias, imprecision | RR 0.81 (0.19 to 3.44) | Moderate | |
| | | | | 73 per 1000 | 14 fewer per 1000 (from 59 fewer to 178 more) |
| Kidney Stones | 208 (3 studies) 1-5 years | VERY LOW[a,b] due to risk of bias, imprecision | Peto OR 0.39 (0.06 to 2.82) | Moderate | |
| | | | | 36 per 1000 | 20 fewer per 1000 (from 60 fewer to 30 more) |

| Outcomes | No of Participants (studies) Follow up | Quality of the evidence (GRADE) | Relative effect (95% CI) | Anticipated absolute effects | |
|---|---|---|---|---|---|
| | | | | Risk with No surgery (in mild PHPT) | Risk difference with Surgery (95% CI) |
| Lumbar spine BMD Z score (final value) | 111 (1 study) 5 years | VERY LOW[a,b] due to risk of bias, imprecision | - | The mean lumbar spine BMD Z score in the control groups was -0.09 | The mean lumbar spine BMD in the intervention groups was 0.48 higher (0.03 lower to 0.99 - higher) |
| Lumbar spine BMD % change from baseline | 49 (1 study) 1 years | VERY LOW[a,b] due to risk of bias, imprecision | - | The mean lumbar spine BMD in the control groups was -1.12% change from baseline | The mean lumbar spine BMD in the intervention groups was 5.28 higher (4.76 to 5.8 higher) |
| Distal radius BMD g/cm$^2$ | 20 (1 study) 17 years | VERY LOW[a,b] due to risk of bias | - | The mean distal radius BMD in the control groups was 1.03 g/cm$^2$ | The mean distal radius BMD in the intervention groups was 0.05 lower (0.22 lower to 0.12 higher) |
| Distal radius BMD % change from baseline | 49 (1 study) 1 years | LOW[a,b] due to risk of bias, imprecision | - | The mean distal radius BMD in the control groups was -0.55% change from baseline | The mean distal radius BMD in the intervention group was 0.21 higher (0.1 lower to 0.52 higher) |
| Radius 33% (BMD, g/cm$^2$) (5 years) | 86 (1 study) 5 years | VERY LOW[a,b] due to risk of bias, imprecision | - | The mean radius 33% BMD in the control groups was 0.584 g/cm$^2$ | The mean radius 33% (BMD, g/cm$^2$) in the intervention groups was 0.03 higher (0.02 lower to 0.08 higher) |
| Ultra-distal radius (BMD, g/cm$^2$) (5 years) | 85 (1 study) 5 years | LOW[a] due to risk of bias | - | The mean ultra-distal radius BMD in the control groups was 0.297 g/cm$^2$ | The mean ultra-distal radius (BMD, g/cm$^2$) in the intervention groups was 0.01 higher (0.03 lower to 0.04 higher) |
| Cardiovascular events | 145 (1 study) 5 years | VERY LOW[a,b] due to risk of bias, imprecision | RR 0.63 (0.22 to 1.85) | Moderate | |
| | | | | 110 per 1000 | 41 fewer per 1000 (from 86 fewer to 94 more) |

| Outcomes | No of Participants (studies) Follow up | Quality of the evidence (GRADE) | Relative effect (95% CI) | Anticipated absolute effects | |
|---|---|---|---|---|---|
| | | | | Risk with No surgery (in mild PHPT) | Risk difference with Surgery (95% CI) |
| Adverse events | 102 (2 studies) 1-2 years | VERY LOW[a,b] due to risk of bias, imprecision | RR 0.75 (0.14 to 4.11) | Moderate 54 per 1000 | 14 fewer per 1000 (from 46 fewer to 168 more) |
| Cancer | 194 (2 studies) 1-5 years | VERY LOW[a,b,] due to risk of bias, imprecision | Peto OR 1.53 (0.26 to 8.97) | Moderate 27 per 1000 | 10 more per 1000 (from 40 fewer to 60 more) |

*a Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias.*
*b Downgraded by 1 increment if the confidence interval crossed 1 MID, and downgraded by 2 increments if the confidence interval crossed both MIDs.*
*c Established MID not available for this domain of the SF-36, therefore default MID used*
*d Manual calculation of absolute risk difference*

*e Downgraded by 1 increments as both studies had 0 events in both arms and sample size was >70<350*

**Table 5: Clinical evidence summary: Surgery versus conservative treatment (non-randomised studies)**

| Outcomes | No of Participants (studies) Follow up | Quality of the evidence (GRADE) | Relative effect (95% CI) | Anticipated absolute effects | |
|---|---|---|---|---|---|
| | | | | Risk with Conservative treatment (NRS) | Risk difference with Surgery (95% CI) |
| Mortality | 3774 (2 studies) 6.1 years | VERY LOW[a] due to risk of bias | HR 0.65 (0.57 to 0.74) | See comment[c] | See comment[c] |

| Outcomes | No of Participants (studies) Follow up | Quality of the evidence (GRADE) | Relative effect (95% CI) | Anticipated absolute effects | |
|---|---|---|---|---|---|
| | | | | Risk with Conservative treatment (NRS) | Risk difference with Surgery (95% CI) |
| Fractures | 3746 (2 studies) 6.1-7.4 years | VERY LOW[a,b] due to risk of bias, imprecision | HR 0.67 (0.55 to 0.82) | See comment[c] | See comment[c] |
| Cancer | 3213 (1 study) 6.1 years | VERY LOW[a,b] due to risk of bias, imprecision | HR 1.11 (0.9 to 1.37) | 65 per 1000 | 10 more per 1000 (from 9 fewer to 32 more) |
| Kidney stones | 3213 (1 study) 6.1 years | VERY LOW[a] due to risk of bias | HR 1.87 (1.3 to 2.69) | 65 per 1000 | 53 more per 1000 (from 19 more to 100 more) |

[a] Downgraded by 1 increment if the majority of studies were at high risk of bias, and downgraded by 2 increments if the majority of studies were at very high risk of bias.
[b] Downgraded by 1 increment if the confidence interval crossed 1 MID, and downgraded by 2 increments if the confidence interval crossed both MIDs.
[c] Control group risk not reported

See appendix F for full GRADE tables.

Narrative results

A modest but significant beneficial effect on quality of life [bodily pain (p=0.001); general health (p=0.008); vitality (p=0.003); and mental health (p=0.017)] was observed in patients after surgery compared with those followed without surgery. No difference was found in the remaining SF-36 and SCL-90R domains (Ambrogini). In comparison with the patients who did not have surgery, a statistically significant beneficial effect of parathyroidectomy was seen in two of the nine domains (social function, group difference p=0.007; and emotional role function, group difference, p=0.012 (Sudhaker). Concerning the physical domains, a slightly, but significant, decrease was observed over the two-year period in the medical observation group (p<0.01), whereas no change was seen in the operation group. The difference over time was significantly different in favour of surgery (p<0.01). The operation group scored slightly higher at year one, compared with baseline in the mental health subdomain and mental component summary score (p<0.05 for both), but not after two years of observation. For the mental health subdomain, the observation group scored higher at two years, compared with baseline (p<0.05). Although no longitudinal differences were observed in any group in the other psychological domains, the differences over time for the domain role emotional were in favour of surgery for both one and two years of observation[13].

# 1.5 Economic evidence

## 1.5.1 Included studies

No relevant health economic studies were identified.

## 1.5.2 Excluded studies

One health economic study was identified relevant to this question, but was excluded due to a combination of limited applicability and methodological limitations.[74] This is listed in appendix I, with reasons for exclusion given.

See also the health economic study selection flow chart in appendix G.

### 1.5.3 Unit costs

Below are unit costs of surgery for primary hyperparathyroidism, from NHS reference costs.

**Table 6: Parathyroid procedures costs (Elective inpatient schedule)**

| HRG code | Description | Activity | National average unit cost | Average cost of excess bed day | Average Length of Stay - Days | No. Data Submissions |
|---|---|---|---|---|---|---|
| KA03C | Parathyroid Procedures with CC Score 2+ | 1,444 | £3,227 | £432 | 1.47 | 189 |
| KA03D | Parathyroid Procedures with CC Score 0-1 | 1,883 | £2,851 | £578 | 1.00 | 186 |
| | Weighted average (including complications and excess bed days) | | | | | |
| KA03C and KA03D | Parathyroid procedures | 3,327 | £3,154 | | 1.2 | |

*Source: NHS reference costs 2016-17*[30]

## 1.6 Resource costs

The recommendations made by the committee based on this review may have a substantial impact on resources.

Additional costs could be incurred where the recommendations lead to a change in practice for NHS providers. At present, people who are mostly asymptomatic are not routinely recommended for surgical intervention. If the recommendation lead to a large increase in the number of surgeries performed for PHPT, there will potentially be a large increase in healthcare resource use. However, it is unclear how widely this will be implemented.

## 1.7 Evidence statements

### 1.7.1 Clinical evidence statements

#### 1.7.1.1 Surgery versus conservative management (randomised studies)

There was a clinically important benefit of surgery for QOL (SF-36 Social functioning subscale; SF-36 Emotional role functioning subscale) (1 study, n=53; follow-up 2 years; Very Low quality) vertebral fractures (3 studies, n=208; follow-up 1-5 years; Low quality); lumbar spine BMD % change from baseline (1 study, n=49; follow up 17 years; Very Low quality); distal radius BMD % change from baseline (1 study, n=49; follow-up 1 year; Low quality and cardiovascular events (1 study, n=145; follow-up 5 years; Very Low quality).

There was no difference between surgery and conservative management for QOL (SF-36 physical functioning subscale; SF-36 physical role functioning subscale; SF-36 mental health subscale; SF-36 vitality subscale; SF-36 bodily pain subscale; SF-36 general health subscale; SF-36 health transition) (1 study, n=53; follow-up 2 years; Very Low quality); mortality (1 study, n=191; follow-up 5 years; Very Low quality); renal dysfunction (2 studies, n=73; follow-up 2-17 years; Low quality); peripheral skeletal fractures (1 study, n=106; follow-up 5 years; Very Low quality); kidney stones (3 studies, n=208; follow-up 1-5 years; Very

Low quality); lumbar spine BMD Z score final value (1 study, n=111; follow-up 5 years; Very Low quality); distal radius BMD (1 study, n=20; follow-up 17 years; Very Low quality); ultra-distal radius BMD (1 study, n=85; follow-up 5 years; Low quality); radius 33% BMD (1 study, n=86; follow-up 5 years; Very Low quality); adverse events (2 studies, n=102; follow-up 1-2 years; Very Low quality); and cancer (2 studies, n=194; follow-up 1-5 years; Very Low quality). No evidence was identified for the outcome of persistent hypercalcaemia.

#### 1.7.1.2 Surgery versus conservative management (non-randomised studies)

There was clinically important benefit of surgery for mortality (2 studies, n=3774; follow-up 6.1 years; Very Low quality) and fractures (2 studies, n=3746; follow-up 6.1-7.4 years; Very Low quality). There was clinical harm of surgery for the outcome kidney stones (1 study, n=3213; follow-up 6.1 years; Very Low quality). There was no difference between surgery and conservative management for cancer (1 study, n=3213; follow-up 6.1 years; Very Low quality). No evidence was identified for the outcomes persistent hypercalcaemia and health related QOL.

#### 1.7.1.3 Surgery versus bisphosphonates

No evidence was identified.

#### 1.7.1.4 Surgery versus calcimimetics

No evidence was identified.

#### 1.7.1.5 Surgery versus combination treatment (calcimimetics and bisphosphonates)

No evidence was identified.

### 1.7.2 Health economic evidence statements

No relevant economic evaluations were identified.

## 1.8 Recommendations

### *Referral for surgery*

#### Indications for referral for surgery

C1. Refer people with primary hyperparathyroidism to a surgeon with expertise in parathyroid surgery if they have:

- symptoms of hypercalcaemia such as thirst, frequent or excessive urination, or constipation **or**
- end-organ disease (renal stones, fragility fractures or osteoporosis) **or**
- an albumin-adjusted serum calcium level of 2.85 mmol/litre or above.

C2. Consider referral to a surgeon with expertise in parathyroid surgery for people with primary hyperparathyroidism irrespective of the features listed in recommendation C1.

# 1.9 The committee's discussion of the evidence

## 1.9.1 Interpreting the evidence

### 1.9.1.1 The outcomes that matter most

The committee considered the outcomes of health-related quality of life, mortality and preservation of end organ function (bone mineral density, fractures, renal stones and renal function) as critical outcomes for decision making. Other important outcomes included adverse events, cancer incidence, cardiovascular events and persistent hypercalcaemia. The committee was interested in cardiovascular and cancer outcomes, as there is some observational prognostic evidence to suggest that the risk of these future events is higher in untreated primary hyperparathyroidism.

From the non-randomised studies (NRSs) no evidence was available for the critical outcome of quality of life. No evidence was identified for the outcome of persistent hypercalcaemia from either the randomised controlled trials (RCTs) or NRSs.

### 1.9.1.2 The quality of the evidence

All the evidence in this review (both RCTs and NRSs) compared surgery with conservative management. No evidence was available for the comparison of surgery with bisphosphonates, calcimimetics or combination treatment from either RCTs or NRSs.

The majority of the studies did not provide any details on conservative management; out of the 8 studies, 6 studies did not provide any details; one study stated 'non-operative conservative management' but did not provide any further details; another study reported 'no surgery' and follow-up every 6 months for at least 24 months with no further details.

All the available RCTs described the population as asymptomatic. The majority of the RCT evidence was in people who overall do not meet the current *National Institutes of Health* (NIH) criteria for surgery (with the exception of one study[34] in which the protocol subgroup criteria were unclear except to say people were free of symptoms). There was another study which included a small number of people with osteoporosis as it was based on the criteria for surgery prior to 2002 – had the criteria of the 2002 Workshop on Asymptomatic primary hyperparathyroidism been adopted, 29 of the 50 participants would have met these criteria for surgery. No studies were available in people with symptomatic disease or in people with asymptomatic disease who would be eligible for surgery under the NIH guidelines. The current NIH criteria [11] for surgery in people with asymptomatic primary hyperparathyroidism are as follows: Serum calcium (>upper limit of normal): 1.0 mg/dL (0.25 mmol/L); BMD by DXA: T-score ≤2.5 at lumbar spine, total hip, femoral neck, or distal 1/3 radius; vertebral fracture by X-ray, CT, MRI, or VFA; creatinine clearance < 60 cc/min; 24-hour urine for calcium >400 mg/d (>10 mmol/d) and increased stone risk by biochemical stone risk analysis; presence of nephrolithiasis or nephrocalcinosis by X-ray, ultrasound, or CT; <50 years old.

For the RCTs comparing surgery with conservative management, the majority of the evidence was of Low to Very Low quality due to risk of bias and imprecision. This decreases our confidence in the estimate of effect of surgery.

For NRSs, details of the severity of primary hyperparathyroidism or to inform our protocol subgroups were not reported, but it is likely that these studies included a mixed population of people who would and would not be eligible for surgery according to the current guidelines (in contrast to the RCT evidence which was in people not currently eligible for surgery).

For the NRSs evidence all outcomes were graded as Very Low quality due to high risk of bias and imprecision.

### 1.9.1.3 Benefits and harms

As there is no one tool to define severity of disease in primary hyperparathyroidism, subgroup populations were included to investigate the populations in which surgery is effective and should be recommended. The guideline committee defined the subgroup populations using the same criteria as set out in the 4th International Guidelines for the Management of Asymptomatic Primary Hyperparathyroidism, in order to determine in whom (the presence of which individual indications) surgery is effective and should be recommended.

The subgroups were: people with end-organ effects versus absence of end-organ effects (end organ effects defined as renal stones, history of fragility fractures or osteoporosis [BMD T-score <-2.5 at any site]); serum adjusted calcium > 0.25 mmol/litre above the ULN (same as ≥2.85mmol/litre and <2.85mmol/litre); reduction in creatinine clearance to <60 mL/minute; and age under 50 years versus ≥50 years. However, there were an insufficient number of studies to perform subgroup analysis for any of the protocol outcomes.

The committee also planned to consider the following population strata: people with normocalcaemic primary hyperparathyroidism (serum adjusted calcium ≤2.6mmol/litre and an elevated PTH that cannot be explained by abnormal renal function or low 25OHD); previous unsuccessful parathyroidectomy (reoperation); and pregnant women. No evidence was identified on the clinical effectiveness of surgery in any of the population strata listed above.

The RCT evidence for the comparison surgery versus conservative management suggested that there was a clinical benefit of surgery for the outcomes quality of life (for 2 domains), vertebral fractures, lumbar spine BMD (% change from baseline); distal radius BMD % change from baseline (1 study, n=49; follow-up 1 year; Low quality) and cardiovascular events. The RCT evidence suggested that there was no difference between the groups surgery and conservative management for the outcomes mortality, quality of life (for 7 domains), renal dysfunction, peripheral skeletal fractures, renal stones, lumbar spine BMD Z score (final value), distal radius (BMD g/cm$^2$), ultra-distal radius (BMD, g/cm$^2$), radius 33% (BMD, g/cm$^2$), adverse events and cancer. The estimates were imprecise for all the above outcomes except for distal radius BMD g/cm$^{2,}$ ultra-distal radius (BMD, g/cm$^2$) and vertebral fractures.

The NRS evidence for the comparison surgery versus conservative management suggested that there was clinical benefit of surgery for the outcomes mortality and fractures. Although there was a clinical benefit for fractures it was noted that the estimate was imprecise. Evidence suggested that there was clinical harm of surgery for the outcome renal stones. Evidence suggested that there was no difference between the groups for the outcome cancer however the estimate was imprecise.

For the non-randomised studies, the committee noted the apparent raised risk of renal stones in people who had surgery but from their experience felt that this was likely to represent their higher risk, as once someone has had a renal stone they remain at higher risk of a recurrence. The non-randomised data on fracture was consistent with the randomised evidence. It was reassuring that there was a significantly lower mortality in the surgical arm but this was largely likely to be due to confounding factors (people selected for surgery tend to be fitter).

The committee felt that some primary hyperparathyroidism patients present with long standing non-specific/undifferentiated symptoms such as fatigue, depression, muscle weakness, abdominal pain, loss of concentration etc. However the committee felt that such symptoms occur in many other diseases and agreed not to make a recommendation for such non-specific symptoms as indications for surgery. The committee noted that primary hyperparathyroidism is associated with a decline in renal function but there is no evidence that parathyroidectomy leads to an improvement. They noted that specific thresholds for

renal dysfunction (creatinine clearance, 24-hour urine calcium) have been used in other countries as indications for surgery, but there are no data available to suggest that these cut-offs in isolation would be an indication for parathyroidectomy. The committee noted that 24-hour calcium is a good predictor of renal stone formation in the future. They felt that renal function thresholds for deteriorating renal function can be considered as part of decision making.

The committee noted that there was no evidence to support a particular cut-off point for adjusted serum calcium requiring surgery but they felt that it was reasonable to define a threshold of 2.85mmol/litre or above at which surgery would be recommended.

The committee felt that the evidence in favour of surgery in patients who do not already have indications for surgery in these trials provided indirect evidence of benefit in the population in whom surgery is currently performed for whom no randomised evidence was found. This is because the currently accepted indications are in people who are at higher risk of the adverse sequellae of primary hyperparathyroidism and therefore would in principle benefit more from the operation.

The committee felt that the absence of randomised evidence in the population that meet the NIH criteria reflects the broad international consensus that surgery is indicated in this group. For people with no symptoms or indications for surgery, the committee based their recommendation on limited evidence together with their clinical experience. The recommendation is for the person to be referred for surgery so that their specific risks and benefits can be discussed. Surgery would not be offered for all of these people. A proportion of these people would meet the current criteria for surgery in the future but the committee proposed to consider surgery earlier to avoid the potential consequences of primary hyperparathyroidism. The committee felt that the benefits of surgery shown in people with no symptoms or other indications for surgery would be magnified for people with more severe disease. The committee from clinical experience noted that primary hyperparathyroidism patients have lower bone density, increased fracture risk, osteoporosis; and surgery reduces the risk of fracture in such patients. The committee from their clinical experience also discussed that kidney stones are one of the end organ effects of primary hyperparathyroidism and the risk of developing renal stones decreases after surgery.The committee felt that surgery should be considered in people who have risk factors which are predictors of end organ disease or progressive disease. Risk factors discussed included younger age with persistent hypercalcaemia but below the 2.85 mmol/litre threshold, and symptoms suggestive of renal stone disease without current stones but with elevated urinary calcium excretion.

The committee discussed that if surgery is to be offered, it is important that the risks and benefits of the procedure are fully explained so that the patient can make an informed choice.

The committee determined that whilst the current NIH criteria separates those who are below 50 and those who are over 50, it would not be appropriate to make this distinction in their recommendations to ensure equality of access to surgery regardless of age. The age of the person is a factor for the clinician to discuss with the person when considering whether surgery is a suitable option for them. The committee emphasised that the consideration is more about life expectancy than age, as performance status is not necessary correlated with age in a linear way.

The committee discussed the other management approaches compared to surgery including calcimimetics and bisphosphonates. The committee noted that cinacalcet (calcimimetics) should be an option in people who are unable to undergo surgery only and not as an alternative to surgery, as parathyroidectomy is the only definitive treatment option in people with primary hyperparathyroidism without surgical contraindication. The committee from their experience stated that cinacalcet does not directly stop bone loss or kidney problems due to primary hyperparathyroidism (for further discussion of this evidence please refer to Evidence

review G). The committee also discussed that as bisphosphonates do not provide a cure for the underlying condition of primary hyperparathyroidism, they should not be considered as an alternative to curative measures such as surgery. However the committee agreed that bisphosphonates should be considered in people with primary hyperparathyroidism and bone end organ effects, to reduce fracture risk (for further discussion of this evidence please refer to Evidence review H).

### 1.9.2 Cost effectiveness and resource use

No relevant economic evaluations were identified for this question.

Unit costs were presented to the committee for consideration. The average cost of an elective inpatient parathyroid procedure is around £3,050, with an average length of stay of 1.5 days. This was estimated using NHS reference costs (2015–16), and takes into account complexity of procedure with regard to complications and comorbidities.

This area was initially identified as being high priority for original economic analysis. However, following the clinical review it was judged that economic modelling for this question would not be possible due to the lack of clinical evidence regarding the effectiveness of parathyroidectomy for people with either symptomatic or asymptomatic disease. Consequently, cost effectiveness of parathyroidectomy could not be calculated and is therefore highly uncertain.

However, the committee discussed that surgery is the only definitive cure for primary hyperparathyroidism. They noted that surgery is likely to cure primary hyperparathyroidism (current national cure rate around 94%) and therefore cure hypercalcaemia and relieve patients of symptoms such as thirst, polyuria and constipation. Furthermore, the committee considered that surgery in this population could also prevent future events such as renal stones and fragility fractures from occurring which will incur both a high cost to the NHS as well as reducing quality of life for the person. Furthermore, surgery would be more cost effective as it requires a one-off high cost with sustained benefit due to cure, whereas for example calcimimetics requires continuous high cost to maintain a similar benefit without providing a definitive cure of the primary hyperparathyroidism.

The committee considered that those with the greatest potential for quality of life gains and cost savings, and hence those for which surgery is most likely to be cost effective, are those who have symptoms of hypercalcaemia, or end organ disease, or those with a serum calcium level of 2.85mmol/litre or above. They therefore agreed to offer surgery to this population. Therefore as mentioned in the benefits and harms section above, the population for which the committee have recommended surgery should be offered reflect broad international consensus, and as a result this recommendation is in line with current practice and therefore will not have a substantial resource impact.

The committee expressed concern that in current practice, people with primary hyperparathyroidism who may potentially be cured by surgery are not currently being referred to have surgery due to not meeting current NIH criteria. It was estimated this might affect around 15–20% of patients. Therefore, the committee also considered the cost effectiveness of surgery for those who do not meet these criteria - an ‘asymptomatic’ population. The committee discussed that as these people are generally ‘asymptomatic’ the likely quality of life gains initially after surgery are likely to be smaller, however they still considered there could be some improvement due to the possible resolution of non-specific symptoms people with ‘asymptomatic’ primary hyperparathyroidism can experience such as fatigue, depression and muscle weakness to name a few. The committee also discussed that if surgery was not considered in this population they would be monitored, which also incurs a cost. Furthermore the committee recognised that people may become eligible according to the recommendations at a later date due to disease progression. The committee discussed

that by this point their quality of life could have worsened due to the development of symptoms of hypercalcaemia or possible due to end organ damage. However, as there are no data available to suggest the rate or proportion of people that are likely to become eligible for surgery according to these criteria, as well as a lack of data available on the effectiveness of monitoring in detecting potential disease progression prior to end organ damage occurring, the cost effectiveness of surgery in this population is highly uncertain. However, the committee considered that because future decrements in quality of life and cost of events associated with end organ damage could be avoided, surgery should be considered in this group.

It is uncertain how many additional surgeries would be performed as a result of this recommendation; hence it is not possible to estimate its impact on healthcare resource use. However, if widely implemented there is potential for there to be a substantial resource impact.

### 1.9.3 Other factors the committee took into account

The committee considered symptomatic primary hyperparathyroidism to include symptoms attributable to hypercalcaemia such as thirst, polyuria and constipation. They also recognised associations with non-specific symptoms such as fatigue, depression, muscle weakness, constipation, abdominal pain, loss of concentration, mild confusion etc. End organ disease refers particularly to disease of the kidney and bones as these are more commonly associated with primary hyperparathyroidism. The committee noted primary hyperparathyroidism was considered as a rare cause of pancreatitis, but there was no evidence to suggest that parathyroid surgery would improve the course of pancreatitis in such patients.

The committee noted that surgery is only offered if the benefits outweigh the risks. People may not be offered surgery if they have a very high operative risk, airway problems, distorted anatomy or short life expectancy.

The committee discussed the terminologies used for parathyroid surgery and stated that parathyroid surgery is surgery targeted at the parathyroid and parathyroidectomy is removal of parathyroid tissue. They noted that there may be failed parathyroidectomy (or unsuccessful) that is still parathyroid surgery.

# References

1. 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

2. Agus ZS. Conservative vs surgical treatment of hyperparathyroidism: which to choose, and when? Cleveland Clinic Journal of Medicine. 1993; 60(3):191-2

3. Alhava EM, Karjalainen P, Paakkonen M. Bone mineral density and surgical treatment of primary hyperparathyroidism. Acta Chirurgica Scandinavica. 1988; 154(5-6):345-7

4. Almqvist EG, Becker C, Bondeson AG, Bondeson L, Svensson J. Early parathyroidectomy increases bone mineral density in patients with mild primary hyperparathyroidism: a prospective and randomized study. Surgery. 2004; 136(6):1281-8

5. Almqvist EG, Bondeson AG, Bondeson L, Nissborg A, Smedgård P, Svensson SE. Cardiac dysfunction in mild primary hyperparathyroidism assessed by radionuclide angiography and echocardiography before and after parathyroidectomy. Surgery. 2002; 132(6):1126-32; discussion 1132

6. Alvarez-Allende CR, Pascual Marrero AM, Castillo CA, Mendez-Latalladi W. Parathyroidectomy outcomes in normocalcemic primary hyperparathyroidism. Journal of the American College of Surgeons. 2014; 219(4 Suppl):e12

7. Ambrogini E, Cetani F, Cianferotti L, Vignali E, Banti C, Viccica G et al. Surgery or surveillance for mild asymptomatic primary hyperparathyroidism: a prospective, randomized clinical trial. Journal of Clinical Endocrinology and Metabolism. 2007; 92(8):3114-21

8. Anonymous. Erratum: A 10-year prospective study of primary hyperparathyroidision with or without parathyroid surgery (New England Journal of Medicine (October 21, 1999) 341 (1249-1255)). New England Journal of Medicine. 2000; 342(2):144

9. Anonymous. Surgery reduces risk of fracture in primary hyperparathyroidism. BMJ. 2000; 321(7261):C

10. Barkun J, Duh QY, Wiseman S, McKenzie M. Canadian Association of General Surgeons and American College of Surgeons Evidence Based Reviews in Surgery. 16. Randomized trial of parathyroidectomy in mild asymptomatic primary hyperparathyroidism. Canadian Journal of Surgery. 2006; 49(1):59-61

11. 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

12. Blanchard C, Mathonnet M, Sebag F, Caillard C, Kubis C, Drui D et al. Quality of life is modestly improved in older patients with mild primary hyperparathyroidism postoperatively: Results of a prospective multicenter study. Annals of Surgical Oncology. 2014; 21(11):3534-40

13. Bollerslev J, Jansson S, Mollerup CL, Nordenström J, Lundgren E, Tørring O et al. Medical observation, compared with parathyroidectomy, for asymptomatic primary hyperparathyroidism: a prospective, randomized trial. Journal of Clinical Endocrinology and Metabolism. 2007; 92(5):1687-92

14. Bollerslev J, Rosen T, Mollerup CL, Nordenström J, Baranowski M, Franco C et al. Effect of surgery on cardiovascular risk factors in mild primary hyperparathyroidism. Journal of Clinical Endocrinology and Metabolism. 2009; 94(7):2255-61

15. Bonzelaar LB, Salapatas AM, Hwang MS, Friedman M. Parathyroidectomy for hyperparathyroidism: Morbidity and mortality. Otolaryngology - Head and Neck Surgery. 2016; 155(1):P57

16. Britton DC, Thompson MH, Johnston ID, Fleming LB. Renal function following parathyroid surgery in primary hyperparathyroidism. Lancet. 1971; 2(7715):74-75

17. Brothers TE, Thompson NW. Surgical treatment of primary hyperparathyroidism in elderly patients. Acta Chirurgica Scandinavica. 1987; 153(3):175-8

18. Broulik PD, Broulikova A, Adamek S, Libansky P, Tvrdon J, Broulikova K et al. Improvement of hypertension after parathyroidectomy of patients suffering from primary hyperparathyroidism. International Journal of Endocrinology. 2011; 2011:309068

19. Bruining HA, Van Houten H, Juttmann JR. Results of operative treatment of 615 patients with primary hyperparathyroidism. World Journal of Surgery. 1981; 5(1):85-90

20. Burney RE, Jones KR, Coon JW, Blewitt DK, Herm AM. Assessment of patient outcomes after operation for primary hyperparathyroidism. Surgery. 1996; 120(6):1013-9

21. Burney RE, Jones KR, Peterson M, Christy B, Thompson NW, Cady B et al. Surgical correction of primary hyperparathyroidism improves quality of life. Surgery. 1998; 124(6):987-92

22. Calo 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. 2016; 29(S1):15-21

23. Carneiro-Pla DM, Irvin GL, III, Chen H. Consequences of parathyroidectomy in patients with "mild" sporadic primary hyperparathyroidism. Surgery. 2007; 142(6):795-9

24. Chen H, Parkerson S, Udelsman R. Parathyroidectomy in the elderly: Do the benefits outweigh the risks? World Journal of Surgery. 1998; 22(6):531-6

25. Cheng SP, Lee JJ, Liu TP, Yang PS, Liu SC, Hsu YC et al. Quality of life after surgery or surveillance for asymptomatic primary hyperparathyroidism: A meta-analysis of randomized controlled trials. Medicine. 2015; 94(23):e931

26. Chigot JP, Menegaux F, Achrafi H. Should primary hyperparathyroidism be treated surgically in elderly patients older than 75 years? Surgery. 1995; 117(4):397-401

27. Clifton-Bligh PB, Nery ML, Supramaniam R, Reeve TS, Delbridge L, Stiel JN et al. Mortality associated with primary hyperparathyroidism. Bone. 2015; 74:121-4

28. Cowie AGA. Morbidity in adult parathyroid surgery. Journal of the Royal Society of Medicine. 1982; 75(12):942-5

29. D'Andrea V, Biancari F, Catania A, Chiarini S, Lippolis G, Falvo L et al. Surgical treatment for hyperparathyroidism. Italian Journal of Mineral and Electrolyte Metabolism. 1996; 10(1):39-43

30. Department of Health. NHS reference costs 2015-16. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2015-to-2016 Last accessed: 17/01/2018.

31. Diaz-Guerra GM, Guadalix S, Garcia F, Melon N, Martinez-Pueyo JI, Ferrero E et al. Parathyroidectomy in primary hyperparathyroidism (PHPT): Retrospective study of 272 patients. Endocrine Reviews. 2015; 36(2 Suppl):SAT-266

32. Dy BM, Grant CS, Wermers RA, Kearns AE, Huebner M, Harmsen WS et al. Changes in bone mineral density after surgical intervention for primary hyperparathyroidism. Surgery. 2012; 152(6):1051-8

33. Edwards ME, Rotramel A, Beyer T, Gaffud MJ, Djuricin G, Loviscek K et al. Improvement in the health-related quality-of-life symptoms of hyperparathyroidism is durable on long-term follow-up. Surgery. 2006; 140(4):655-64

34. Elvius M, Lagrelius A, Nygren A, Alveryd A, Christensson TA, Nordenström J. Seventeen year follow-up study of bone mass in patients with mild asymptomatic hyperparathyroidism some of whom were operated on. European Journal of Surgery. 1995; 161(12):863-9

35. Espiritu RP, Kearns AE, Vickers KS, Grant C, Ryu E, Wermers RA. Depression in primary hyperparathyroidism: Prevalence and benefit of surgery. Journal of Clinical Endocrinology and Metabolism. 2011; 96(11):E1737-E1745

36. Falkheden T, Ohlsson L, Sjogren B. Renal function in primary hyperparathyroidism. A follow-up study two to eleven years after surgery comprising 139 patients. Scandinavian Journal of Urology and Nephrology. 1980; 14(2):167-75

37. Fang WL, Tseng LM, Chen JY, Chiou SY, Chou YH, Wu CW et al. The management of high-risk patients with primary hyperparathyroidism - Minimally invasive parathyroidectomy vs. medical treatment. Clinical Endocrinology. 2008; 68(4):520-8

38. Farnebo LO, Trigonis C, Forsgren L. Surgery for primary hyperparathyroidism. Experience with 400 patients during 10 years (1972-1981). Acta Chirurgica Scandinavica. 1984; 150(Suppl. 520):11-6

39. Freaney R, Casey OM, Muldowney FP. The long-term effect of parathyroidectomy on renal function. Irish Journal of Medical Science. 1978; 147(6):205-9

40. Ghose RR, Morgan WD. Improvement in renal function in primary hyperparathyroidism following parathyroidectomy. Postgraduate Medical Journal. 1981; 57(663):28-30

41. Hagstrom E, Lundgren E, Mallmin H, Rastad J, Hellman P. Positive effect of parathyroidectomy on bone mineral density in mild asymptomatic primary hyperparathyroidism. Journal of Internal Medicine. 2006; 259(2):191-8

42. Hedback G, Oden A, Tisell LE. The influence of surgery on the risk of death in patients with primary hyperparathyroidism. World Journal of Surgery. 1991; 15(3):399-407

43. Hedback G, Tisell LE, Bengtsson BA, Hedman I, Oden A. Premature death in patients operated on for primary hyperparathyroidism. World Journal of Surgery. 1990; 14(6):829-36

44. Horiuchi T, Onouchi T, Inoue J, Shionoiri A, Hosoi T, Orimo H. A strategy for the management of elderly women with primary hyperparathyroidism: a comparison of etidronate therapy with parathyroidectomy. Gerontology. 2002; 48(2):103-8

45. Jansson S, Mollerup C, Nordenstrom J, Varhaug JE, Isaksen G, Bollerslev J. The SIPH study: surgery vs. medical observation in mild, asymptomatic phpt preliminary results from a prospective, randomized study. Langenbeck's Archives of Surgery. 2006; 391(3):242

46. Khosla S, Melton LJ, III, Wermers RA, Crowson CS, O'Fallon WM, Riggs BL. Primary hyperparathyroidism and the risk of fracture: A population-based study. Journal of Bone and Mineral Research. 1999; 14(10):1700-7

47. Lafferty FW, Hubay CA. Primary hyperparathyroidism. A review of the long-term surgical and nonsurgical morbidities as a basis for a rational approach to treatment. Archives of Internal Medicine. 1989; 149(4):789-96

48. Larsson K, Ljunghall S, Krusemo UB, Naessen T, Lindh E, Persson I. The risk of hip fractures in patients with primary hyperparathyroidism: A population-based cohort study with a follow-up of 19 years. Journal of Internal Medicine. 1993; 234(6):585-93

49. Leong KJ, Sam RC, Garnham AW. Health-related quality of life improvement following surgical treatment of primary hyperparathyroidism in a United Kingdom population. Surgeon. 2010; 8(1):5-8

50. Lundstam K, Heck A, Godang K, Mollerup C, Baranowski M, Pernow Y et al. Effect of surgery versus observation: Skeletal 5-year outcomes in a randomized trial of patients with primary HPT (the SIPH study). Journal of Bone and Mineral Research. 2017; 32(9):1907-14

51. Lundstam K, Heck A, Mollerup C, Godang K, Baranowski M, Pernow Y et al. Effects of parathyroidectomy versus observation on the development of vertebral fractures in mild primary hyperparathyroidism. Journal of Clinical Endocrinology and Metabolism. 2015; 100(4):1359-67

52. McDow AD, Sippel RS. Should symptoms be considered an indication for parathyroidectomy in primary hyperparathyroidism? Clinical Medicine Insights. 2018; 11:1179551418785135

53. Melton LJ, 3rd, Atkinson EJ, O'Fallon WM, Heath H, III Risk of age-related fractures in patients with primary hyperparathyroidism. Archives of Internal Medicine. 1992; 152(11):2269-73

54. Mole PA, Walkinshaw MH, Gunn A, Paterson CR. Bone mineral content in patients with primary hyperparathyroidism: A comparison of conservative management with surgical treatment. British Journal of Surgery. 1992; 79(3):263-5

55. Morris GS, Grubbs EG, Hearon CM, Gantela S, Lee JE, Evans DB et al. Parathyroidectomy improves functional capacity in "asymptomatic" older patients with primary hyperparathyroidism: a randomized control trial. Annals of Surgery. 2010; 251(5):832-7

56. National Institute for Health and Care Excellence. Developing NICE guidelines: the manual. London. National Institute for Health and Care Excellence, 2014. Available from: http://www.nice.org.uk/article/PMG20/chapter/1%20Introduction%20and%20overview

57. Nomura R, Sugimoto T, Tsukamoto T, Yamauchi M, Sowa H, Chen Q et al. Marked and sustained increase in bone mineral density after parathyroidectomy in patients with primary hyperparathyroidism; a six-year longitudinal study with or without parathyroidectomy in a Japanese population. Clinical Endocrinology. 2004; 60(3):335-42

58. Nordenstrom E, Westerdahl J, Bergenfelz A. Recovery of bone mineral density in 126 patients after surgery for primary hyperparathyroidism. World Journal of Surgery. 2004; 28(5):502-7

59. Oucharek JJ, O'Neill CJ, Suliburk JW, Sywak MS, Delbridge LW, Sidhu SB. Durability of focused minimally invasive parathyroidectomy in young patients with sporadic primary hyperparathyroidism. Annals of Surgical Oncology. 2011; 18(5):1290-2

60. Paloyan E, Lawrence AM, Oslapas R, Shah KH, Ernst K, Hofmann C. Subtotal parathyroidectomy for primary hyperparathyroidism. Long-term results in 292 patients. Archives of Surgery. 1983; 118(4):425-31

61. Perrier ND, Balachandran D, Wefel JS, Jimenez C, Busaidy N, Morris GS et al. Prospective, randomized, controlled trial of parathyroidectomy versus observation in patients with "asymptomatic" primary hyperparathyroidism. Surgery. 2009; 146(6):1116-22

62. Persson A, Bollerslev J, Rosen T, Mollerup CL, Franco C, Isaksen GA et al. Effect of surgery on cardiac structure and function in mild primary hyperparathyroidism. Clinical Endocrinology. 2011; 74(2):174-80

63. Posen S, Clifton-Bligh P, Reeve TS. Is parathyroidectomy of benefit in primary hyperparathyroidism? Quarterly Journal of Medicine. 1985; 54(215):241-51

64. Rao DS, Phillips ER, Divine GW, Talpos GB. Randomized controlled clinical trial of surgery versus no surgery in patients with mild asymptomatic primary hyperparathyroidism. Journal of Clinical Endocrinology and Metabolism. 2004; 89(11):5415-22

65. Rao DS, Wallace EA, Antonelli RF, Talpos GB, Ansari MR, Jacobsen G et al. Forearm bone density in primary hyperparathyroidism: Long-term follow-up with and without parathyroidectomy. Clinical Endocrinology. 2003; 58(3):348-54

66. Richmond BK, Eads K, Flaherty S, Belcher M, Runyon D. Complications of thyroidectomy and parathyroidectomy in the rural community hospital setting. American Surgeon. 2007; 73(4):332-6

67. Rolighed L, Vestergaard P, Heickendorff L, Sikjaer T, Rejnmark L, Mosekilde L et al. Bone mineral density improvements after operation for primary hyperparathyroidism. Langenbeck's Archives of Surgery. 2012; 397 (5):845

68. Rubin MR, Bilezikian JP, 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

69. Sankaran S, Gamble G, Bolland M, Reid IR, Grey A. Skeletal effects of interventions in mild primary hyperparathyroidism: A meta-analysis. Journal of Clinical Endocrinology and Metabolism. 2010; 95(4):1653-62

70. Sanzenbacher LJ, Pak CY, Bartter FC. Preoperative and postoperative evaluation of patients with normocalcemic primary hyperparathyroidism. Surgical Forum. 1970; 21(4):96-8

71. Saponaro F, Faggiano A, Grimaldi F, Borretta G, Brandi ML, Minisola S et al. Cinacalcet in the management of primary hyperparathyroidism: Post marketing experience of an Italian multicentre group. Clinical Endocrinology. 2013; 79(1):20-6

72. Schneider DF, Mazeh H, Chen H, Sippel RS. Predictors of recurrence in primary hyperparathyroidism: An analysis of 1386 cases. Annals of Surgery. 2014; 259(3):563-8

73. Scott HW, Jr, Richie RE, Crane JM, Rosenfeld L, Jacobs JK, Ginn E et al. Surgical experience with hyperparathyroidism. American Surgeon. 1981; 47(2):54-62

74. Sejean K, Calmus S, Durand-Zaleski I, Bonnichon P, Thomopoulos P, Cormier C et al. Surgery versus medical follow-up in patients with asymptomatic primary hyperparathyroidism: a decision analysis European Journal of Endocrinology. 2005; 153(6):915-27

75. Silverberg SJ, Gartenberg F, Jacobs TP, Shane E, Siris E, Staron RB et al. Increased bone mineral density after parathyroidectomy in primary hyperparathyroidism. Journal of Clinical Endocrinology and Metabolism. 1995; 80(3):729-34

76. Silverberg SJ, Shane E, Jacobs TP, Siris E, Bilezikian JP. A 10-year prospective study of primary hyperparathyroidism with or without parathyroid surgery. New England Journal of Medicine. 1999; 341(17):1249-55

77. Singh Ospina N, Maraka S, Rodriguez-Gutierrez R, Espinosa de Ycaza AE, Jasim S, Gionfriddo M et al. Comparative efficacy of parathyroidectomy and active surveillance in patients with mild primary hyperparathyroidism: a systematic review and meta-analysis. Osteoporosis International. 2016; 27(12):3395-407

78. Singh Ospina NM, Rodriguez-Gutierrez R, Maraka S, Espinosa de Ycaza AE, Jasim S, Castaneda-Guarderas A et al. Outcomes of parathyroidectomy in patients with primary hyperparathyroidism: A systematic review and meta-analysis. World Journal of Surgery. 2016; 40(10):2359-77

79. Siperstein AE, Shen W, Chan AK, Duh QY, Clark OH, Giuliano AE et al. Normocalcemic hyperparathyroidism: Biochemical and symptom profiles before and after surgery. Archives of Surgery. 1992; 127(10):1157-63

80. Solorzano CC, Mendez W, Lew JI, Rodgers SE, Montano R, Carneiro-Pla DM et al. Long-term outcome of patients with elevated parathyroid hormone levels after successful parathyroidectomy for sporadic primary hyperparathyroidism. Archives of Surgery. 2008; 143(7):659-63

81. Soreide JA, Van Heerden JA, Grant CS, Lo CY, Schleck C, Ilstrup DM. Survival after surgical treatment for primary hyperparathyroidism. Surgery. 1997; 122(6):1117-23

82. Strewler GJ. Indications for surgery in patients with minimally symptomatic primary hyperparathyroidism. Surgical Clinics of North America. 1995; 75(3):439-47

83. Talpos GB, Bone HG, Kleerekoper M, Phillips ER, Alam M, Honasoge M et al. Randomized trial of parathyroidectomy in mild asymptomatic primary hyperparathyroidism: patient description and effects on the SF-36 health survey. Surgery. 2000; 128(6):1013-20;discussion 1020-1

84. Tay YK, Khoo J, Chandran M. Surgery or no surgery: What works best for the kidneys in primary hyperparathyroidism? A study in a multi-ethnic Asian population. Indian Journal of Endocrinology and Metabolism. 2016; 20(1):55-61

85. Tisell LE. Results of surgical and medical treatment in primary hyperparathyroidism. Annales Chirurgiae et Gynaecologiae. 1983; 72(3):129-34

86. Trombetti A, Christ ER, Henzen C, Gold G, Brandle M, Herrmann FR et al. Clinical presentation and management of patients with primary hyperparathyroidism of the Swiss Primary Hyperparathyroidism Cohort: A focus on neuro-behavioral and cognitive symptoms. Journal of Endocrinological Investigation. 2016; 39(5):567-76

87. VanderWalde LH, Liu IL, O'Connell TX, Haigh PI. The effect of parathyroidectomy on bone fracture risk in patients with primary hyperparathyroidism. Archives of Surgery. 2006; 141(9):885-9

88. VanderWalde LH, Liu ILA, Haigh PI. Effect of bone mineral density and parathyroidectomy on fracture risk in primary hyperparathyroidism. World Journal of Surgery. 2009; 33(3):406-11

89. Vera L, Accornero M, Dolcino M, Oddo S, Giusti M. Five-year longitudinal evaluation of mild primary hyperparathyroidism - Medical treatment versus clinical observation. Endokrynologia Polska. 2014; 65(6):456-63

90. Vestergaard P, Mosekilde L. Cohort study on effects of parathyroid surgery on multiple outcomes in primary hyperparathyroidism. BMJ. 2003; 327(7414):530-3

91. Vestergaard P, Mosekilde L. Fractures in patients with primary hyperparathyroidism: Nationwide follow-up study of 1201 patients. World Journal of Surgery. 2003; 27(3):343-9

92. Wagner K, Miskulin J. Effect of parathyroidectomy on osteoporosis and bone metabolism. Clinical Reviews in Bone and Mineral Metabolism. 2007; 5(2):115-21

93. Wermers RA, Khosla S, Atkinson EJ, Grant CS, Hodgson SF, O'Fallon WM et al. Survival after the diagnosis of hyperparathyroidism: A population-based study. American Journal of Medicine. 1998; 104(2):115-22

94. Witteveen JE, Kievit J, Morreau H, Romijn JA, Hamdy NAT. No recurrence of sporadic primary hyperparathyroidism when cure is established 6 months after parathyroidectomy. European Journal of Endocrinology. 2010; 162(2):399-406

95. Wu B, Haigh PI, Hwang R, Ituarte PHG, Liu ILA, Hahn TJ et al. Underutilization of parathyroidectomy in elderly patients with primary hyperparathyroidism. Journal of Clinical Endocrinology and Metabolism. 2010; 95(9):4324-30

96. Yeh MW, Zhou H, Adams AL, Ituarte PHG, Li N, Liu ILA et al. The relationship of parathyroidectomy and bisphosphonates with fracture risk in primary hyperparathyroidism: An observational study. Annals of Internal Medicine. 2016; 164(11):715-23

97. Yu N, Donnan PT, Flynn RWV, Murphy MJ, Smith D, Rudman A et al. Increased mortality and morbidity in mild primary hyperparathyroid patients. the Parathyroid Epidemiology and Audit Research Study (PEARS). Clinical Endocrinology. 2010; 73(1):30-4

98. Zhao DW, Yen TW, Doffek K, Evans DB, Wang TS. Changes in bone mineral density after parathyroidectomy in elderly patients with primary hyperparathyroidism. Journal of Surgical Research. 2014; 186 (2):557

# 1 Appendices

# 2 Appendix A: Review protocols

3 **Table 7: Review protocol: Surgery**

| Field | Content |
|---|---|
| Review question | What is the clinical and cost effectiveness of surgery (parathyroidectomy) in people with primary hyperparathyroidism? |
| Type of review question | Intervention |
| Objective of the review | To determine the clinical and cost effectiveness of parathyroidectomy versus conservative management or pharmacological intervention. To determine whether surgery should be recommended in all people with PHPT, or only subgroups of people with certain indications and poorer prognosis. |
| Eligibility criteria – population | Adults (18 years or over) with confirmed primary hyperparathyroidism<br><br>Strata (report the following groups separately):<br>• People with normocalcaemic PHPT (serum adjusted calcium ≤2.6mmol/L and an elevated PTH that cannot be explained by abnormal renal function or low 25OHD)<br>• Previous unsuccessful parathyroidectomy (reoperation)<br>• Pregnant women<br><br>Exclude people:<br>• with secondary and tertiary HPT<br>• with multiple endocrine neoplasia (MEN)<br>• with familial hyperparathyroidism<br>• with parathyroid carcinoma<br>• Taking medications interfering with calcium metabolism (for example, lithium).<br><br>Studies including mixed populations of people with primary and secondary or tertiary hyperparathyroidism will be excluded unless subgroups reported separately by type of hyperparathyroidism. |
| Eligibility criteria – intervention(s) | Parathyroid surgery (all types of surgery grouped within class, to include minimally invasive surgeries or unilateral or bilateral exploratory surgery) |
| Eligibility criteria – comparator(s) | • no surgery (surveillance/conservative management)<br>• calcimimetic treatment<br>• bisphosphonate treatment<br>• combination pharmacological treatment (calcimimetics and bisphosphonates)<br><br>The above comparators will not be pooled in the analysis |
| Outcomes and prioritisation | **Report all outcomes separately for <6 months and ≥6 months**<br><br>**Critical outcomes:**<br>HRQOL (continuous outcome)<br>Mortality (dichotomous outcome)<br>Preservation of end organ function (bone mineral density, fractures, renal stones and renal function) (dichotomous for fractures, renal function, renal stones and continuous for BMD)<br>**Important outcomes:**<br>Adverse events (to include voice change, hypoparathyroidism; dichotomous |

| | |
|---|---|
| | outcome)<br>Cancer incidence (dichotomous outcome)<br>Cardiovascular events (dichotomous outcome)<br>Persistent hypercalcaemia (dichotomous outcome) |
| Eligibility criteria – study design | RCTs and systematic reviews of RCTs<br><br>In the absence of RCT evidence for the critical outcomes, NRS will be included (only if the following key confounders are matched for or adjusted for in the analysis)<br>Key confounders:<br>• Age<br>• Absence/presence of end-organ effects<br>• Adjusted serum calcium level |
| Other inclusion exclusion criteria | Non-English language articles<br>Conference abstracts |
| Proposed sensitivity / subgroup analysis, or meta-regression | Subgroups will be investigated in the following order if there is heterogeneity in the data:<br>• People with end-organ effects vs absence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)<br>• serum adjusted calcium > 0.25 mmol/L above the ULN (same as ≥2.85mmol/L and <2.85mmol/L)<br>• reduction in creatinine clearance to < 60 mL/min<br>• age under 50 years vs ≥50 years |
| Selection process – duplicate screening / selection / analysis | Studies are sifted by title and abstract. Potentially significant publications obtained in full text are then assessed against the inclusion criteria specified in this protocol |
| Data management (software) | • Pairwise meta-analyses were performed using Cochrane Review Manager (RevMan5).<br>• GRADEpro was used to assess the quality of evidence for each outcome.<br>• Endnote for bibliography, citations, sifting and reference management<br>• Data extractions performed using EviBase, a platform designed and maintained by the National Guideline Centre (NGC) |
| Information sources – databases and dates | Clinical search databases to be used: Medline, Embase, Cochrane Library, CINAHL, PsycINFO<br>Date: all years<br><br>Health economics search databases to be used: Medline, Embase, NHSEED, HTA<br>Date: Medline, Embase from 2002<br>NHSEED, HTA – all years<br><br>Language: Restrict to English only<br>Supplementary search techniques: backward citation searching |
| Identify if an update | N/A |
| Author contacts | https://www.nice.org.uk/guidance/indevelopment/gid-ng10051 |
| Highlight if amendment to | N/A |

| | |
|---|---|
| previous protocol | |
| Search strategy – for one database | For details please see appendix B |
| Data collection process – forms / duplicate | A standardised evidence table format will be used, and published as appendix D of the evidence report. |
| Data items – define all variables to be collected | For details please see evidence tables in Appendix D (clinical evidence tables) or H (health economic evidence tables). |
| Methods for assessing bias at outcome / study level | Standard study checklists were used to critically appraise individual studies. For details please see section 6.2 of Developing NICE guidelines: the manual<br>The risk of bias across all available evidence was evaluated for each outcome using an adaptation of the 'Grading of Recommendations Assessment, Development and Evaluation (GRADE) toolbox' developed by the international GRADE working group http://www.gradeworkinggroup.org/ |
| Criteria for quantitative synthesis | For details please see section 6.4 of Developing NICE guidelines: the manual. |
| Methods for quantitative analysis – combining studies and exploring (in)consistency | For details please see the separate Methods report for this guideline. |
| Meta-bias assessment – publication bias, selective reporting bias | For details please see section 6.2 of Developing NICE guidelines: the manual. |
| Confidence in cumulative evidence | For details please see sections 6.4 and 9.1 of Developing NICE guidelines: the manual. |
| Rationale / context – what is known | For details please see the introduction to the evidence review. |
| Describe contributions of authors and guarantor | A multidisciplinary committee developed the evidence review. The committee was convened by the National Guideline Centre (NGC) and chaired by Jonathan Mant in line with section 3 of Developing NICE guidelines: the manual.<br>Staff from NGC undertook systematic literature searches, appraised the evidence, conducted meta-analysis and cost-effectiveness analysis where appropriate, and drafted the evidence review in collaboration with the committee. For details please see Developing NICE guidelines: the manual. |
| Sources of funding / support | NGC is funded by NICE and hosted by the Royal College of Physicians. |
| Name of sponsor | NGC is funded by NICE and hosted by the Royal College of Physicians. |
| Roles of sponsor | NICE funds NGC to develop guidelines for those working in the NHS, public health and social care in England. |
| PROSPERO registration | Not registered |

| | |
|---|---|
| number | |

1

2 **Table 8: Health economic review protocol**

| **Review question** | **All questions – health economic evidence** |
|---|---|
| **Objectives** | To identify health economic studies relevant to any of the review questions. |
| **Search criteria** | • Populations, interventions and comparators must be as specified in the clinical review protocol above.<br>• Studies must be of a relevant health economic study design (cost–utility analysis, cost-effectiveness analysis, cost–benefit analysis, cost–consequences analysis, comparative cost analysis).<br>• Studies must not be a letter, editorial or commentary, or a review of health economic evaluations. (Recent reviews will be ordered although not reviewed. The bibliographies will be checked for relevant studies, which will then be ordered.)<br>• Unpublished reports will not be considered unless submitted as part of a call for evidence.<br>Studies must be in English. |
| **Search strategy** | A health economic study search will be undertaken using population-specific terms and a health economic study filter – see appendix B below. |
| **Review strategy** | Studies not meeting any of the search criteria above will be excluded. Studies published before 2002, abstract-only studies and studies from non-OECD countries or the USA will also be excluded.<br>Each remaining study will be assessed for applicability and methodological limitations using the NICE economic evaluation checklist which can be found in appendix H of Developing NICE guidelines: the manual (2014).[56]<br><br>**Inclusion and exclusion criteria**<br>• If a study is rated as both 'Directly applicable' and with 'Minor limitations' then it will be included in the guideline. A health economic evidence table will be completed and it will be included in the health economic evidence profile.<br>• If a study is rated as either 'Not applicable' or with 'Very serious limitations' then it will usually be excluded from the guideline. If it is excluded then a health economic evidence table will not be completed and it will not be included in the health economic evidence profile.<br>• If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included.<br><br>**Where there is discretion**<br>The health economist will make a decision based on the relative applicability and quality of the available evidence for that question, in discussion with the guideline committee if required. The ultimate aim is to include health economic studies that are helpful for decision-making in the context of the guideline and the current NHS setting. If several studies are considered of sufficiently high applicability and methodological quality that they could all be included, then the health economist, in discussion with the committee if required, may decide to include only the most applicable studies and to selectively exclude the remaining studies. All studies excluded on the basis of applicability or methodological limitations will be listed with explanation in the excluded health economic studies appendix below.<br><br>The health economist will be guided by the following hierarchies. |

| Review question | All questions – health economic evidence |
|---|---|
| | *Setting:*<br>• UK NHS (most applicable).<br>• OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden).<br>• OECD countries with predominantly private health insurance systems (for example, Switzerland).<br>• Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations.<br>*Health economic study type:*<br>• Cost–utility analysis (most applicable).<br>• Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis).<br>• Comparative cost analysis.<br>• Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations.<br>*Year of analysis:*<br>• The more recent the study, the more applicable it will be.<br>• Studies published in 2002 or later but that depend on unit costs and resource data entirely or predominantly from before 2002 will be rated as 'Not applicable'.<br>• Studies published before 2002 will be excluded before being assessed for applicability and methodological limitations.<br>*Quality and relevance of effectiveness data used in the health economic analysis:*<br>• The more closely the clinical effectiveness data used in the health economic analysis match with the outcomes of the studies included in the clinical review the more useful the analysis will be for decision-making in the guideline. |

# Appendix B: Literature search strategies

The literature searches for this review are detailed below and complied with the methodology outlined in Developing NICE guidelines: the manual 2014, updated 2017 https://www.nice.org.uk/guidance/pmg20/resources/developing-nice-guidelines-the-manual-pdf-72286708700869

*For more detailed information, please see the Methodology Review*.

## B.1 Clinical search literature search strategy

Searches were constructed using a PICO framework where population (P) terms were combined with Intervention (I) and in some cases Comparison (C) terms. Outcomes (O) are rarely used in search strategies for interventions as these concepts may not be well described in title, abstract or indexes and therefore difficult to retrieve. Search filters were applied to the search where appropriate.

**Table 9: Database date parameters and filters used**

| Database | Dates searched | Search filter used |
|---|---|---|
| Medline (OVID) | 1946 – 06 August 2018 | Exclusions |
| Embase (OVID) | 1974 – 06 August 2018 | Exclusions |

| Database | Dates searched | Search filter used |
|---|---|---|
| The Cochrane Library (Wiley) | Cochrane Reviews to 2018 Issue 8 of 12<br>CENTRAL to 2018 Issue 7 of 12<br>DARE, and NHSEED to 2015 Issue 2 of 4<br>HTA to 2016 Issue 4 of 4 | None |
| CINAHL, Current Nursing and Allied Health Literature (EBSCO) | Inception – 06 August 2018 | Exclusions |
| PsycINFO (ProQuest) | Inception – 06 August 2018 | Exclusions |

1 **Medline (Ovid) search terms**

| | |
|---|---|
| 1. | hyperparathyroidism/ or hyperparathyroidism, primary/ |
| 2. | ((primary or asymptomatic or symptomatic or mild or familial or maternal) adj6 (HPT or hyperparathyroidis*)).ti,ab. |
| 3. | PHPT.ti,ab. |
| 4. | Parathyroid Neoplasms/ |
| 5. | (parathyroid* adj3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumo?r* or cancer* or metasta* or hypercalc?emi*)).ti,ab. |
| 6. | or/1-5 |
| 7. | letter/ |
| 8. | editorial/ |
| 9. | news/ |
| 10. | exp historical article/ |
| 11. | Anecdotes as Topic/ |
| 12. | comment/ |
| 13. | case report/ |
| 14. | (letter or comment*).ti. |
| 15. | or/7-14 |
| 16. | randomized controlled trial/ or random*.ti,ab. |
| 17. | 15 not 16 |
| 18. | animals/ not humans/ |
| 19. | exp Animals, Laboratory/ |
| 20. | exp Animal Experimentation/ |
| 21. | exp Models, Animal/ |
| 22. | exp Rodentia/ |
| 23. | (rat or rats or mouse or mice).ti. |
| 24. | or/17-23 |
| 25. | 6 not 24 |
| 26. | limit 25 to English language |

2 **Embase (Ovid) search terms**

| | |
|---|---|
| 1. | hyperparathyroidism/ or primary hyperparathyroidism/ |
| 2. | ((primary or asymptomatic or symptomatic or mild or familial or maternal) adj6 (HPT or hyperparathyroidis*)).ti,ab. |
| 3. | PHPT.ti,ab. |
| 4. | parathyroid tumor/ or parathyroid adenoma/ or parathyroid carcinoma/ |

| | |
|---|---|
| 5. | (parathyroid* adj3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumo?r* or cancer* or metasta* or hypercalc?emi*)).ti,ab. |
| 6. | or/1-5 |
| 7. | letter.pt. or letter/ |
| 8. | note.pt. |
| 9. | editorial.pt. |
| 10. | Case report/ or Case study/ |
| 11. | (letter or comment*).ti. |
| 12. | or/7-11 |
| 13. | randomized controlled trial/ or random*.ti,ab. |
| 14. | 12 not 13 |
| 15. | animal/ not human/ |
| 16. | Nonhuman/ |
| 17. | exp Animal Experiment/ |
| 18. | exp Experimental animal/ |
| 19. | Animal model/ |
| 20. | exp Rodent/ |
| 21. | (rat or rats or mouse or mice).ti. |
| 22. | or/14-21 |
| 23. | 6 not 22 |
| 24. | limit 23 to English language |

1 **Cochrane Library (Wiley) search terms**

| | |
|---|---|
| #1. | MeSH descriptor: [Hyperparathyroidism] explode all trees |
| #2. | MeSH descriptor: [Hyperparathyroidism, Primary] explode all trees |
| #3. | ((primary or asymptomatic or symptomatic or mild or familial or maternal) near/6 (HPT or hyperparathyroidis*)):ti,ab |
| #4. | PHPT:ti,ab |
| #5. | MeSH descriptor: [Parathyroid Neoplasms] explode all trees |
| #6. | (parathyroid* near/3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumo?r* or cancer* or metasta* or hypercalc?emi*)):ti,ab |
| #7. | (or #1-#6) |

2 **CINAHL (EBSCO) search terms**

| | |
|---|---|
| S1. | (MH "Hyperparathyroidism") |
| S2. | ( (primary or asymptomatic or symptomatic or mild or familial or maternal) n6 HPT ) OR ( (primary or asymptomatic or symptomatic or mild or familial or maternal) n6 hyperparathyroidis* ) |
| S3. | PHPT |
| S4. | (MH "Parathyroid Neoplasms") |
| S5. | (parathyroid* n3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumor* or tumour* or cancer* or metasta* or hypercalcemi* or hypercalcaemi*)) |
| S6. | S1 OR S2 OR S3 OR S4 OR S5 |
| S7. | PT anecdote or PT audiovisual or PT bibliography or PT biography or PT book or PT book review or PT brief item or PT cartoon or PT commentary or PT computer program or PT editorial or PT games or PT glossary or PT historical material or PT interview or PT letter or PT listservs or PT masters thesis or PT obituary or PT pamphlet or PT pamphlet chapter or PT pictorial or PT poetry or PT proceedings or PT “questions and answers” or PT response or PT software or PT teaching materials or PT website |

| | |
|---|---|
| S8. | S6 NOT S7 |

**PsycINFO (ProQuest) search terms**

| | |
|---|---|
| 1. | su.Exact("parathyroid neoplasms" OR "hyperparathyroidism" OR "hyperparathyroidism, primary") |
| 2. | PHPT |
| 3. | ((primary or asymptomatic or symptomatic or mild or familial or maternal) Near/6 (HPT or hyperparathyroidis*)) |
| 4. | (parathyroid* near/3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumor* or tumour* or cancer* or metasta* or hypercalcaemi* or hypercalcemi*)) |
| 5. | 1 or 2 or 3 or 4 |
| 6. | (su.exact.explode("rodents") or su.exact.explode("mice") or (su.exact("animals") not (su.exact("human males") or su.exact("human females"))) or ti(rat or rats or mouse or mice)) |
| 7. | (s1 or s2 or s3 or s4) NOT (su.exact.explode("rodents") or su.exact.explode("mice") or (su.exact("animals") not (su.exact("human males") or su.exact("human females"))) or ti(rat or rats or mouse or mice)) |

# B.2 Health Economics literature search strategy

Health economic evidence was identified by conducting a broad search relating to primary hyperparathyroidism population in NHS Economic Evaluation Database (NHS EED – this ceased to be updated after March 2015) and the Health Technology Assessment database (HTA) with no date restrictions. NHS EED and HTA databases are hosted by the Centre for Research and Dissemination (CRD). Additional searches were run on Medline and Embase for health economics papers published since 2002.

**Table 10: Database date parameters and filters used**

| Database | Dates searched | Search filter used |
|---|---|---|
| Medline | 2002 – 06 August 2018 | Exclusions<br>Health economics studies |
| Embase | 2002 – 06 August 2018 | Exclusions<br>Health economics studies |
| Centre for Research and Dissemination (CRD) | HTA - Inception – 06 August 2018<br>NHSEED - Inception to March 2015 | None |

**Medline (Ovid) search terms**

| | |
|---|---|
| 1. | hyperparathyroidism/ or hyperparathyroidism, primary/ |
| 2. | ((primary or asymptomatic or symptomatic or mild or familial or maternal) adj6 (HPT or hyperparathyroidis*)).ti,ab. |
| 3. | PHPT.ti,ab. |
| 4. | Parathyroid Neoplasms/ |
| 5. | (parathyroid* adj3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumo?r* or cancer* or metasta* or hypercalc?emi*)).ti,ab. |
| 6. | or/1-5 |
| 7. | letter/ |
| 8. | editorial/ |
| 9. | news/ |
| 10. | exp historical article/ |

| | |
|---|---|
| 11. | Anecdotes as Topic/ |
| 12. | comment/ |
| 13. | case report/ |
| 14. | (letter or comment*).ti. |
| 15. | or/7-14 |
| 16. | randomized controlled trial/ or random*.ti,ab. |
| 17. | 15 not 16 |
| 18. | animals/ not humans/ |
| 19. | exp Animals, Laboratory/ |
| 20. | exp Animal Experimentation/ |
| 21. | exp Models, Animal/ |
| 22. | exp Rodentia/ |
| 23. | (rat or rats or mouse or mice).ti. |
| 24. | or/17-23 |
| 25. | 6 not 24 |
| 26. | limit 25 to English language |
| 27. | Economics/ |
| 28. | Value of life/ |
| 29. | exp "Costs and Cost Analysis"/ |
| 30. | exp Economics, Hospital/ |
| 31. | exp Economics, Medical/ |
| 32. | Economics, Nursing/ |
| 33. | Economics, Pharmaceutical/ |
| 34. | exp "Fees and Charges"/ |
| 35. | exp Budgets/ |
| 36. | budget*.ti,ab. |
| 37. | cost*.ti. |
| 38. | (economic* or pharmaco?economic*).ti. |
| 39. | (price* or pricing*).ti,ab. |
| 40. | (cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab. |
| 41. | (financ* or fee or fees).ti,ab. |
| 42. | (value adj2 (money or monetary)).ti,ab. |
| 43. | or/27-42 |
| 44. | 26 and 43 |

1 **Embase (Ovid) search terms**

| | |
|---|---|
| 1. | hyperparathyroidism/ or primary hyperparathyroidism/ |
| 2. | ((primary or asymptomatic or symptomatic or mild or familial or maternal) adj6 (HPT or hyperparathyroidis*)).ti,ab. |
| 3. | PHPT.ti,ab. |
| 4. | parathyroid tumor/ or parathyroid adenoma/ or parathyroid carcinoma/ |
| 5. | (parathyroid* adj3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumo?r* or cancer* or metasta* or hypercalc?emi*)).ti,ab. |
| 6. | or/1-5 |
| 7. | letter.pt. or letter/ |

| | |
|---|---|
| 8. | note.pt. |
| 9. | editorial.pt. |
| 10. | Case report/ or Case study/ |
| 11. | (letter or comment*).ti. |
| 12. | or/7-11 |
| 13. | randomized controlled trial/ or random*.ti,ab. |
| 14. | 12 not 13 |
| 15. | animal/ not human/ |
| 16. | Nonhuman/ |
| 17. | exp Animal Experiment/ |
| 18. | exp Experimental animal/ |
| 19. | Animal model/ |
| 20. | exp Rodent/ |
| 21. | (rat or rats or mouse or mice).ti. |
| 22. | or/14-21 |
| 23. | 6 not 22 |
| 24. | limit 23 to English language |
| 25. | health economics/ |
| 26. | exp economic evaluation/ |
| 27. | exp health care cost/ |
| 28. | exp fee/ |
| 29. | budget/ |
| 30. | funding/ |
| 31. | budget*.ti,ab. |
| 32. | cost*.ti. |
| 33. | (economic* or pharmaco?economic*).ti. |
| 34. | (price* or pricing*).ti,ab. |
| 35. | (cost* adj2 (effective* or utilit* or benefit* or minimi* or unit* or estimat* or variable*)).ab. |
| 36. | (financ* or fee or fees).ti,ab. |
| 37. | (value adj2 (money or monetary)).ti,ab. |
| 38. | or/25-37 |
| 39. | 24 and 38 |

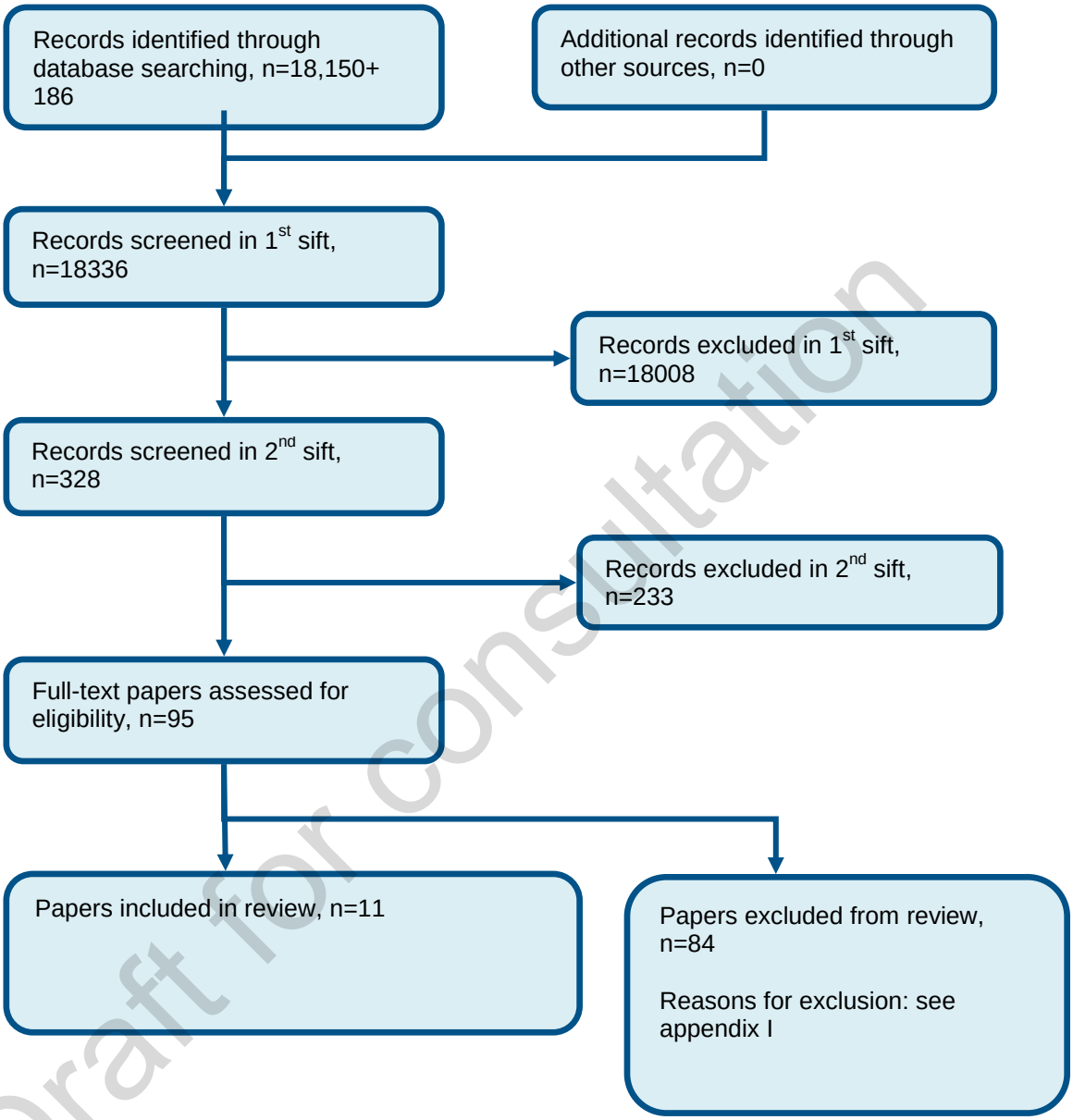
## 1 NHS EED and HTA (CRD) search terms

| | |
|---|---|
| #1. | MeSH DESCRIPTOR Hyperparathyroidism EXPLODE ALL TREES |
| #2. | MeSH DESCRIPTOR Hyperparathyroidism, Primary EXPLODE ALL TREES |
| #3. | (((primary or asymptomatic or symptomatic or mild or familial or maternal) adj6 (HPT or hyperparathyroidis*))) |
| #4. | (PHPT) |
| #5. | MeSH DESCRIPTOR Parathyroid Neoplasms EXPLODE ALL TREES |
| #6. | ((parathyroid* adj3 (adenoma* or carcinoma* or hyperplasia* or neoplas* or tumo?r* or cancer* or metasta* or hypercalc?emi*))) |

| | |
|---|---|
| #7. | #1 OR #2 OR #3 OR #4 OR #5 OR #6 |
| #8. | * IN NHSEED |
| #9. | * IN HTA |
| #10. | #7 AND #8 |
| #11. | #7 AND #9 |

# Appendix C: Clinical evidence selection

**Figure 1: Flow chart of clinical study selection for the review of surgery**



# Appendix D: Clinical evidence tables

| Study | Ambrogini 2007[7] |
|---|---|
| Study type | RCT (Patient randomised; Parallel) |
| Number of studies (number of participants) | N/A (n=50) |
| Countries and setting | Conducted in Italy; Setting: Referral centre |
| Line of therapy | Mixed line |
| Duration of study | Intervention + follow up: Patients followed up to 1 year post-surgery (6 month intervals) |
| Method of assessment of guideline condition | Adequate method of assessment/diagnosis: The PHPT diagnosis was based on increased ionised (>132mmol/L) or albumin-corrected serum calcium (>10.2mg/dL [2.55mmol/L]), with increased (>65pg/mL [65ng/L]) or inappropriately normal intact parathyroid hormone. |
| Stratum | Overall |
| Subgroup analysis within study | Post-hoc subgroup analysis: Presence/absence of osteoporosis |
| Inclusion criteria | Patients with mild PHPT who did not meet any of the National Institutes of Health (NIH) criteria for surgery. Asymptomatic PHPT; albumin-corrected serum calcium of <1mg/dL above the upper limit of normal (11.2mg/dL [2.8mmol/L]) on ≥3 occasions; 24-hour urine calcium excretion <400mg (10mmol); creatinine clearance in the normal range or reduce by ≤30% compared with age-matched normal people; age- and sex-matched BMD at the distal third of radius to be Z>-2.0; age between 50 and 75 years |
| Exclusion criteria | Symptomatic disease (nephrolithiasis, osteitis fibrosa cystica, prevalent fragility fractures); familial PHPT; menopause <3 years; disease/therapies affecting the skeleton; current thyroid disease requiring surgery; contraindications to surgery; previous neck surgery |
| Recruitment/selection of patients | Between January 2002 and September 2005, 412 consecutive patients with PHPT were referred to the Department of Endocrinology at the University Hospital of Pisa. Of these individuals, 198 already met the National Institutes of Health (NIH) criteria for surgery. Of the 214 potentially eligible patients, 161 were excluded for several reasons, and the remaining 53 were asked to participate in the study |
| Age, gender and ethnicity | Age - Mean (SD): Intervention = 64 (6) vs. Control = 65 (6). Gender (M:F): 4:46. Ethnicity: Not reported |
| Further population details | 1. Adjusted serum calcium: <2.85mmol/L 2. Age: ≥50 years old 3. Creatinine clearance: ≥ 60 mL/min (study reports as not less than 30% age-matched value). 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): mixed (people with kidney stones and fractures excluded, some people had osteoporosis but subgroups analysis done within study) (Based on guidelines prior to 2002 so does not exclude people with osteoporosis |

| Study | Ambrogini 2007[7] |
|---|---|
| | [subgroup analysis done of people with osteoporosis]. Does not exclude people with osteoporosis based on the T score but does exclude people with low BMD Z score <-2). |
| Extra comments | [The study began before the 2002 Workshop on Asymptomatic PHPT, therefore, the older guidelines formed the basis for the inclusion criteria. Had the criteria of the 2002 Workshop on Asymptomatic PHPT been adopted, 29 of the 50 participants would have met these criteria for surgery.] |
| Indirectness of population | No indirectness |
| Interventions | (n=24) Intervention 1: Surgery (parathyroidectomy) - minimally invasive surgery. Two experienced parathyroid surgeons performed all surgery, using the minimally invasive approach when the abnormal gland was identified by pre-operative imaging. Four of the 24 subjects who underwent surgery required standard neck exploration because of equivocal or negative pre-operative imaging studies. Duration Single surgery. Concurrent medication/care: No patient was given oral calcium supplements. Indirectness: No indirectness<br><br>(n=26) Intervention 2: Conservative management. Not described. Duration N/A. Concurrent medication/care: Not described. Indirectness: No indirectness<br>Comments: Details about care have not been provided for this control group. |
| Funding | Academic or government funding (Ministero dell'Istruzione, dell'Universita e della Ricerca Scientifica Rome and the University of Pisa) |

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: MINIMALLY INVASIVE SURGERY versus NO SURGERY

Protocol outcome 1: Quality of life
- Actual outcome: Quality of life (SF-36) at 6 months post-surgery; 0 - 100 Top=High is good outcome; The results were reported as graphs and not as numerical values. Significant beneficial effect of surgery on QOL for the following domains: bodily pain (P=0.001), general health (P=0.008), vitality (P=0.003), mental health (P=0.017);
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation.
- Actual outcome: Psychosocial well-being (SCL-90R) at 6 months post-surgery; 0 - 100 Top=High is good outcome; The results were reported as statements about whether there were any differences between the two groups (and p values for some of the domains), and no numerical values were reported.;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - High, Measurement - Low, Crossover - Low, Comments - It is indicated that no difference was found between the two groups but this is neither supported by numbers nor charts.; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation.

| Study | Ambrogini 2007[7] |
|---|---|

Protocol outcome 2: Fractures (vertebral or long bone)
- Actual outcome: Clinical vertebral fragility fracture at During 1 year post-surgery; Group 1: 0/24, Group 2: 1/25
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation.

Protocol outcome 3: Occurrence of kidney stones
- Actual outcome: Kidney stones at During 1 year post-surgery; Group 1: 0/24, Group 2: 1/25
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation.

Protocol outcome 4: Bone mineral density (BMD; distal radius or lumbar spine)
- Actual outcome: Lumbar spine (L1-L4) BMD at 1 year post-surgery (change score – described as % change from baseline (% change of g/cm$^2$ presumed)); Group 1: mean 4.16 % (SD 1.1); n=24, Group 2: mean -1.12 % (SD 0.71); n=25; Comments: p=0.0002
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation.
- Actual outcome: Distal radius BMD at 1 year post-surgery (change score - described as % change from baseline (% change of g/cm$^2$ presumed)); Group 1: mean -0.34 % (SD 0.59); n=24, Group 2: mean -0.55 % (SD 0.53); n=25; Comments: p=0.68
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation.

Protocol outcome 5: Adverse events (including voice change and hypoparathyroidism)
- Actual outcome: Surgical complications (such as laryngeal nerve dysfunction) at During 1 year post-surgery; Group 1: 0/24, Group 2: 0/25
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation and not analysed due to chemotherapy

Protocol outcome 6: Cancer
- Actual outcome: chronic myeloid leukaemia at During 1 year post-surgery; Group 1: 0/24, Group 2: 1/25
Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 0, Reason: N/A; Group 2 Number missing: 1, Reason: One patient developed chronic myeloid leukaemia 4 months after randomisation.

| Study | Ambrogini 2007[7] |
|---|---|
| Protocol outcomes not reported by the study | Mortality; Deterioration in renal function; Persistent hypercalcaemia; Cardiovascular events |

1

| Study | Clifton-Bligh 2015[27] |
|---|---|
| Study type | Non-randomised comparative study |
| Number of studies (number of participants) | 1 (n=561) |
| Countries and setting | Conducted in Australia; Setting: Hospital |
| Line of therapy | 1st line |
| Duration of study | Follow up (post intervention): average follow-up not reported |
| Method of assessment of guideline condition | Adequate method of assessment/diagnosis: Before 1972 the diagnosis of PHPT was made if surgical removal of a parathyroid tumour restored eucalcaemia, or if full investigation failed to find another cause of hypercalcaemia; after 1972 the diagnosis of PHPT was made if the serum calcium and serum PTH were above the upper limit of the reference range. |
| Stratum | Overall |
| Subgroup analysis within study | Not applicable |
| Inclusion criteria | Diagnosed with PHPT (before 1972 the diagnosis of PHPT was made if surgical removal of a parathyroid tumour restored eucalcaemia, or if full investigation failed to find another cause of hypercalcaemia; after 1972 the diagnosis of PHPT was made if the serum calcium and serum PTH were above the upper limit of the reference range). |
| Exclusion criteria | Not reported |
| Recruitment/selection of patients | All patients diagnosed with PHPT between 1961 and 1994. Medical records were obtained and death registers checked. |
| Age, gender and ethnicity | Age - Mean (SD): Surgery: 52.9 (14.7); non-surgery: 55.5 (15.9). Gender (M:F): Not reported. Ethnicity: not reported |
| Further population details | 1. Adjusted serum calcium: Not stated / Unclear 2. Age: Not stated / Unclear 3. Creatinine clearance: Not stated / Unclear 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): Not stated / Unclear |
| Indirectness of population | No indirectness |
| Interventions | (n=448) Intervention 1: Surgery (parathyroidectomy) - 4-gland or bilateral exploration: not reported. Duration one off surgery (average follow-up not reported). Concurrent medication/care: not reported. Indirectness: No indirectness |

| Study | Clifton-Bligh 2015[27] |
|---|---|
| | (n=113) Intervention 2: Conservative management. Duration average follow-up not reported. Concurrent medication/care: not reported. Indirectness: No indirectness |
| Funding | Funding not stated |
| RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARATHYROIDECTOMY versus CONSERVATIVE MANAGEMENT<br><br>Protocol outcome 1: Mortality<br>- Actual outcome: Death register record at not reported; Group 1: n=448 ; Group 2: n=113; HR 0.67; Lower CI 0.38 to Upper CI 1.18; Comments: Compared with the non-surgically treated group, the hazard ratio of death for the surgically treated group adjusted for age sex and time of diagnosis was 0.67 (0.38-1.18; P=0.167) (Cox proportional hazard multivariate analysis)<br>Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: There was no significant difference in age between groups but the serum calcium and the serum PTH were significantly lower in the non-surgically treated group.; Group 1 Number missing: 0; Group 2 Number missing: 0 | |
| Protocol outcomes not reported by the study | Quality of life; Deterioration in renal function; Fractures (vertebral or long bone); Occurrence of kidney stones; Persistent hypercalcaemia; Bone mineral density (BMD; distal radius or lumbar spine); Cardiovascular events; Adverse events (including voice change and hypoparathyroidism); Cancer |

1

| Study | Elvius 1995[34] |
|---|---|
| Study type | RCT (Patient randomised; Parallel) |
| Number of studies (number of participants) | N/A (n=48) |
| Countries and setting | Conducted in Sweden; Setting: Health screening programme |
| Line of therapy | 1st line |
| Duration of study | Intervention + follow up: Single surgery then 3 years of follow-up |
| Method of assessment of guideline condition | Method of assessment /diagnosis not stated: No detail given on how hyperparathyroidism was diagnosed, except to report that female patients with moderately raised serum calcium concentrations who were free of symptoms of the disease were randomised. |
| Stratum | Overall |
| Subgroup analysis within study | Not applicable |
| Inclusion criteria | Not provided |

| Study | Elvius 1995[34] |
|---|---|
| Exclusion criteria | Not provided |
| Recruitment/selection of patients | Between 1971 and 1973, 15,903 employees of the City and County of Stockholm took part in a health screening survey. Hyperparathyroidism was diagnosed in 68 of the subjects. Twenty of these underwent elective operations and the remaining 48 female patients who were free of symptoms were randomised to two treatment groups. |
| Age, gender and ethnicity | Age - Mean (SD): 58 (3). Gender (M:F): All women. Ethnicity: Not reported |
| Further population details | 1. Adjusted serum calcium: Not stated / Unclear 2. Age: Not stated / Unclear 3. Creatinine clearance: Not stated / Unclear 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): Not stated / Unclear |
| Extra comments | Female patients with moderately raised serum calcium concentrations who were free of symptoms of the disease. No details given for subgroups except that women were diagnosed with asymptomatic HPT |
| Indirectness of population | Serious indirectness: Not specified whether the participants had 'primary' HPT or other types of HPT |
| Interventions | (n=26) Intervention 1: Surgery (parathyroidectomy) - 4-gland or bilateral exploration. No detail given. Duration Single surgery. Concurrent medication/care: Not reported. Indirectness: No indirectness Comments: In each surgery case, a parathyroid adenoma was removed.<br><br>(n=22) Intervention 2: Conservative management. Non-operative conservative management. Duration Up to 3 years. Concurrent medication/care: Not reported. Indirectness: No indirectness |
| Funding | Academic or government funding (Serafimer Hospital Research Fund) |

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARATHYROIDECTOMY versus CONSERVATIVE MANAGEMENT

Protocol outcome 1: Deterioration in renal function

- Actual outcome: Narrative comment that kidney function remained within normal limits during the study period at 17 years; Group 1: 0/12, Group 2: 0/8

Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: For baseline characteristics (age, BMI, postmenopausal age), comparison was only made between the two intervention groups combined and the selected, healthy control population. The baseline characteristics between the two intervention arms were not compared.; Group 1 Number missing: 14, Reason: Oestriol taken by one patient. Other reasons not reported. ; Group 2 Number missing: 14, Reason: Eight had undergone parathyroidectomy during the follow-up (in the absence of evidence of aggregated hypercalcaemia or development of symptomatic disease). Oestriol taken by two patients. Other reasons not reported

Protocol outcome 2: Bone mineral density (BMD; distal radius or lumbar spine)

| Study | Elvius 1995[34] |
|---|---|
| - Actual outcome: Bone mineral content (described in paper as g/cm but g/cm$^2$ presumed) at 17 years; Group 1: mean 0.98 g/cm (SD 0.21); n=12, Group 2: mean 1.03 g/cm (SD 0.18); n=8<br>Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: For baseline characteristics (age, BMI, postmenopausal age), comparison was only made between the two intervention groups combined and the selected, healthy control population. The baseline characteristics between the two intervention arms were not compared.; Group 1 Number missing: 14, Reason: Oestriol taken by one patient. Other reasons not reported. ; Group 2 Number missing: 14, Reason: Eight had undergone parathyroidectomy during the follow-up (in the absence of evidence of aggregated hypercalcaemia or development of symptomatic disease). Oestriol taken by two patients. Other reasons not reported | |
| Protocol outcomes not reported by the study | Quality of life; Mortality; Fractures (vertebral or long bone); Occurrence of kidney stones; Persistent hypercalcaemia; Cardiovascular events; Adverse events (including voice change and hypoparathyroidism); Cancer |

1

| Study | Rao 2004[64] |
|---|---|
| Study type | RCT (Patient randomised; Parallel) |
| Number of studies (number of participants) | N/A (n=53) |
| Countries and setting | Conducted in USA; Setting: hospital |
| Line of therapy | 1st line |
| Duration of study | Intervention + follow up: Single surgery + Minimum of 24 months follow-up |
| Method of assessment of guideline condition | Adequate method of assessment/diagnosis: Hypercalcaemia was defined as serum Ca>10.1mg/dL or >2.52mmo/L. See inclusion criteria for more detail. |
| Stratum | Overall |
| Subgroup analysis within study | Not applicable |
| Inclusion criteria | Age 50-75 years; mean of ≥3 albumin-adjusted serum calcium levels 10.1-11.5 mg/dL (2.52-2.87 mmol/L); intact parathyroid hormone level >20pg/mL (>20ng/L); normal renal function (serum creatinine <1.5mg/dL); forearm bone mineral density within 2 S.D. adjusted for age, sex and race (Z-scores); absence of relevant symptoms and complications directly attributable to either hypercalcaemia or excess parathyroid hormone secretion; willingness to participate and ability to give informed consent for a randomised trial of parathyroidectomy; living within a 150-mile radius of the Henry Ford Hospital. |
| Exclusion criteria | Familial hyperparathyroidism; previous neck surgery or current thyroid disease requiring surgical intervention; non-traumatic vertebral/hip fractures; nephrolithiasis in past 2 years; women within 5 years of menopause; taking medications known to affect bone and mineral metabolism (e.g. glucocorticoids, |

| Study | Rao 2004[64] |
|---|---|
| | anticonvulsants, bisphosphonates); unexpected echocardiographic findings that precluded surgery |
| Recruitment/selection of patients | Patients were recruited between June 1994 and March 1997 from within the Henry Ford Health System by either physician referral or centralised laboratory computer tracking of all patients with hypercalcaemia. |
| Age, gender and ethnicity | Age - Mean (SD): Surgery = 67 (7) vs. Observation = 63 (7). Gender (M:F): 11:42. Ethnicity: Black:White = 25:28 |
| Further population details | 1. Adjusted serum calcium: <2.85mmol/L 2. Age: ≥50 years old 3. Creatinine clearance: ≥ 60 mL/min (study states serum creatinine <1.5mg/dL (<133umol/L)). 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): Absence of end-organ effects (excluded people with non-traumatic vertebral or hip fractures and nephrolithiasis. Forearm bone mineral density within 2 S.D. adjusted for age, sex and race (Z-scores)). |
| Extra comments | Patients with mild asymptomatic PHPT generally representative of the vast majority of patients with contemporary PHPT. |
| Indirectness of population | No indirectness |
| Interventions | (n=25) Intervention 1: Surgery (parathyroidectomy) - 4-gland or bilateral exploration. The surgery was performed by an experienced parathyroid surgeon, who attempted to identify 4 parathyroid glands in each patient and resected only the grossly abnormal parathyroid gland(s). No localising imaging study was performed. Duration One-off surgery. Concurrent medication/care: No detail given. Indirectness: No indirectness<br>Comments: Majority of the participants (23/25) underwent parathyroidectomy within 3 months of randomisation. One participant refused surgery after randomisation but had successful parathyroidectomy a year later, and the other participant did not have surgery in the end. At least one abnormal parathyroid gland was found in each patient.<br><br>(n=28) Intervention 2: Conservative management. No surgery. The participants were followed up every 6 months for at least 24 months. Duration Minimum of 24 months. Concurrent medication/care: No detail given. Indirectness: No indirectness<br>Comments: Ultimately, 3 of the 28 participants in the observation group had parathyroidectomy during the follow-up period because one patient developed a small kidney stone 2 years after randomisation; another patient developed pancreatitis; and a third patient developed fatigue, irritability and depression. |
| Funding | Academic or government funding (NIH Grant DK 43858) |
| RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARATHYROIDECTOMY versus OBSERVATION | |

| Study | Rao 2004[64] |
|---|---|
| Protocol outcome 1: Quality of life<br>- Actual outcome: Quality of life at Minimum of 24 months; SF-36 assessed the following nine domains: [1] physical functioning, [2] social functioning, [3] physical problem, [4] emotional problem, [5] mental health, [6] energy/fatigue, [7] pain, [8] health perception, [9] health change. In comparison with the patients who did not have surgery a statistically significant beneficial effect of parathyroidectomy was seen in 2/9 domains: social function (group difference: p=0.007) and emotional role function. A small decline was seen in 6/9 domains but only that of physical function was significant (p=0.022). In the observation group, a significant worsening occurred in 5/9 domains: social functioning, physical problem, emotional problem, energy, and health perception (p=0.013 to <0.0001). Apart from nine graphs (i.e. nine domains) charting annual changes over 36 months in the two groups and the earlier descriptive text, no other data (e.g. numerical values) were provided in relation to SF-36. ;<br>Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Baseline details: The mean age of the parathyroidectomy group was older than that of the observation group (p=0.03). All other demographic and biochemical features were similar between the two groups.; Group 1 Number missing: ; Group 2 Number missing:<br><br>Protocol outcome 2: Deterioration in renal function<br>- Actual outcome: Renal dysfunction at Minimum of 24 months; Group 1: 0/25, Group 2: 0/28<br>Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: The mean age of the parathyroidectomy group was older than that of the observation group (p=0.03). All other demographic and biochemical features were similar between the two groups.; Group 1 Number missing: ; Group 2 Number missing:<br><br>Protocol outcome 3: Fractures (vertebral or long bone)<br>- Actual outcome: Skeletal fractures (X-ray performed to assess vertebral fractures) at Minimum of 24 months; Group 1: 0/25, Group 2: 0/28<br>Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: The mean age of the parathyroidectomy group was older than that of the observation group (p=0.03). All other demographic and biochemical features were similar between the two groups.; Group 1 Number missing: ; Group 2 Number missing:<br><br>Protocol outcome 4: Occurrence of kidney stones<br>- Actual outcome: Development of kidney stones at Minimum of 24 months; Group 1: 0/25, Group 2: 1/28<br>Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: The mean age of the parathyroidectomy group was older than that of the observation group (p=0.03). All other demographic and biochemical features were similar between the two groups.; Group 1 Number missing: ; Group 2 Number missing:<br><br>Protocol outcome 5: Bone mineral density (BMD; distal radius or lumbar spine)<br>- Actual outcome: Annual change in lumbar spine BMD at Minimum of 24 months; mean values given but without measure of variance (1.2% and 0.5%, respectively). BMD increase significance: parathyroidectomy p<0.001 vs. observation p=0.087 |

| Study | Rao 2004[64] |
|---|---|
| Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Baseline details: The mean age of the parathyroidectomy group was older than that of the observation group (p=0.03). All other demographic and biochemical features were similar between the two groups.; Group 1 Number missing: ; Group 2 Number missing:<br>- Actual outcome: Annual change in forearm BMD at Minimum of 24 months; mean values given but without measure of variance (0.4% and 0.2%, respectively). BMD increase significance: parathyroidectomy p<0.001 vs. observation p=0.047<br>Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness; Baseline details: The mean age of the parathyroidectomy group was older than that of the observation group (p=0.03). All other demographic and biochemical features were similar between the two groups.; Group 1 Number missing: ; Group 2 Number missing:<br><br>Protocol outcome 6: Adverse events (including voice change and hypoparathyroidism)<br>- Actual outcome: Number of participants developing any adverse events at Minimum of 24 months; Group 1: 2/25, Group 2: 3/28; Comments: p = 0.67<br>Risk of bias: All domain - High, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Baseline details: The mean age of the parathyroidectomy group was older than that of the observation group (p=0.03). All other demographic and biochemical features were similar between the two groups.; Group 1 Number missing: ; Group 2 Number missing: | |
| Protocol outcomes not reported by the study | Mortality; Persistent hypercalcaemia ; Cardiovascular events ; Cancer |

1

| Study (subsidiary papers) | Scandinavian Investigation on Primary Hyperparathyroidism (SIPH) trial: Bollerslev 2007[13] (Lundstam 2015[51 50]) |
|---|---|
| Study type | RCT (Patient randomised; Parallel) |
| Number of studies (number of participants) | N/A (n=191) |
| Countries and setting | Conducted in Denmark, Norway, Sweden; Setting: hospital |
| Line of therapy | Mixed line |
| Duration of study | Intervention + follow up: Single surgery then follow-up at 2, 5 and 10 years (end of study) |
| Method of assessment of guideline condition | Adequate method of assessment/diagnosis: The diagnosis of PHPT was based on elevated fasting serum calcium values on 3 occasional days corrected for variation in albumin levels, and ≥2 serum measurements of intact parathyroid hormone to be above the mean of the reference interval at the local laboratory. |
| Stratum | Overall |
| Subgroup analysis within study | Not applicable |

| Study (subsidiary papers) | Scandinavian Investigation on Primary Hyperparathyroidism (SIPH) trial: Bollerslev 2007[13] (Lundstam 2015[51] [50]) |
|---|---|
| Inclusion criteria | Untreated & asymptomatic PHPT; 2.60 ≤ serum calcium ≤ 2.85mmol/L; age between 50 and 80 years; no medications interfering with calcium metabolism; informed consent |
| Exclusion criteria | Hyperparathyroid bone disease; previous neck operation; impaired kidney function (creatinine level > 130µmol/L); kidney stones; complicating medical conditions; psychiatric disorders; multiple endocrine neoplasia / familial hypocalciuric hypercalcaemia / familial hyperparathyroidism |
| Recruitment/selection of patients | The participants were recruited between 1999 and 2005 in Sweden (n=126), Norway (n=55) and Denmark (n=10). |
| Age, gender and ethnicity | Age - Mean (SD): 64.2 (7.4). Gender (M:F): 26:165. Ethnicity: Not reported |
| Further population details | 1. Adjusted serum calcium: <2.85mmol/L 2. Age: ≥50 years old 3. Creatinine clearance: Not stated / Unclear (excluded impaired kidney function (creatinine level > 130umol/l)). 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): Absence of end-organ effects |
| Extra comments | Adults with mild asymptomatic PHPT. |
| Indirectness of population | No indirectness |
| Interventions | (n=96) Intervention 1: Surgery (parathyroidectomy) - minimally invasive surgery. Parathyroidectomy by an experienced parathyroid surgeon. Duration N/A. Concurrent medication/care: In the surgery group, 14 were on oestrogens and 2 on bisphosphonates. Indirectness: No indirectness<br>Comments: Participants in the surgery group were seen 3 months after surgery for safety reasons and then once yearly. Complications of surgery (e.g. hypocalcaemia), were treated according to local traditions. In the case of unsuccessful primary operation, a secondary operation was offered according to the protocol. However, no patients were operated on more than once.<br><br>(n=95) Intervention 2: Conservative management. No details given. Duration N/A. Concurrent medication/care: In the medical observation group, 9 patients received oestrogens and 3 bisphosphonates. Indirectness: No indirectness<br>Comments: Participants in the medical observation group were seen 3 months after randomization for safety reasons and then yearly. If conservatively followed patients developed symptoms or indications for surgery or demanded surgery, they were offered surgery. By the end of the inclusion period, a total of 10 patients randomized to medical observation were surgically treated. In the statistical analyses, they were regarded as medical observation patients (Intention-to-Treat). |
| Funding | Academic or government funding (The study was supported by the Norwegian Research Council. Several of the authors had received lecture fees from industry (Amgen, Biovitrum, Novartis, Novo Nordisk, Pfizer, |

| Study (subsidiary papers) | Scandinavian Investigation on Primary Hyperparathyroidism (SIPH) trial: Bollerslev 2007[13] (Lundstam 2015[51] [50]) |
|---|---|
| | Nycomed)) |

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARATHYROIDECTOMY versus OBSERVATION

Protocol outcome 1: Quality of life
- Actual outcome: Quality of life at 1 year and 2 years; 0 - 100 Top=High is good outcome; The quality of life results based on SF-36 scores are reported as charts and not as numerical values. Statistical significance was provided for selected domains and time points only. ;
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - High, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - The quality of life results are reported as charts and specific numerical values are not given.; Indirectness of outcome: No indirectness ; Group 1 Number missing: ; Group 2 Number missing:

Protocol outcome 2: Mortality
- Actual outcome: Number of deaths in 5 years at 5 years; Group 1: 2/96, Group 2: 1/95
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - After 5 years, a total of 145 participants were still included in the protocol. For various reasons, 32 participants withdrew from the study (Surgery = 15 vs. Observation = 17), data were missing from 11 participants (Surgery = 7 vs. Observation = 4), and 3 participants died (Surgery = 2 vs. Observation = 1). Over the 5 years, 12 participants in the observation group underwent parathyroidectomy (5 due to increasing calcium levels and 7 for personal reasons). As the intention-to-treat analysis was applied, these 12 participants remained in the observation group.
; Indirectness of outcome: No indirectness ; Group 1 Number missing: 22, Reason: 15 withdrew from the study and 7 are missing; Group 2 Number missing: 21, Reason: 17 withdrew from the study and 4 are missing

Protocol outcome 3: Fractures (vertebral or long bone)
- Actual outcome: Number of new vertebral fractures in 5 years (assessed by radiograph) at 5 years; Group 1: 0/51, Group 2: 5/55; Comments: Group difference: p=0.058. 5 new vertebral fractures in 5 patients, all females in the OBS group. Four of the new vertebral fractures occurred in patients with no previous history of vertebral fractures. One of the new fractures was a progression of a fracture present already at baseline, in a vertebra containing a hemangioma, with an increase in score from 1 to 2.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - After 5 years, a total of 145 participants were still included in the protocol. For various reasons, 32 participants withdrew from the study (Surgery = 15 vs. Observation = 17), data were missing from 11 participants (Surgery = 7 vs. Observation = 4), and 3 participants died (Surgery = 2 vs. Observation = 1). Over the 5 years, 12 participants in the observation group underwent parathyroidectomy (5 due to increasing calcium levels and 7 for personal reasons). As the intention-to-treat analysis was applied, these 12 participants remained in the observation group.
; Indirectness of outcome: No indirectness ; Group 1 Number missing: 43, Reason: 15 withdrew from the study, 7 are missing, 21 did not have a follow-up X-ray; Group 2 Number missing: 39, Reason: 17 withdrew from the study, 4 are missing, 18 did not have a follow-up X-ray
- Actual outcome: Number of patients experiencing minor traumatic peripheral skeletal fractures in 5 years at 5 years; Group 1: 3/51, Group 2: 4/55
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - After 5 years, a total of 145 participants were still included in the protocol. For various reasons, 32 participants withdrew

| Study (subsidiary papers) | Scandinavian Investigation on Primary Hyperparathyroidism (SIPH) trial: Bollerslev 2007[13] (Lundstam 2015[51] [50]) |
|---|---|

from the study (Surgery = 15 vs. Observation = 17), data were missing from 11 participants (Surgery = 7 vs. Observation = 4), and 3 participants died (Surgery = 2 vs. Observation = 1). Over the 5 years, 12 participants in the observation group underwent parathyroidectomy (5 due to increasing calcium levels and 7 for personal reasons). As the intention-to-treat analysis was applied, these 12 participants remained in the observation group.
; Indirectness of outcome: No indirectness ; Group 1 Number missing: 43, Reason: 15 withdrew from the study, 7 are missing, 21 did not have a follow-up X-ray; Group 2 Number missing: 39, Reason: 17 withdrew from the study, 4 are missing, 18 did not have a follow-up X-ray

Protocol outcome 4: Occurrence of kidney stones
- Actual outcome: Number of patients developing radiological signs of new kidney stones in 5 years at 5 years; Group 1: 1/51, Group 2: 1/55; Comments: These were radiological signs of new stones in the urinary tract. No patients experienced clinical symptoms of renal calculi during the study period.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - After 5 years, a total of 145 participants were still included in the protocol. For various reasons, 32 participants withdrew from the study (Surgery = 15 vs. Observation = 17), data were missing from 11 participants (Surgery = 7 vs. Observation = 4), and 3 participants died (Surgery = 2 vs. Observation = 1). Over the 5 years, 12 participants in the observation group underwent parathyroidectomy (5 due to increasing calcium levels and 7 for personal reasons). As the intention-to-treat analysis was applied, these 12 participants remained in the observation group.
; Indirectness of outcome: No indirectness ; Group 1 Number missing: 43, Reason: 15 withdrew from the study, 7 are missing, 21 did not have a follow-up X-ray; Group 2 Number missing: 39, Reason: 17 withdrew from the study, 4 are missing, 18 did not have a follow-up X-ray

Protocol outcome 5: Bone mineral density (BMD; distal radius or lumbar spine)
- Actual outcome: Lumbar spine BMD Z-score at 5 years at 5 years; Group 1: mean 0.39 (SD 1.4); n=58, Group 2: mean -0.09 (SD 1.35); n=53; Comments: Validated DXA scans were only available for 111 participants. Difference in change between groups after 5 years: p=0.024.
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - After 5 years, a total of 145 participants were still included in the protocol. For various reasons, 32 participants withdrew from the study (Surgery = 15 vs. Observation = 17), data were missing from 11 participants (Surgery = 7 vs. Observation = 4), and 3 participants died (Surgery = 2 vs. Observation = 1). Over the 5 years, 12 participants in the observation group underwent parathyroidectomy (5 due to increasing calcium levels and 7 for personal reasons). As the intention-to-treat analysis was applied, these 12 participants remained in the observation group.
; Indirectness of outcome: No indirectness ; Group 1 Number missing: 35, Reason: 15 withdrew from the study, 7 are missing, 2 died and 14 were missing DXA scans at follow-up; Group 2 Number missing: 42, Reason: 17 withdrew from the study, 4 are missing, 1 died and 20 were missing DXA scans at follow-up
Actual outcome: Radius 33% (BMD, g/cm2) at 5 years; Group 1: mean 0.614 (SD 0.11); n=40, Group 2: mean 0.584 (SD 0.11); n=46
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Reasons for withdrawals during the inclusion period are explained, however, reasons for cases lost to follow-ups are not provided. There are discrepancies between the numbers provided in the text and those provided on the patient flow chart (Appendix 1, Supplemental Data). ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 40; Group 2 Number missing: 36
- Actual outcome: Ultra-distal radius (BMD, g/cm2) at 5 years; Group 1: mean 0.304 (SD 0.08); n=39, Group 2: mean 0.297 (SD 0.08); n=46
Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - Reasons for withdrawals during the inclusion period are explained, however, reasons for cases lost to follow-ups are not

| Study (subsidiary papers) | Scandinavian Investigation on Primary Hyperparathyroidism (SIPH) trial: Bollerslev 2007[13] (Lundstam 2015[51 50]) |
|---|---|
| | provided. There are discrepancies between the numbers provided in the text and those provided on the patient flow chart (Appendix 1, Supplemental Data). ; Indirectness of outcome: No indirectness ; Group 1 Number missing: 40; Group 2 Number missing: 36<br><br>Protocol outcome 6: Cardiovascular events<br>- Actual outcome: Number of patients with cardiovascular complications in 5 years at 5 years; Group 1: 5/72, Group 2: 8/73<br>Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - After 5 years, a total of 145 participants were still included in the protocol. For various reasons, 32 participants withdrew from the study (Surgery = 15 vs. Observation = 17), data were missing from 11 participants (Surgery = 7 vs. Observation = 4), and 3 participants died (Surgery = 2 vs. Observation = 1). Over the 5 years, 12 participants in the observation group underwent parathyroidectomy (5 due to increasing calcium levels and 7 for personal reasons). As the intention-to-treat analysis was applied, these 12 participants remained in the observation group.<br>; Indirectness of outcome: No indirectness ; Group 1 Number missing: 24, Reason: 15 withdrew from the study, 7 are missing and 2 died; Group 2 Number missing: 22, Reason: 17 withdrew from the study, 4 are missing and 1 died<br><br>Protocol outcome 7: Cancer<br>- Actual outcome: Number of patients developing malignancies in 5 years at 5 years; Group 1: 3/72, Group 2: 1/73<br>Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Very high, Outcome reporting - Low, Measurement - Low, Crossover - Low, Comments - After 5 years, a total of 145 participants were still included in the protocol. For various reasons, 32 participants withdrew from the study (Surgery = 15 vs. Observation = 17), data were missing from 11 participants (Surgery = 7 vs. Observation = 4), and 3 participants died (Surgery = 2 vs. Observation = 1). Over the 5 years, 12 participants in the observation group underwent parathyroidectomy (5 due to increasing calcium levels and 7 for personal reasons). As the intention-to-treat analysis was applied, these 12 participants remained in the observation group.<br>; Indirectness of outcome: No indirectness ; Group 1 Number missing: 24, Reason: 15 withdrew from the study, 7 are missing and 2 died; Group 2 Number missing: 22, Reason: 17 withdrew from the study, 4 are missing and 1 died |
| Protocol outcomes not reported by the study | Deterioration in renal function ; Persistent hypercalcaemia; Adverse events |

1

| Study | Talpos 2000[83] |
|---|---|
| Study type | RCT (Patient randomised; Parallel) |
| Number of studies (number of participants) | N/A (n=53) |
| Countries and setting | Conducted in USA; Setting: Secondary care |
| Line of therapy | 1st line |
| Duration of study | Intervention + follow up: Single surgery + Up to 2 years of follow-up |

| Study | Talpos 2000[83] |
|---|---|
| Method of assessment of guideline condition | Adequate method of assessment/diagnosis: See inclusion criteria |
| Stratum | Overall |
| Subgroup analysis within study | Stratified then randomised |
| Inclusion criteria | Age 50 - 75 years; persistent albumin-adjusted serum calcium level 10.1 - 11.5 mg/dL (2.52 - 2.87mmol/L) (normal level < 10.1mg/dL) from at least 3 measurements over a period of at least 3 months; intact parathyroid hormone level > 20pg/mL; no other cause for hypercalcaemia; women at least 5 years after menopause; willingness to participate and ability to give consent to a RCT; living within 150-mile radius of downtown Detroit; not currently enrolled in any other clinical trial. |
| Exclusion criteria | Polyuria/Polydipsia/Anorexia/Nausea/Vomiting; pancreatitis in the past 1 year; symptomatic peptic ulcer disease; objective muscle weakness; history of non-traumatic vertebral/hip fractures; nephrolithiasis in the past 2 years; history of glucocorticoid/anticonvulsant drug therapy; thiazide diuretic therapy for hypertension cannot be changed; family history of PHPT / multiple endocrine neoplasia / benign hypocalciuric hypercalcaemia; evidence of thyroid disease requiring surgery; history of childhood irradiation to head/neck; presence of any of the following abnormalities (mean of 3 corrected serum calcium > 11.5mg/dL, mean of 3 serum creatinine determinations > 1.5mg/dL, creatinine clearance level < 70%, forearm BMD >2 SD below the expected value, phalangeal sub periosteal resorption on hand radiographs, vertebral compression fractures, urolithiasis on kidneys/ureter/bladder, unexpected findings on echocardiogram that preclude surgery) |
| Recruitment/selection of patients | All patients who were referred to the Division of Bone and Mineral Metabolism or the Department of Surgery between April 1994 and March 1997, who met the criteria were invited to participate in the study. |
| Age, gender and ethnicity | Age - Other: Mean age for operative group = 66.7 vs. observation group = 62.6; p<0.03. Gender (M: F): 11:42. Ethnicity: White = 28; Black = 25 |
| Further population details | 1. Adjusted serum calcium: <2.85mmol/L 2. Age: ≥50 years old 3. Creatinine clearance: ≥ 60 mL/min (study reports an exclusion criteria of having a creatinine clearance level < 70%). 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): Absence of end-organ effects (Exclusion criteria were forearm BMD >2 SD below the expected value, vertebral compression fractures, urolithiasis on kidneys/ureter/bladder, history of non-traumatic vertebral/hip fractures; nephrolithiasis in the past 2 years.). |
| Extra comments | Asymptomatic patients with confirmed PHPT. |
| Indirectness of population | No indirectness |
| Interventions | (n=25) Intervention 1: Surgery (parathyroidectomy) - 4-gland or bilateral exploration. All patients randomised to surgery underwent standard parathyroidectomy with a bilateral approach by a single experienced surgeon who had performed >600 parathyroid procedures before the start of the study. Duration Single surgery. |

| Study | Talpos 2000[83] |
|---|---|
| | Concurrent medication/care: Routine postoperative care was provided which included frequent calcium determinations during the average 2-day hospitalisation. Calcium carbonate and magnesium supplements were administered as needed before and after discharge. Indirectness: No indirectness<br><br>(n=28) Intervention 2: Conservative management. No detail given. Duration Up to 2 years. Concurrent medication/care: No detail given. Indirectness: No indirectness |
| Funding | Academic or government funding (National Institutes of Health grant) |

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARATHYROIDECTOMY versus OBSERVATION

Protocol outcome 1: Quality of life

- Actual outcome: Annual change estimate for SF-36 physical functioning at 2 years; MD; -2.103 (SE: 1.70), Comments: SE calculated from P value of the mean difference);

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0

- Actual outcome: Annual change estimate for SF-36 social functioning at 2 years; MD; 3.918 (SE: 1.39), Comments: SE calculated from P value of the mean difference);

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0

- Actual outcome: Annual change estimate for SF-36 physical role functioning at 2 years; MD; 0.392 (SE: 3.17), Comments: SE calculated from P value of the mean difference);

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0

- Actual outcome: Annual change estimate for SF-36 emotional role functioning at 2 years; MD; 5.955 (SE: 2.29), Comments: SE calculated from P value of the mean difference);

Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; G Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0

- Actual outcome: Annual change estimate for SF-36 mental health at 2 years; MD; 0.225 (SE: 0.92), Comments: SE calculated from P value of the mean difference);

| Study | Talpos 2000[83] |
|---|---|
| | Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0<br>- Actual outcome: Annual change estimate for SF-36 vitality at 2 years; MD; 0.970 (SE: 1.10), Comments: SE calculated from P value of the mean difference);<br>Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0<br>- Actual outcome: Annual change estimate for SF-36 bodily pain at 2 years; MD; 0.649 (SE: 1.63), Comments: SE calculated from P value of the mean difference);<br>Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0<br>- Actual outcome: Annual change estimate for SF-36 general health at 2 years; MD; 1.815 (SE: 1.12), Comments: SE calculated from P value of the mean difference);<br>Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0<br>- Actual outcome: Annual change estimate for SF-36 health transition at 2 years; MD; 0.116 (SE: 1.64), Comments: SE calculated from P value of the mean difference);<br>Risk of bias: All domain - Very high, Selection - High, Blinding - High, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Low, Crossover - Low; Indirectness of outcome: No indirectness ; Group 1 Number missing: 2, Reason: 2 people refused surgery but were analysed in the intervention group on an ITT basis; Group 2 Number missing: 0 |
| Protocol outcomes not reported by the study | Mortality; Deterioration in renal function; Fractures (vertebral or long bone); Occurrence of kidney stones; Persistent hypercalcaemia; Bone mineral density (BMD; distal radius or lumbar spine); Cardiovascular events; Adverse events (including voice change and hypoparathyroidism); Cancer |

1

| Study (subsidiary papers) | Vanderwalde 2006[87] (Vanderwalde 2009[88]) |
|---|---|
| Study type | Non-randomised comparative study |
| Number of studies (number of participants) | 1 (n=1569) |
| Countries and setting | Conducted in USA; Setting: Hospital |
| Line of therapy | 1st line |

| Study (subsidiary papers) | Vanderwalde 2006[87] (Vanderwalde 2009[88]) |
|---|---|
| Duration of study | Follow up (post intervention): Retrospective cohort study with a follow-up of 7.4 years (range: 13 days to 10 years). |
| Method of assessment of guideline condition | Adequate method of assessment/diagnosis: People on the database defined as having PHPT if they had an intact parathyroid hormone (PTH) level greater than 65 pg/mL, a calcium level greater than 10.5 mg/dL (>2.6 mmol/L), and a creatinine level less than 2.5 mg/dL (<221.0 µmol/L). Excluded patients likely to have tertiary HPT or with a history of chronic renal failure requiring dialysis (see exclusion criteria). |
| Stratum | Overall |
| Subgroup analysis within study | Not applicable |
| Inclusion criteria | People with an intact parathyroid hormone (PTH) level greater than 65 pg/mL, a calcium level greater than 10.5 mg/dL (>2.6 mmol/L), and a creatinine level less than 2.5 mg/dL (<221.0 µmol/L) |
| Exclusion criteria | <20 years old. To ensure that no patient was included who had tertiary HPT, any patient who had at least 2 separate blood samples drawn for measurement of cyclosporine (laboratory procedure code 8718671), tacrolimus (FK 506;laboratory procedure code 8203004), or sirolimus (laboratory procedure code 8718652) levels was considered to be a probable kidney transplant recipient and excluded. A second database, the Southern California Kaiser Permanente Discharge Abstract Database, was used to exclude patients with any history of chronic renal failure requiring dialysis (International Classification of Diseases, Ninth Revision [ICD-9] code 585.6). |
| Recruitment/selection of patients | Retrospective cohort study. Screened the Southern California Kaiser Permanente Laboratory Management System database to identify all southern California Kaiser Permanente members eligible for inclusion between January 1, 1995, and December 31, 2000. |
| Age, gender and ethnicity | Age - Other: Age ≥50 years: parathyroidectomy 138 (87%); conservative management 334 (89%). Gender (M:F): 72/461. Ethnicity: |
| Further population details | 1. Adjusted serum calcium: Not stated / Unclear 2. Age: ≥50 years old (89% ≥ 50 years old). 3. Creatinine clearance: Not stated / Unclear 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): Not stated / Unclear (22% had osteoporosis at baseline; kidney stones or history of fragility fractures not reported). |
| Extra comments | . 2006 paper is the primary study reporting the overall cohort of 1569 people. 2009 paper reports data for N=533 who had BMD data available (hazard ratio also adjusted for BMD). |
| Indirectness of population | No indirectness |
| Interventions | (n=159) Intervention 1: Surgery (parathyroidectomy) - 4-gland or bilateral exploration: not reported. Duration average follow-up of 7.4 years (range:<br>13 days to 10 years). Concurrent medication/care: not reported. Indirectness: No indirectness |

| Study (subsidiary papers) | Vanderwalde 2006[87] (Vanderwalde 2009[88]) |
|---|---|
| | (n=374) Intervention 2: Conservative management. Duration average follow-up of 7.4 years (range: 13 days to 10 years). Concurrent medication/care: not reported. Indirectness: No indirectness |
| Funding | Funding not stated |
| RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PARATHYROIDECTOMY versus CONSERVATIVE MANAGEMENT<br><br>Protocol outcome 1: Fractures (vertebral or long bone)<br>- Actual outcome: Hospitalised fracture at average follow-up of 7.4 years; Group 1: n=159 ; Group 2: n=374; HR 0.41; Lower CI 0.18 to Upper CI 0.93; Comments: Multivariate analysis confirmed that parathyroidectomy was independently associated with a decreased fracture risk (HR = 0.41; 95% CI 0.18,0.93; p = 0.03) after accounting for all other variables (age, sex, Charlson comorbidity index (CCI); levels of calcium, PTH, and creatinine; BMD (femurT-score).<br>Risk of bias: All domain - Very high, Selection - High, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - High, Crossover - Low, Comments - Outcome of fracture taken from records of hospitalised fractures (so would not pick up all vertebral fractures on radiograph or outpatient fractures of the extremities); Indirectness of outcome: No indirectness ; Baseline details: Patients who were treated operatively were similar with regard to age, gender, and race, but were more likely to have higher calcium (p= 0.001) and PTH levels (p = 0.001) than patients who were observed. Furthermore, those who were observed were more likely to have osteoporosis (p =0.018); Key confounders: Age, sex, Charlson comorbidity index (CCI); levels of calcium, PTH, and creatinine; BMD (T score femur); Group 1 Number missing: 0; Group 2 Number missing: 0 | |
| Protocol outcomes not reported by the study | Quality of life; Mortality; Deterioration in renal function; Occurrence of kidney stones; Persistent hypercalcaemia; Bone mineral density (BMD; distal radius or lumbar spine); Cardiovascular events; Adverse events (including voice change and hypoparathyroidism); Cancer |

1

| Study (subsidiary papers) | Vestergaard 2003[90] |
|---|---|
| Study type | Prospective cohort study |
| Number of studies (number of participants) | 1 (n=3213) |
| Countries and setting | Conducted in Denmark; Setting: Nationwide Danish cohort. |
| Line of therapy | 1st line |
| Duration of study | Intervention + follow up: Data collected from 1 January 1980 to 31 December 1999. 6.1 years (median follow up after diagnosis) |
| Method of assessment of guideline condition | Adequate method of assessment/diagnosis |

| Study (subsidiary papers) | Vestergaard 2003[90] |
|---|---|
| Stratum | Overall |
| Subgroup analysis within study | Not applicable |
| Inclusion criteria | Patients with a first time diagnosis of primary hyperparathyroidism for the period 1 January 1980 to 31 December 1999. |
| Exclusion criteria | not stated |
| Recruitment/selection of patients | Patients were identified through the Danish National Hospital Discharge Register, which is a nationwide computer-based register of all contacts to Danish hospitals |
| Age, gender and ethnicity | Age - Mean (SD): surgery - 58.3 (15.2) ; no surgery 64.2 (17.4). Gender (M:F): Men-surgery- 500 (26%); no surgery- 293 (23%) ; Women - surgery 1434 (74%) ; no surgery -986 (77%). Ethnicity: not stated |
| Further population details | 1. Adjusted serum calcium: Not stated / Unclear 2. Age: Not stated / Unclear 3. Creatinine clearance: Not stated / Unclear 4. Presence of end-organ effects (end organ effects defined as kidney stones, history of fragility fractures or osteoporosis (BMD T-score <-2.5 at any site)): Not stated / Unclear |
| Extra comments | -- |
| Indirectness of population | No indirectness |
| Interventions | (n=1934) Intervention 1: Surgery (parathyroidectomy) - 4-gland or bilateral exploration. Median time to surgery was 31 days from diagnosis (range 0-14 years). Duration 6.1 years (median follow up after diagnosis). Concurrent medication/care: No further details. Indirectness: No indirectness<br><br>(n=1279) Intervention 2: Conservative management. Conservative management, no further details. Duration 6.1 years (median follow up after diagnosis). Concurrent medication/care: No details. Indirectness: No indirectness |
| Funding | No funding |

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: SURGERY versus CONSERVATIVE MANAGEMENT

Protocol outcome 1: Mortality
- Actual outcome: Mortality at 6.1 years (estimated); ; Group 1: n=1934 ; Group 2: n=1279; HR 0.65; Lower CI 0.57 to Upper CI 0.93
Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Very high, Crossover - Low, Subgroups - Low; Indirectness of outcome: No indirectness ; Baseline details: Matched for age and gender; Key confounders: Only adjusted for age key confounder; Group 1 Number missing: ; Group 2 Number missing:

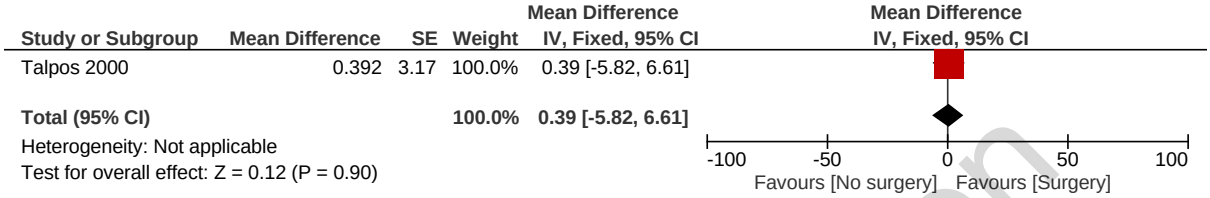
Protocol outcome 2: Kidney stones at 6.1 years (estimated); ; Group 1: n=1934 ; Group 2: n=1279; HR 1.87; Lower CI 1.3 to Upper CI 2.69
- Actual outcome: Kidney stones at 6.1 years (estimated);

| Study (subsidiary papers) | Vestergaard 2003[90] |
|---|---|
| Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Very high, Crossover - Low, Subgroups - Low; Indirectness of outcome: No indirectness ; Baseline details: Matched for age and gender; Key confounders: Only adjusted for age key confounder; Group 1 Number missing: ; Group 2 Number missing:<br><br>Protocol outcome 3: Fractures (vertebral or long bone) at 6.1 years (estimated); ; Group 1: n=1934 ; Group 2: n=1279; HR 0.69; Lower CI 0.56 to Upper CI 0.82<br>- Actual outcome: Fractures at 6.1 years (estimated);<br>Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Very high, Crossover - Low, Subgroups - Low; Indirectness of outcome: No indirectness ; Baseline details: Matched for age and gender; Key confounders: Only adjusted for age key confounder; Group 1 Number missing: ; Group 2 Number missing:<br><br>Protocol outcome 4: Cancer at 6.1 years (estimated)<br>- Actual outcome: Cancer at 6.1 years (estimated); ; Group 1: n=1934 ; Group 2: n=1279; HR 1.11; Lower CI 0.9 to Upper CI 1.37<br>Risk of bias: All domain - Very high, Selection - Very high, Blinding - Low, Incomplete outcome data - Low, Outcome reporting - Low, Measurement - Very high, Crossover - Low, Subgroups - Low; Indirectness of outcome: No indirectness ; Baseline details: Matched for age and gender; Key confounders: Only adjusted for age key confounder; Group 1 Number missing: ; Group 2 Number missing: | |
| Protocol outcomes not reported by the study | Quality of life; Occurrence of kidney stones; Persistent hypercalcaemia; Bone mineral density (BMD; distal radius or lumbar spine); Cardiovascular events; Adverse events (including voice change and hypoparathyroidism) |

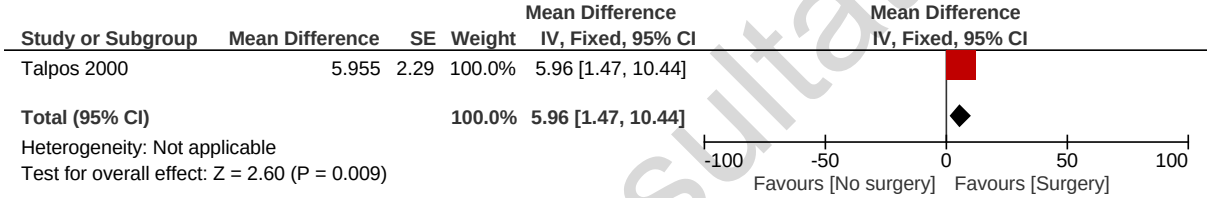
# Appendix E: Forest plots

## E.1 Surgery versus conservative management

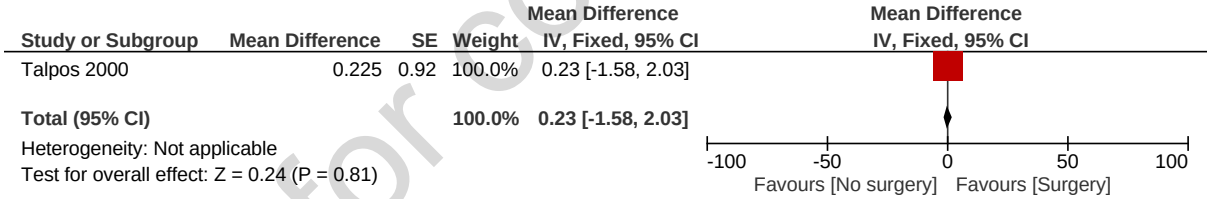
### Figure 2: QOL (SF-36 Physical role functioning subscale)



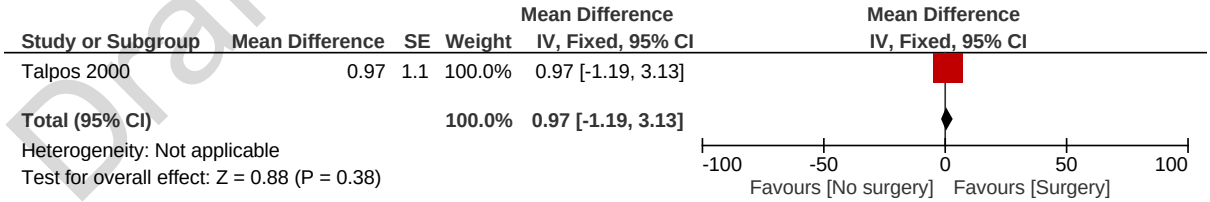

### Figure 3: QOL (Emotional role functioning subscale)



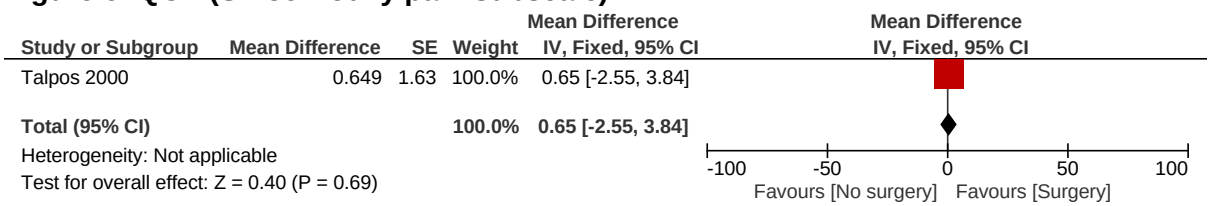

### Figure 4: QOL (SF-36 mental health subscale)




### Figure 5: QOL (SF-36 vitality subscale)



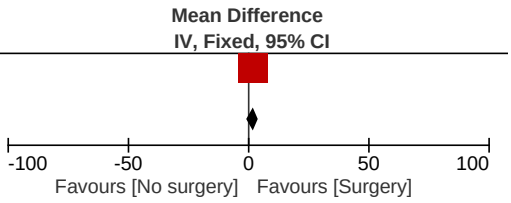

### Figure 6: QOL (SF-36 Bodily pain subscale)



## Figure 7: QOL (SF-36 General health subscale)

| Study or Subgroup | Mean Difference | SE | Weight | Mean Difference IV, Fixed, 95% CI |
|---|---|---|---|---|
| Talpos 2000 | 1.815 | 1.12 | 100.0% | 1.81 [-0.38, 4.01] |
| **Total (95% CI)** | | | **100.0%** | **1.81 [-0.38, 4.01]** |

Heterogeneity: Not applicable
Test for overall effect: Z = 1.62 (P = 0.11)

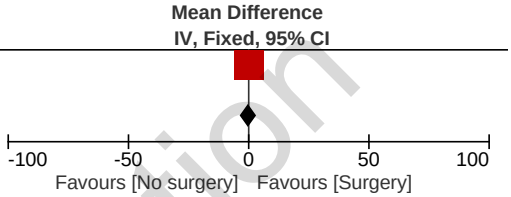



1

## Figure 8: QOL (SF-36 Health transition subscale)

| Study or Subgroup | Mean Difference | SE | Weight | Mean Difference IV, Fixed, 95% CI |
|---|---|---|---|---|
| Talpos 2000 | 0.116 | 1.64 | 100.0% | 0.12 [-3.10, 3.33] |
| **Total (95% CI)** | | | **100.0%** | **0.12 [-3.10, 3.33]** |

Heterogeneity: Not applicable
Test for overall effect: Z = 0.07 (P = 0.94)

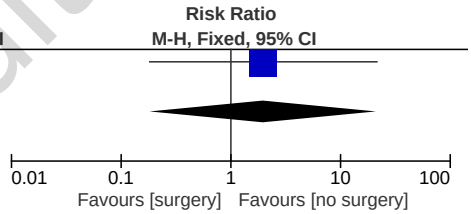



2

## Figure 9: Mortality

| Study or Subgroup | Surgery Events | Surgery Total | No surgery Events | No surgery Total | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|---|
| Bollerslev 2007 | 2 | 96 | 1 | 95 | 100.0% | 1.98 [0.18, 21.46] |
| **Total (95% CI)** | | **96** | | **95** | **100.0%** | **1.98 [0.18, 21.46]** |
| Total events | 2 | | 1 | | | |

Heterogeneity: Not applicable
Test for overall effect: Z = 0.56 (P = 0.57)

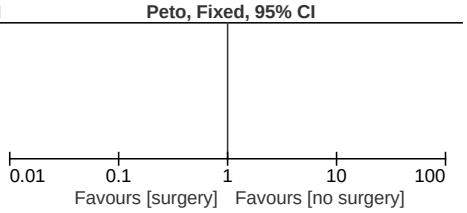



3

## Figure 10: Renal dysfunction

| Study or Subgroup | Surgery Events | Surgery Total | No surgery Events | No surgery Total | Weight | Peto Odds Ratio Peto, Fixed, 95% CI |
|---|---|---|---|---|---|---|
| Elvius 1995 | 0 | 12 | 0 | 8 | | Not estimable |
| Rao 2004 | 0 | 25 | 0 | 28 | | Not estimable |
| **Total (95% CI)** | | **37** | | **36** | | **Not estimable** |
| Total events | 0 | | 0 | | | |

Heterogeneity: Not applicable
Test for overall effect: Not applicable

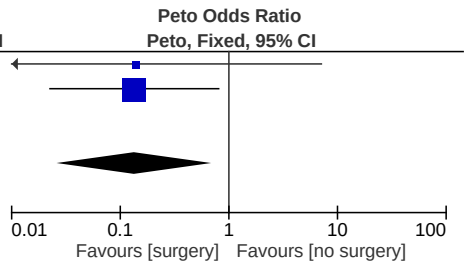



4

## Figure 11: Vertebral fractures

| Study or Subgroup | Surgery Events | Surgery Total | No surgery Events | No surgery Total | Weight | Peto Odds Ratio Peto, Fixed, 95% CI |
|---|---|---|---|---|---|---|
| Ambrogini 2007 | 0 | 24 | 1 | 25 | 17.2% | 0.14 [0.00, 7.10] |
| Bollerslev 2007 | 0 | 51 | 5 | 55 | 82.8% | 0.13 [0.02, 0.81] |
| Rao 2004 | 0 | 25 | 0 | 28 | | Not estimable |
| **Total (95% CI)** | | **100** | | **108** | **100.0%** | **0.14 [0.03, 0.69]** |
| Total events | 0 | | 6 | | | |

Heterogeneity: Chi² = 0.00, df = 1 (P = 0.98); I² = 0%
Test for overall effect: Z = 2.40 (P = 0.02)

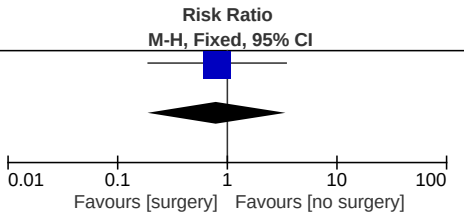



5

## Figure 12: Peripheral skeletal fractures

| Study or Subgroup | Surgery Events | Surgery Total | No surgery Events | No surgery Total | Weight | Risk Ratio M-H, Fixed, 95% CI |
|---|---|---|---|---|---|---|
| Bollerslev 2007 | 3 | 51 | 4 | 55 | 100.0% | 0.81 [0.19, 3.44] |
| **Total (95% CI)** | | **51** | | **55** | **100.0%** | **0.81 [0.19, 3.44]** |
| Total events | 3 | | 4 | | | |

Heterogeneity: Not applicable
Test for overall effect: Z = 0.29 (P = 0.77)

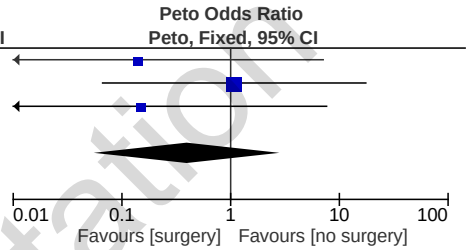



1

## Figure 13: Kidney stones

| Study or Subgroup | Surgery Events | Surgery Total | No surgery Events | No surgery Total | Weight | Peto Odds Ratio Peto, Fixed, 95% CI |
|---|---|---|---|---|---|---|
| Ambrogini 2007 | 0 | 24 | 1 | 25 | 25.1% | 0.14 [0.00, 7.10] |
| Bollerslev 2007 | 1 | 51 | 1 | 55 | 49.8% | 1.08 [0.07, 17.52] |
| Rao 2004 | 0 | 25 | 1 | 28 | 25.1% | 0.15 [0.00, 7.64] |
| **Total (95% CI)** | | **100** | | **108** | **100.0%** | **0.39 [0.06, 2.82]** |
| Total events | 1 | | 3 | | | |

Heterogeneity: Chi² = 1.00, df = 2 (P = 0.61); I² = 0%
Test for overall effect: Z = 0.93 (P = 0.35)

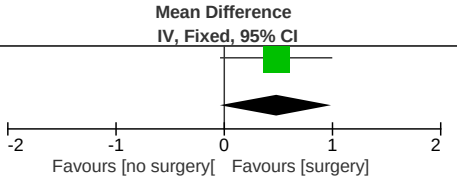



2
3

## Figure 14: Lumbar spine BMD (Z score)

| Study or Subgroup | Surgery Mean | Surgery SD | Surgery Total | No surgery Mean | No surgery SD | No surgery Total | Weight | Mean Difference IV, Fixed, 95% CI |
|---|---|---|---|---|---|---|---|---|
| Bollerslev 2007 | 0.39 | 1.4 | 58 | -0.09 | 1.35 | 53 | 100.0% | 0.48 [-0.03, 0.99] |
| **Total (95% CI)** | | | **58** | | | **53** | **100.0%** | **0.48 [-0.03, 0.99]** |

Heterogeneity: Not applicable
Test for overall effect: Z = 1.84 (P = 0.07)

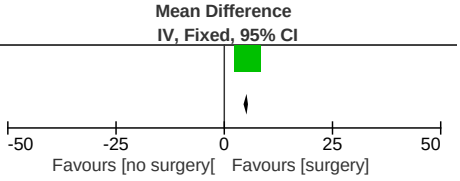



4

## Figure 15: Lumbar spine BMD (% change from baseline)

| Study or Subgroup | Surgery Mean | Surgery SD | Surgery Total | No surgery Mean | No surgery SD | No surgery Total | Weight | Mean Difference IV, Fixed, 95% CI |
|---|---|---|---|---|---|---|---|---|
| Ambrogini 2007 | 4.16 | 1.1 | 24 | -1.12 | 0.71 | 25 | 100.0% | 5.28 [4.76, 5.80] |
| **Total (95% CI)** | | | **24** | | | **25** | **100.0%** | **5.28 [4.76, 5.80]** |

Heterogeneity: Not applicable
Test for overall effect: Z = 19.87 (P < 0.00001)

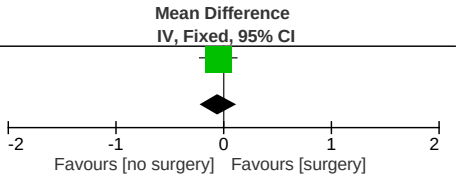



5

## Figure 16: Distal radius BMD (g/cm$^2$)

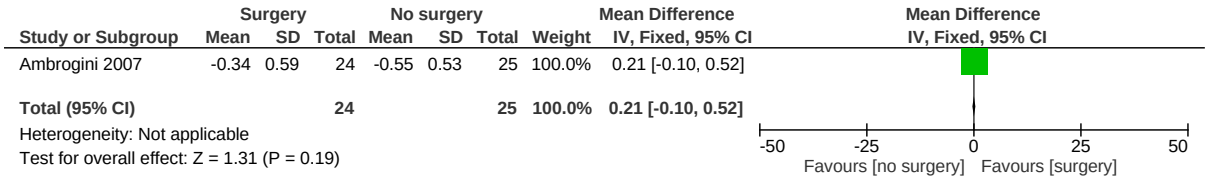
| Study or Subgroup | Surgery Mean | Surgery SD | Surgery Total | No surgery Mean | No surgery SD | No surgery Total | Weight | Mean Difference IV, Fixed, 95% CI |
|---|---|---|---|---|---|---|---|---|
| Elvius 1995 | 0.98 | 0.21 | 12 | 1.03 | 0.18 | 8 | 100.0% | -0.05 [-0.22, 0.12] |
| **Total (95% CI)** | | | **12** | | | **8** | **100.0%** | **-0.05 [-0.22, 0.12]** |

Heterogeneity: Not applicable
Test for overall effect: Z = 0.57 (P = 0.57)



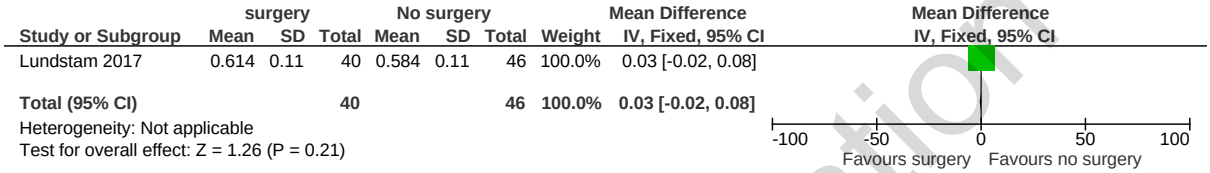

6

## Figure 17: Distal radius BMD (% change from baseline)

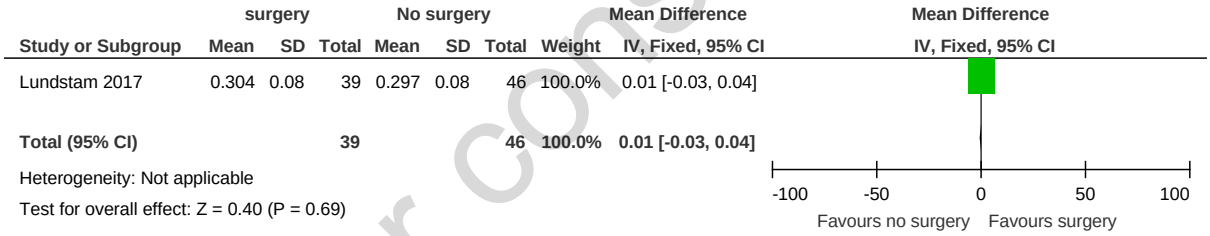



1

## Figure 18: Radius 33% (BMD, g/cm$^2$) (5 years)




## Figure 19: Ultradistal radius (BMD, g/cm$^2$) (5 years)

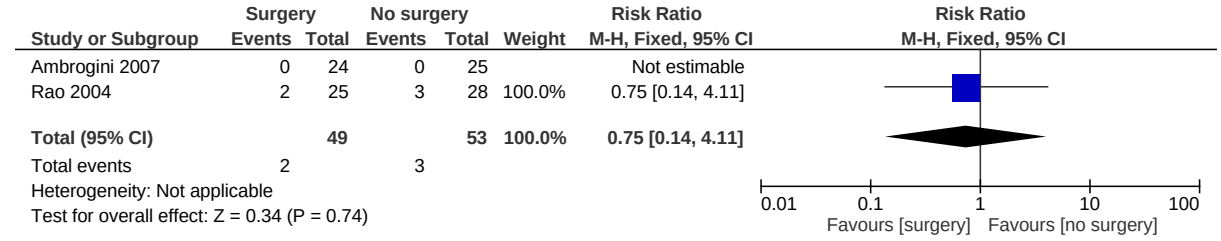



2
3
4

## Figure 20: Cardiovascular events

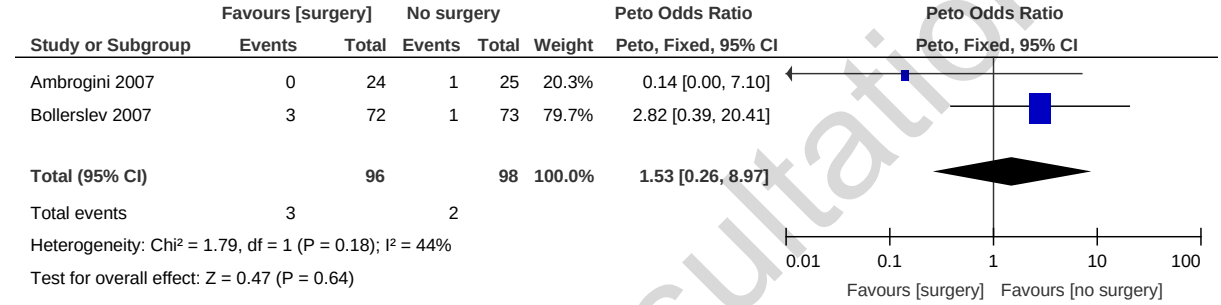

5

**Figure 21: Adverse events**




1

**Figure 22: Cancer**




2

3

4

# E.25 Surgery versus conservative treatment (non-randomised)

6

**Figure 23: Mortality (median follow-up 6.1 years)**

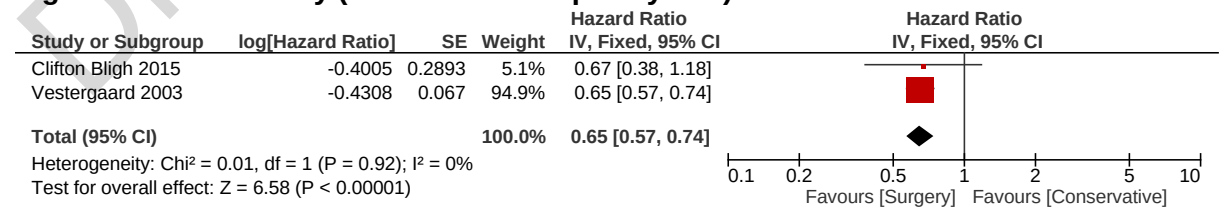


**Figure 24: Fractures (median follow up from diagnosis 6.1 to 7.4 years)**

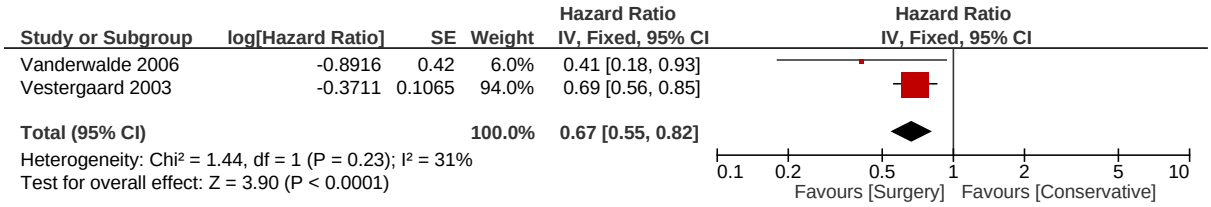



**Figure 25: Cancer (median follow up from diagnosis 6.1 years)**

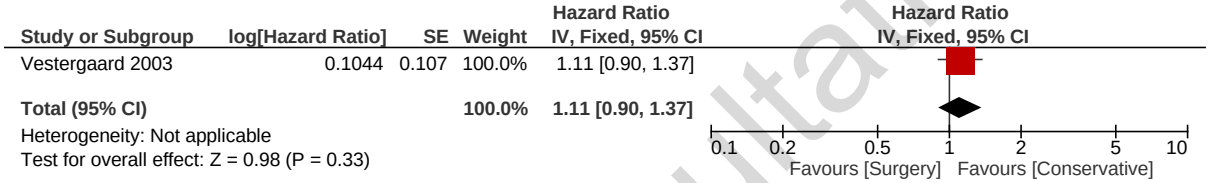
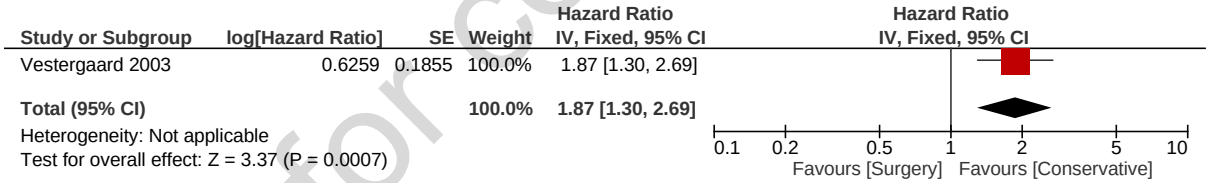



**Figure 26: Kidney stones (median follow up from diagnosis 6.1 years)**



# Appendix F: GRADE tables

**Table 11: Clinical evidence profile: Surgery versus conservative management**

| Quality assessment | | | | | | | No of patients | | Effect | | Quality | Importance |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| No of studies | Design | Risk of bias | Inconsistency | Indirectness | Imprecision | Other considerations | Surgery | No surgery (in mild PHPT) | Relative (95% CI) | Absolute | | |
| **Quality of life (SF-36 Physical functioning subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 25 | 28 | - | MD 2.1 lower (5.43 lower to 1.23 higher) | VERY LOW | CRITICAL |
| **Quality of life (SF-36 Social functioning subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 25 | 28 | - | MD 3.92 higher (1.19 to 6.64 higher) | VERY LOW | CRITICAL |
| **Quality of life (SF-36 Physical role functioning subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 25 | 28 | - | MD 0.39 higher (5.82 lower to 6.61 higher) | VERY LOW | CRITICAL |
| **Quality of life (SF-36 Emotional role functioning subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 25 | 28 | - | MD 5.96 higher (1.47 to 10.44 higher) | VERY LOW | CRITICAL |
| **Quality of life (SF-36 mental health subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | no serious imprecision | none | 25 | 25 | - | MD 0.23 higher (1.58 lower to 2.03 higher) | LOW | CRITICAL |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Quality of life (SF-36 vitality subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 25 | 28 | - | MD 0.97 higher (1.19 lower to 3.13 higher) | VERY LOW | CRITICAL |
| **Quality of life (SF-36 Bodily pain subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 25 | 28 | - | MD 0.65 higher (2.55 lower to 3.84 higher) | VERY LOW | CRITICAL |
| **Quality of life (SF-36 General health subscale) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 25 | 28 | - | MD 1.81 higher (0.38 lower to 4.01 higher) | VERY LOW | CRITICAL |
| **Quality of life (SF-36 Health transition) (follow-up 2 years; measured with: annual change estimate; range of scores: 0-100; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | very serious[b,c] | none | 25 | 28 | - | MD 0.12 higher (3.1 lower to 3.33 higher) | VERY LOW | CRITICAL |
| **Mortality (follow-up 5 years)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 2/96 (2.1%) | 1.1% | RR 1.98 (0.18 to 21.46) | 11 more per 1000 (from 9 fewer to 225 more) | VERY LOW | CRITICAL |
| **Renal Dysfunction (follow-up 2-17 years)** | | | | | | | | | | | | |
| 2 | randomised trials | Serious[a] | no serious inconsistency | no serious indirectness | serious imprecision[f] | none | 0/37 (0%) | 0% | - | 0 more per 1000 (from 180 fewer to 180 more)[d] | LOW | CRITICAL |
| **Vertebral fractures (follow-up 1-5 years)** | | | | | | | | | | | | |
| 3 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | no serious imprecision | none | 0/100 (0%) | 4% | OR 0.14 (0.03 to 0.69) | 60 fewer per 1000 (from 110 fewer to 0 more)[d] | LOW | CRITICAL |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Peripheral skeletal fractures (follow-up 5 years)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 3/51 (5.9%) | 7.3% | RR 0.81 (0.19 to 3.44) | 14 fewer per 1000 (from 59 fewer to 178 more) | VERY LOW | CRITICAL |
| **Kidney Stones (follow-up 1-5 years)** | | | | | | | | | | | | |
| 3 | randomised trials | Very serious[a] | no serious inconsistency | no serious indirectness | serious[b] | none | 1/100 (1%) | 3.6% | Peto OR 0.39 (0.06 to 2.82) | 20 fewer per 1000 (from 60 fewer to 30 more) | VERY LOW | CRITICAL |
| **Lumbar spine BMD (follow-up 5 years; measured with: Z score (final value); Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 58 | 53 | - | MD 0.48 higher (0.03 lower to 0.99 higher) | VERY LOW | CRITICAL |
| **Lumbar spine BMD (follow-up 1 years; measured with: % change from baseline; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | Serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 24 | 25 | - | MD 5.28 higher (4.76 to 5.8 higher) | VERY LOW | CRITICAL |
| **Distal radius BMD (follow-up 17 years; measured with: g/cm2; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 12 | 18 | - | MD 0.05 lower (0.22 lower to 0.12 higher) | VERY LOW | CRITICAL |
| **Distal radius BMD (follow-up 1 years; measured with: % change from baseline; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | Serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | 24 | 25 | - | MD 0.21 higher (0.1 lower to 0.52 higher) | LOW | CRITICAL |
| **Radius 33% (BMD, g/cm2) (follow-up 5 years; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised | very | no serious | no serious | Serious[b] | none | 40 | 46 | - | MD 0.03 higher (0.02 | VERY | CRITICAL |

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | trials | serious[a] | inconsistency | indirectness | | | | | | lower to 0.08 higher) | LOW | |
| **Ultradistal radius (BMD, g/cm2) (follow-up 5 years; Better indicated by higher values)** | | | | | | | | | | | | |
| 1 | randomised trials | Serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 39 | 46 | - | MD 0.01 higher (0.03 lower to 0.04 higher) | LOW | CRITICAL |
| **Cardiovascular events (follow-up 5 years)** | | | | | | | | | | | | |
| 1 | randomised trials | very serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 5/72 (6.9%) | 11% | RR 0.63 (0.22 to 1.85) | 41 fewer per 1000 (from 86 fewer to 94 more) | VERY LOW | IMPORTANT |
| **Adverse events (follow-up 1-2 years)** | | | | | | | | | | | | |
| 2 | randomised trials | Serious[a] | no serious inconsistency | no serious indirectness | very serious[b] | none | 2/49 (4.1%) | 5.4% | RR 0.75 (0.14 to 4.11) | 14 fewer per 1000 (from 46 fewer to 168 more) | VERY LOW | IMPORTANT |
| **Cancer (follow-up 1-5 years)** | | | | | | | | | | | | |
| 2 | randomised trials | serious[1] | No serious inconsistency[e] | no serious indirectness | very serious[b] | none | 3/96 (3.1%) | 2.7% | Peto OR 1.53 (0.26 to 8.97) | 10 more per 1000 (from 40 fewer to 60 more) | VERY LOW | IMPORTANT |

*a Downgraded by 1 increment if the majority of the evidence was at high risk of bias, and downgraded by 2 increments if the majority of the evidence was at very high risk of bias.*
*b Downgraded by 1 increment if the confidence interval crossed 1 MID, and downgraded by 2 increments if the confidence interval crossed both MIDs.*
*c Established MID not available for this domain of the SF-36, therefore default MID used*
*d Manual calculation of absolute risk difference*
*e inconsistency is not applicable due to zero events in one arm of one study*
*f Downgraded by 1 increment as both studies had 0 events in both arms and sample size was >70<350*

**Table 12: Clinical evidence profile: Surgery versus conservative treatment (non-randomised)**

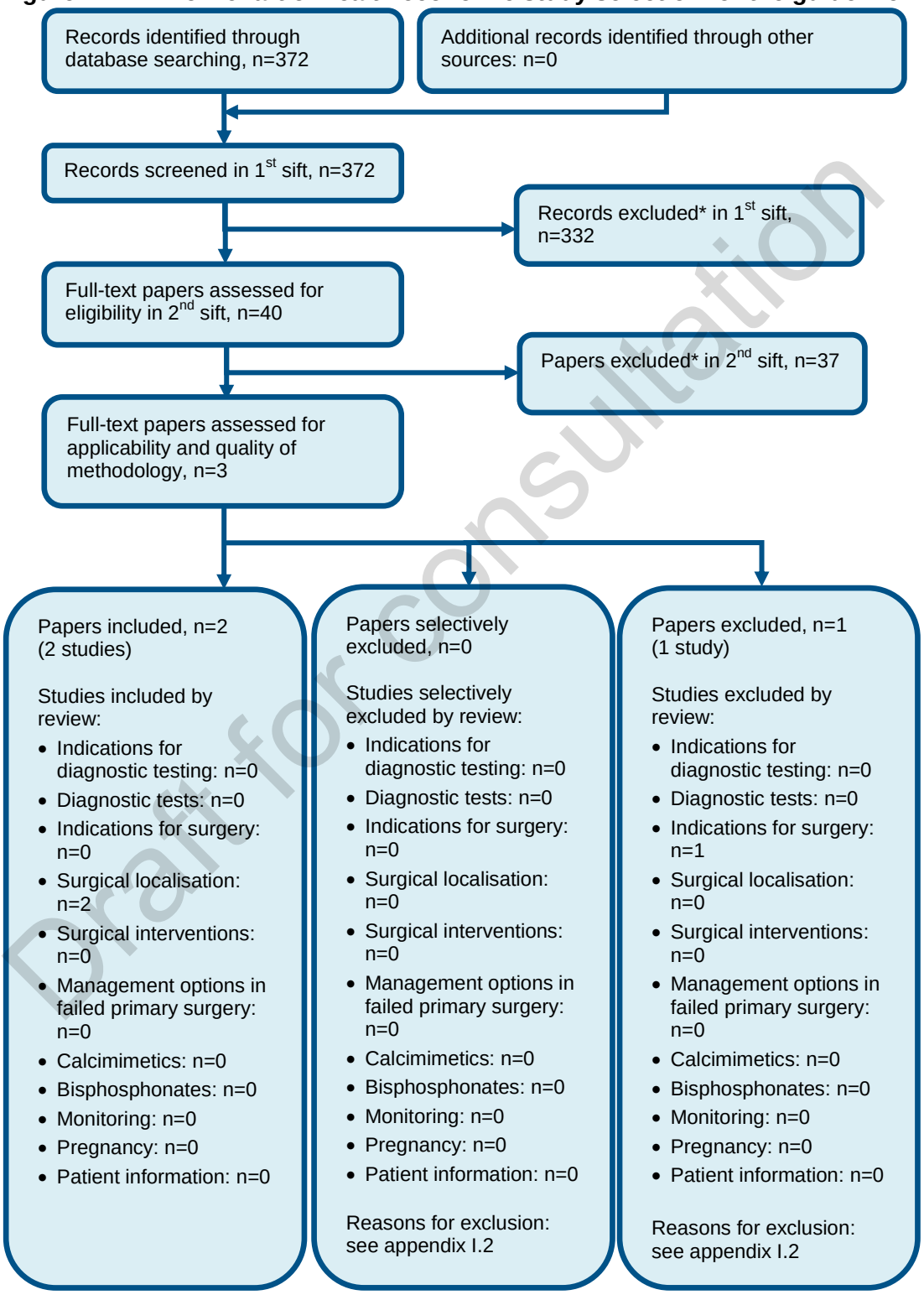
| Quality assessment | | | | | | No of patients | | Effect | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | Quality | Importance |

| No of studies | Design | Risk of bias | Inconsistency | Indirectness | Imprecision | Other considerations | Surgery | Conservative treatment (NRS) | Relative (95% CI) | Absolute | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Mortality** | | | | | | | | | | | | |
| 2 | observational studies | very serious[a] | no serious inconsistency | no serious indirectness | no serious imprecision | none | - | -[c] | HR 0.65 (0.57 to 0.74) | -[3] | VERY LOW | CRITICAL |
| **Fractures** | | | | | | | | | | | | |
| 2 | observational studies | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | none | - | -[c] | HR 0.67 (0.55 to 0.82) | -[3] | VERY LOW | CRITICAL |
| **Cancer** | | | | | | | | | | | | |
| 1 | observational studies | very serious[a] | no serious inconsistency | no serious indirectness | Serious[b] | None | 135/1934 (7%) | 119/1279 (9.3%) | HR 1.11 (0.9 to 1.37) | 10 more per 1000 (from 9 fewer to 32 more) | VERY LOW | IMPORTANT |
| **Kidney stones** | | | | | | | | | | | | |
| 1 | observational studies | very serious[a] | no serious inconsistency | no serious indirectness | no serious imprecision | None | 297/1934 (15.4%) | 83/1279 (6.5%) | HR 1.87 (1.3 to 2.69) | 53 more per 1000 (from 19 more to 100 more) | VERY LOW | CRITICAL |

1 [a] Downgraded by 1 increment if the majority of studies were at high risk of bias, and downgraded by 2 increments if the majority of studies were at very high risk of bias.
2 [b] Downgraded by 1 increment if the confidence interval crossed 1 MID, and downgraded by 2 increments if the confidence interval crossed both MIDs.
3 [c] Control group rate not reported

# Appendix G: Health economic evidence selection

**Figure 27: Flow chart of health economic study selection for the guideline**




* Non-relevant population, intervention, comparison, design or setting; non-English language

# Appendix H: Health economic evidence tables

None.

# 1 Appendix I: Excluded studies

## I.1 2 Excluded clinical studies

3 **Table 13: Studies excluded from the clinical review**

| Study | Exclusion reason |
|---|---|
| Adler 2008[1] | Inappropriate comparison – study compares different types of surgery |
| Agus 1993[2] | An opinion piece |
| Alhava 1988[3] | Non-comparative before and after study |
| Almqvist 2002[5] | No relevant outcomes |
| Almqvist 2004[4] | Inappropriate comparison. Incorrect interventions. Comparison of different timings of surgery. |
| Alvarez-Allende 2014[6] | Conference abstract |
| Anonymous 2000[9] | Not a primary study – article |
| Anonymous 2000[8] | Not a primary study – article |
| Barkun 2006[10] | Commentary of an included RCT |
| Blanchard 2014[12] | Non-comparative before and after study |
| Bollerslev 2009[14] | No relevant outcomes |
| Bonzelaar 2016[15] | Conference abstract |
| Britton 1971[16] | Non-comparative study |
| Brothers 1987[17] | Non-comparative study |
| Broulik 2011[18] | Non-comparative before and after study |
| Bruining 1981[19] | Non-comparative study |
| Burney 1996[20] | Non-comparative study |
| Burney 1998[21] | Non-comparative study |
| Calo 2016[22] | Inappropriate comparison |
| Carneiro-pla 2007[23] | Non-comparative study (all patients underwent surgery) |
| Chen 1998[24] | Non-comparative study |
| Cheng 2015[25] | Systematic review. Screened for relevant references. |
| Chigot 1995[26] | Non-comparative study (all patients underwent surgery) |
| Cowie 1982[28] | Incorrect study design – case series |
| D'Andrea 1996[29] | Non-comparative study (all patients underwent surgery) |
| Diaz-Guerra 2015[31] | Conference abstract |
| Dy 2012[32] | Non-comparative study (all patients underwent surgery) |
| Edwards 2006[33] | Non-comparative study (all patients underwent surgery) |
| Espiritu 2011[35] | No relevant outcomes reported |
| Falkheden 1980[36] | Non-comparative study (all patients underwent surgery) |
| Fang 2008[37] | NRS - no multivariate analysis or adjustment for confounders |
| Farnebo 1984[38] | Non-comparative study (all patients underwent surgery) |
| Freaney 1978[39] | Non-comparative study (all patients underwent surgery) |
| Ghose 1981[40] | Non-comparative before and after study |
| Hagstrom 2006[41] | Non-comparative before and after study |
| Hedback 1990[43] | Non-comparative retrospective study |
| Hedback 1991[42] | Non-comparative retrospective study |

| | |
|---|---|
| Horiuchi 2002[44] | Inappropriate intervention – 2-week administration only of oral etidronate. This bisphosphonate is no longer used. |
| Jansson 2006[45] | Conference abstract |
| Khosla 1999[46] | NRS – only reports the effect of surgery on fracture risk from a univariate model and not the adjusted HR for this factor from the MV model. |
| Lafferty 1989[47] | Non-comparative study (all patients underwent surgery) |
| Larsson 1993[48] | NRS with no adjustment for confounders |
| Leong 2010[49] | Non-comparative study (all patients underwent surgery) |
| McDow, 2018 [52] | Review. Screened for relevant references. |
| Melton 1992[53] | NRS – surgery effect on fracture risk only reported from a univariate model (risk adjusted for confounders not reported). |
| Mole 1992[54] | NRS with no adjustment for confounders. Study also provides an analysis of eight people who underwent surgery compared with eight age-matched conservatively managed people (but other key confounders not matched). |
| Morris 2010[55] | No relevant outcomes reported – for some outcomes results are only reported for the intervention group. Paper includes a statement that there was no morbidity or mortality but it is unclear if this refers to both the intervention and control group or just the control group. |
| Nomura 2004[57] | NRS with no adjustment for confounders |
| Nordenstrom 2004[58] | Non-comparative before and after study |
| Oucharek 2011[59] | Non-comparative study (all patients underwent surgery) |
| Paloyan 1983[60] | Non-comparative study (all patients underwent surgery) |
| Perrier 2009[61] | No relevant outcomes |
| Persson 2011[62] | Follow-up study of an included RCT but with no relevant outcomes |
| Posen 1985[63] | NRS with no adjustment for confounders |
| Rao 2003[65] | NRS with no adjustment for confounders |
| Richmond 2007[66] | Non-comparative study |
| Rolighed 2012[67] | Conference abstract |
| Rubin 2008[68] | NRS with no adjustment for confounders |
| Sankaran 2010[69] | A literature review not specified as systematic review and without quality assessment of the studies included |
| Sanzenbacher 1970[70] | Inappropriate study design |
| Saponaro 2013[71] | Incorrect interventions |
| Schneider 2014[72] | Inappropriate comparison. Incorrect interventions. |
| Scott Jr 1981[73] | Inappropriate study design |
| Sejean 2005[74] | Incorrect study design – decision analysis |
| Silverberg 1995[75] | Non-comparative study (all patients underwent surgery) |
| Silverberg 1999[76] | NRS – study performed a multivariate analysis but factors included are unclear and no adjusted risk given for the effect of surgery on the outcome |
| Singh Ospina 2016[77] | Systematic review screened for references |
| Singh Ospina 2016[78] | Systematic review screened for relevant references |
| Siperstein 1992[79] | Non-comparative study (all patients underwent surgery) |
| Solorzano 2008[80] | Non-comparative retrospective case series |
| Soreide 1997[81] | Non-comparative study (all patients underwent surgery) |
| Strewler 1995[82] | Literature review with commentary and opinion |
| Tay 2016[84] | NRS with multivariate analysis but no relevant outcomes |

| | |
|---|---|
| Tisell 1983[85] | Inappropriate comparison. Inappropriate study design. |
| Trombetti 2016[86] | NRS with no adjustment for confounders |
| Vera 2014[89] | NRS with no adjustment for confounders |
| Vestergaard 2003[91] | Overlap in recruitment of participants with an already included study (Vestergaard 2003) – larger study included in this review |
| Wagner 2007[92] | Review |
| Wermers 1998[93] | NRS with mulitvariate analysis but the effect of surgery on risk of death is not reported from the univariate or multivariate analysis |
| Witteveen 2010[94] | Non-comparative study (all patients underwent surgery) |
| Wu 2010[95] | Inappropriate comparison |
| Yeh 2016[96] | NRS – adjusted relative risk for the effect of surgery on fracture risk not reported |
| Yu 2010[97] | Inappropriate comparison |
| Zhao 2014[98] | Conference abstract |

# I.2 Excluded health economic studies

**Table 14: Studies excluded from the health economic review**

| Reference | Reason for exclusion |
|---|---|
| Sejean 2005 [74] | This study was assessed as partially applicable with very serious limitations. The study took a non-UK perspective, and quality of life was not reported directly from patients. Furthermore, the analysis was based on multiple clinical studies (mostly cohort or case-series studies) that have been excluded from this review. In addition, it was considered that there were some assumptions that were likely to be biasing the results, namely that there is no resource use impact from progression, only that some people would then have surgery. Therefore this study was selectively excluded. |