National Institute for Health and Care Excellence

Final

Hyperparathyroidism (primary): diagnosis, assessment and initial management

[I] Evidence review for monitoring

NICE guideline NG132

Intervention evidence review

May 2019

Final

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 All NICE guidance is subject to regular review and may be updated or withdrawn.

ISBN: 978-1-4731-3415-7

Contents

1 Monitoring

1.1 Review question: What is the optimum type and frequency of monitoring for people with primary hyperparathyroidism (for example, pre-operative, postoperative, non-surgical)?

1.2 Introduction

There is uncertainty regarding the long-term sequelae of primary hyperparathyroidism (PHPT) even in people who have undergone successful parathyroidectomy. This is reflected in variation in practice regarding what should be monitored and for how long. Monitoring for end organ damage is more established although monitoring for cardiovascular events and cancer is less clear. The purpose of this review is to identify the optimum type and frequency of monitoring for people with PHPT. One approach to this question is to understand the long-term outcomes in people with PHPT.

1.2.1 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 primary hyperparathyroidism Strata: • Pre-operative • Post-operative • Non-surgical • Previous surgery • Pregnant women
Interventions	Different techniques/tools/variables being monitored • blood tests (adjusted serum calcium, serum creatinine), • imaging (DXA for bone disease (±VFA) • US for kidney stones, renal tract calcification • x-ray for fragility fracture for vertebral fracture • other (24-hour urinary calcium, creatinine clearance) • CV variables (BP, lipids, ECG), • vitamin D (for post-operative monitoring) • renal function Different frequencies of monitoring (this may differ for the different tests – e.g. bloods 6–12 months, DXA 2–3 years) Different durations of monitoring (for post-surgery stratum, e.g. for 3 months or 6 months – to include the optimum timing of serum calcium assessment post-surgery to determine cure)
Comparisons	• Comparing types of monitoring strategies to each other • Comparing different frequencies of the same strategy
Outcomes	Critical outcomes: • HRQOL (continuous outcome) • Mortality (dichotomous outcome)

	Important outcomes: • Deterioration in renal function (continuous – study may also report renal replacement) • Fractures (vertebral or long bone) (dichotomous outcome) • Occurrence of kidney stones (dichotomous outcome) • Persistent hypercalcaemia (dichotomous outcome) • BMD of the distal radius or the lumbar spine (continuous) • Cardiovascular events (dichotomous outcome) • Adverse events (to include voice change, hypoparathyroidism, hypothyroidism/hyperthyroidism; dichotomous outcome) • Cancer incidence (dichotomous outcome) • Reoperation (for post-surgery stratum)
Study design	RCTs and systematic reviews of RCTs In the absence of RCT evidence for the critical outcomes, NRSs will be included (only if the following key confounders are matched for or adjusted for in the analysis) Key confounders: • Age • Absence/presence of end-organ effects • Adjusted serum calcium level

1.3 Review question: What are the long-term outcomes in people with primary hyperparathyroidism?

1.3.1 PICO table

For full details see the review protocol in appendix A.

Table 2: PICO characteristics of review question

Population	Adults (18 years or over) with primary hyperparathyroidism Strata: • Preoperative • Non-surgical • Post-operative • People on calcimimetics • People on bisphosphonates • Previous surgery • Normocalcaemic patients Subgroup: • People on HRT
Interventions and comparisons	This review will be looking at the incidence of outcomes in people with PHPT compared with healthy controls
Outcomes	Critical outcomes: • mortality (dichotomous outcome) • fragility fracture (dichotomous outcome) • renal stones (dichotomous outcome)

	• renal tract calcification (dichotomous outcome) • pancreatitis (dichotomous outcome) • stroke (dichotomous outcome) • hypertension (dichotomous outcome) • myocardial infarction (dichotomous outcome) • Number of people who become eligible for surgery/meet the criteria for surgery (dichotomous) • serum calcium (>2.85 mmol/l) (dichotomous) (continuous if dichotomous not available) • 24-hour urine for calcium (>10 mmol/dl) (dichotomous) (continuous if dichotomous not available) • BMD of proximal femur (T-score <2.5; Z score <2) (dichotomous) (continuous if dichotomous not available) Follow-up: minimum 2 years
Study design	Prospective cohort studies Retrospective cohort studies will be included only if insufficient prospective cohort studies are identified Key confounders: • Age • Absence/presence of end-organ effects • Adjusted serum calcium level

1.4 Clinical evidence

1.4.1 Included studies

A search was conducted for assessing the optimum type and frequency of monitoring for people with PHPT. No evidence was identified for this review question. See the study selection flow chart in appendix C.

A second search of the original PHPT search was conducted to determine whether PHPT is associated with poor long-term outcomes and to determine what monitoring strategies they need to undergo. The aim of the review was to look at the incidence of outcomes in people with PHPT compared with healthy controls. It was thought that any difference in all/some of the outcomes would mean that someone with PHPT would need to be monitored.

Eleven comparative studies: Clifton-Bligh 2015 [21]; De Geronimo 2006 [26]; Hedback 1998 [38]; Kenny 1995 [41]; Khosla 1999 [42]; Larsson 1993 [45]; Melton 1992 [49]; Ronni-Sivula 1985 [60]; Su 2008 [69]; Wilson 1988 [79]; Yu 2011 [80] were included.

Only one prospective cohort study, Yu 2011 [80], adjusted for all key confounders. This study evaluated the risk of mortality and morbidity among untreated mild PHPT patients compared with a matched cohort. The study adjusted for a number of potential confounding variables (multiple deprivation index [SIMD], history of bisphosphonates prescription, history of hospital admitted CVD, cerebrovascular disease, hypertension, renal failure, renal stones, psychiatric disease, fractures, cancer and diabetes) and the propensity of having calcium checked in the analysis. In 5 studies (Clifton-Bligh, 2015 [21]; De Geronimo, 2006 [26]; Hedback 1998 [38]; Melton, 1992 [49]; Ronni-Sivula, 1985 [60]) the control group was matched for factors such as age and gender but these studies did not adjust for serum calcium level and absence/presence of end

organ effects (no multivariate analysis conducted). The remaining 5 studies did not match or adjust for any key confounders (no multivariate analysis conducted).

The definition of mild and non-mild PHPT was not consistent across the studies. The definitions reported in the studies are noted in their respective evidence tables.

The studies were stratified as non-surgical, pre-operative, mixed pre and post-operative (including surgery and non-surgery patients) and post-operative.

There were 2 studies in the strata non-surgical (mild asymptomatic patients) (Yu 2011; Wilson 1988); 1 study in pre-operative (Suh 2008); 5 studies in the mixed pre and post-operative (mild and non-mild patients): (Melton 1992; De Geronimo 2006; Clifton-Bligh 2015; Khosla 1999; Larsson 1993); and 3 studies in the post-operative (Ronni-Sivula 1985; Kenny 1995; Hedback 1998). No evidence was available for people on calcimimetics, bisphosphonates and normocalcaemic patients.

Evidence from the studies are summarised in the clinical evidence summary tables below (Table 4, Table 5, Table 6, Table 7, Table 8, Table 9). 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.

1.4.2 Excluded studies

See the excluded studies list in appendix I.

1.4.3 Summary of clinical studies included in the evidence review

Table 3: Summary of studies included in the evidence review (comparative studies)

Study	Comparison	Population	Outcomes	Comments
Yu 2011 [80] Prospective cohort study UK Setting: Tayside, Scotland , 1997-2006	Each of the selected patients was then matched with five individuals, or comparators, by age, gender and calendar year of PHPT diagnosis, from the general Tayside population, with either no calcium records or normal serum calcium concentration during the study period. The calendar year of the matching was the index date for each comparator.	All Tayside residents over 20 years of age were considered as potential cases. In conjunction with hospital admission records, nuclear medicine and histology data, the PHPT cohort was primarily diagnosed if they met either of the following biochemical criteria: albumin-corrected serum calcium >2.55 mm (10.22 mg/dl, reference range 2.10–2.55 mm) on at least two occasions, with plasma parathyroid hormone (PTH) concentration >3 pm (13.5 ng/l, reference range 1.0–6.9 pm); or, albumin-corrected serum calcium >2.55 mm (10.22 mg/dl) on a single occasion, with plasma PTH concentration >6.9 pm (31.05 ng/l). A subgroup of 'mild PHPT' patients was further selected for this study, these being patients with untreated	• All-cause mortality • Fatal CVD • Non-fatal CVD • Cerebrovascular disease • Hypertension • Renal failure • Renal stones • All fractures • Osteoporotic fractures	PHPT patients not fitting NIH criteria for referral. Study used individual data for both cases and comparators.

Study	Comparison	Population	Outcomes	Comments
		PHPT, low concentrations of hypercalcaemia and an absence of renal stones, renal failure and osteoporotic fractures, at diagnosis. Mild untreated group were defined as untreated PHPT patients whose serum calcium concentrations were <2.9 mmol/l within the first 6 months after a positive diagnosis with absence of previous (prior to PHPT diagnosis) fracture fragility, renal stones and renal failure and not treated with cinacalcet. Further exclusion criteria were applied to the mild untreated group: serum calcium was followed up for >6 months; less than two serum calcium measurements within the first 6 months.		
Clifton-Bligh, 2015 [21] Retrospective cohort study	n=561 patients with PHPT Patients diagnosed with PHPT between 1961 and 1994 were identified, medical records were obtained and examined, and a determination was made as to	Survival data of individuals with PHPT compared with expected survival in the general Australian population was obtained from the Life Tables.	Survival rate over a 10 and 20 year period	Follow-up: 10 years and 20 years The relative survival over a 20 year time interval was calculated for the patients studied between 1972 and 2011. The group was divided into 2 cohorts: those

Study	Comparison	Population	Outcomes	Comments
Australia	whether or not they were alive at the end of 1994.	The control population was the Australian population at large for whom Life Tables from 1961-1994 existed at the time (Life Tables are published by the Australian Government Actuary). Control population matched for age, sex, the year observation began, and the duration of the observation.		diagnosed between 1972 and 1981 and those diagnosed between 1982 and 1991, and 20 year relative survival was calculated for each cohort. Before 1972 diagnosis of PHPT was made if surgical removal of a parathyroid gland restored eucalcaemia, or if a full investigation failed to find another cause for hypercalcaemia. After 1972, the diagnosis of PHPT was made if the serum calcium and serum PTH was above the upper limit of the reference range. Because of the concept that a person with mild PHPT might not require surgery, 113 of the patients with mild PHPT were not subjected to neck exploration and 448 patients had parathyroid surgery.
De Geronimo, 2006 [26] Cohort study Italy	n=98 post-menopausal women with PHPT and 89 healthy women. 98 consecutive post-menopausal patients with PHPT. They were grouped as mild or non-mild according to criteria established by the	Mild or non-mild PHPT post-menopausal women vs Healthy subjects - 89 healthy postmenopausal women, matched for age, years since menopause, BMI. They were	• Vertebral fracture • Non-vertebral fractures • Lumbar spine – BMD • Femoral neck – BM • Total femur – BMD	Control patients matched for age, years since menopause and BMI. Not adjusted for absence/presence of end-organ effects and serum calcium level. Diagnosis of PHPT was made according to the conventional clinical and lab data, including a

Study	Comparison	Population	Outcomes	Comments
Setting: Mineral metabolic centre	Consensus Development conference on the management of asymptomatic primary hyperparathyroidism: Serum calcium greater than 1m g/dl above the upper limits of normal; 24-h total urine calcium excretion of more than 400 mg; creatinine clearance reduced by more than 30% compared to age-matched persons; bone density at lumbar spine, hip or distal radius that is more than 2.5 SD below peak bone mass; patients under 50 years of age. Based on these criteria, only 25 of the 98 patients were considered as suffering from mild disease. In this group of patients hypercalcaemia was occasionally detected in the course of the standard biochemical evaluation performed on all subjects undergoing BMD measurement. None of the mild patients reported either height loss or clinical vertebral fractures at history. In the remaining 73 patients, a severe disease was present; 13 of them had a history of non-vertebral fractures; 7 referred	randomly selected from ambulatory post-menopausal women sent by their GPs to the hospital as part of a menopause-screening programme.		history of at least 1 year of prolonged hypercalcaemia without evidence of non-parathyroid aetiology and unsuppressed serum levels of immunoreactive PTH.

Study	Comparison	Population	Outcomes	Comments
	by GPs for nephrolithiasis; 21 referred by GPs for osteoporosis; 15 patients renal stones shown by ultrasonography; 3 were referred for hypercalcaemia; 2 had a life threatening episode of pancreatitis. The remaining 12 patients complained of bone pain and/or neuromuscular symptoms.			
Headback 1998 [38] Retrospective cohort study Sweden	n=4461 (915 men and 3546 women) Patients diagnosed as having PHPT or parathyroid adenoma according to the International Classification of Diseases during 1987-94 from the Swedish national patient registry. The patient series studied consisted of those individuals who at the same time were reported to have undergone removal of a parathyroid gland or adenoma. The inclusion date of a patient was the date of arrival at the hospital before surgery.	PHPT patients vs Control Whole Swedish population matched for age, sex and calendar year. Strata: Post-operative	Death Number of patients operated for PHPT	Total number of observation years was 3205 giving mean follow-up times of 3.6 and 3.5 years respectively (range 0-8 years)
Kenny 1995 [41]	n=46 Post-menopausal women who had undergone	PHPT patients vs control	Fracture incidence	Follow-up: 5 years

Study	Comparison	Population	Outcomes	Comments
Retrospective cohort study USA	parathyroidectomy for hyperparathyroidism during a 5 year period (1986 to 1991) Control – n=44 postmenopausal women without hyperparathyroidism contacted by random digit dialling and interviewed as controls.	Strata : Post-operative		The women in the control were group were similar to those with PHPT in regard to age, weight, and height. Not adjusted for key confounders.
Khosla 1999 [42] USA Retrospective Cohort study	n=407 patients with PHPT Patients with PHPT: The database the Rochester Epidemiology Project was used to identify 435 Rochester residents with PHPT during period, 1965-1992. Seven of these patients refused subsequent authorisation for chart review, and 21 had no follow-up after age 35 years, resulting in a final cohort of 407 subjects. These subjects were then followed forward in time through their linked medical records in the community (retrospective cohort study) until death or the most recent clinical contact and backward in time to the first medical record	Age related fractures in patients with PHPT (1965-1992) vs Age related fractures in general population The majority were women (344, 775) and most were 45 years of age or older at the diagnosis of HPT (335, 82%). The mean age at diagnosis was 57.8 years. The average maximum serum calcium level was 10.9 (0.6) mg/dl. Median serum PTH, measured by a C-terminal assay, was 47 µl.eq/ml (25-75% interval, 33-71 µl.eq/ml; normal ≤50 µl.eq/ml). The majority of the patients were asymptomatic.	Fractures	Study - 28 year period No multivariate analysis Diagnosis of vertebral fracture was accepted on the basis of a radiologists' report of compression or collapse of one or more thoracic or lumbar vertebrae. All fractures were classified according to the circumstances of the injury.

Study	Comparison	Population	Outcomes	Comments
	entry in the community.			
Larsson, 1993 [45] Sweden Retrospective cohort study	n=1373 patients with PHPT All patients who were admitted to hospital with the diagnosis of PHPT. 1318 patients (69%) had been subjected to parathyroidectomy at the same or a subsequent admission. Mean serum calcium at the time of diagnosis for women with PHPT was 2.91 (0.22), and for men 2.79 (0.18) mmol/l; p<0.01. The serum calcium levels at diagnosis were stable during the study period. For each patient the observation period started in 1965, or at the age of 30 for those below this age in 1965 and ended at the date of hip fracture diagnosis, death, or at the end of 1983.	Patients with PHPT vs Entire background population	Hip fracture	Follow-up - 19 years No multivariable analysis
Melton, 1992 [49] Rochester, USA	n=90 patients with PHPT; n=90 matched control subjects 90 cases of PHPT newly	Patients with PHPT vs matched control subjects	Fractures Age interaction (narrative)	Proportional hazards model was used to determine the relative influence of various clinical characteristics on subsequent fracture risk among those with

Study	Comparison	Population	Outcomes	Comments
Population based retrospective cohort study	diagnosed during the life among Rochester during 1965 through 1976; 83 subjects had histopathologic proof of parathyroid adenoma or hyperplasia or had hypercalcaemia with pathognomonic radiographic signs and/or elevated serum parathyroid hormone concentrations. The remainder had hypercalcaemia for more than a year without another cause being found after careful evaluation. Patients with an incidental autopsy diagnosis of parathyroid adenoma or hyperplasia were excluded. The 90 patients were matched by age and gender to control subjects from the local population who had no evidence of PHPT.	Controls matched by age and gender.		HPT. The factors evaluated included age at diagnosis, gender, initial serum calcium (<2.74 vs >2.74 mmol/L), comorbid conditions related to HPT and parathyroid surgery. Not adjusted for serum calcium level and absence/presence of end organ effects.
Ronni-Sivula, 1985 [60] Finland Retrospective cohort study	n=334 (PHPT); n= 334 healthy controls 334 patients (83 men and 251 women) were operated on for PHPT in the years 1956-79. A follow-up study of this material was performed in the years 1980-82. 34 patients had died before the end of the year	PHPT patients operated in the years 1956-79 vs Control.	Mortality	No multivariate analysis Not adjusted for serum calcium level and absence/presence of end organ effects.

Study	Comparison	Population	Outcomes	Comments
Setting: department of surgery	1980. All clinical data as well as causes of death of these patients were collected and examined.	Each PHPT patient in the original material was given a pair who was sex and age matched and operated on for appendicitis, varicose veins or haemorrhoids in the same year Strata : Post-operative		
Suh 2008 [69] Retrospective cohort study USA	n=271 All patients who had undergone renal imaging within 6 months before parathyroid surgery for PHPT	n=500 Age matched subjects who had right upper quadrant sonograms obtained for various reasons Strata : Post-operative	Renal stones	No multivariate analysis
Wilson, 1988 [79] Retrospective cohort study USA Setting: Outpatient department of a bone and mineral metabolism clinic.	n=174 PHPT; n=200 control All patients with mild asymptomatic PHPT. Diagnostic criteria for PHPT: persistent hypercalcaemia (calcium ≥2.65 mmol/L), with no clinical indication of another cause, and evidence of parathyroid hormone hypersecretion, with increased or non-suppressed values of radioimmunoassay of	PHPT group vs Healthy white women. These data had been obtained in 200 ambulatory white female patients having routine annual physical examination. In almost every patient hypercalcaemia was an	Prevalence of vertebral fractures Follow-up: 10 years A reduction in anterior height of more than 20% compared with an adjacent vertebra was classified as a fracture.	Prospectively collected data were retrospectively analysed and compared with data from historical control group. No multivariate analysis

Study	Comparison	Population	Outcomes	Comments
	parathyroid hormone or nephrogenous cyclic adenosine monophosphate excretion per unit of glomerular filtrate. The diagnosis was made at the clinic between 1 January 1976 and December 1985. None of the post-menopausal women were receiving oestrogen replacement therapy at the time of initial diagnosis. Inclusion criteria: Absence of symptoms due to PHPT, no current kidney stone disease, a plasma calcium level of less than 3.00 mmol/L, a plasma creatinine level of less than 133µmol/L, no radiographic evidence of osteitis fibrosa, and a forearm bone densitometry not more than 2.5 standard deviations below the mean value expected for age, sex and race.	incidental finding. The mean plasma calcium level was 2.77 (0.09) mmol/L (reference value, 2.40 [0.08]).		

See appendix D for full evidence tables.

1.4.4 Quality assessment of clinical studies included in the evidence review

Table 4: Clinical evidence summary: PHPT cases compared to matched comparators (adjusted for key confounders) (Stratum-Non-surgical)

Outcomes	No of Participants	Quality of the evidence	Relative effect	Anticipated absolute effects

	(studies) Follow up	(GRADE)	(95% CI)	Risk with Control	Risk difference with Cases (95% CI)
All-cause mortality	8544 (1 study) median 2.9 years	LOW[c]	HR 1.64 (1.43 to 1.87) [b]	Moderate	
					-[a]
Fatal CVD	8544 (1 study) median 2.9 years	LOW[c]	HR 1.64 (1.32 to 2.04) [b]	Moderate	
					-[a]
Nonfatal CVD	8544 (1 study) median 2.9 years	LOW[c]	HR 2.48 (2.13 to 2.89) [b]	Moderate	
					-[a]
Hypertension	8544 (1 study) median 2.9 years	LOW[c]	HR 2.60 (2.04 to 3.31) [b]	Moderate	
					-[a]
Cerebrovascular disease	8544 (1 study) median 2.9 years	LOW[c]	HR 2.51 (1.95 to 3.22) [b]	Moderate	
					-[a]
Renal failure	8544 (1 study) median 2.9 years	LOW[c]	HR 13.83 (10.41 to 18.37) [b]	Moderate	
					-[a]
Renal stones	8544 (1 study) median 2.9 years	LOW[c]	HR 5.15 (2.69 to 9.83) [b]	Moderate	
					-[a]
All fractures	8544 (1 study) median 2.9 years	LOW[c]	HR 1.75 (1.36 to 2.26) [b]	Moderate	
					-[a]
Osteoporotic fractures	8544 (1 study) median 2.9 years	LOW[c]	HR 1.63 (1.22 to 2.19)[b]	Moderate	
					-[a]

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Cases (95% CI)

[a] Absolute effect cannot be calculated as adjusted control group event rate not reported.
[b] Confounding covariates considered were multiple deprivation index, history of bisphosphonates prescription, history of hospital-admitted CVD, cerebrovascular disease, hypertension, renal failure, renal stones, psychiatric disease, fractures, cancer and diabetes.
[c] 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 (used the non-randomised studies checklist on evibase for assessment for risk of bias).
Note: All patients with diagnosed but untreated, mild asymptomatic PHPT.

Table 5: Clinical evidence summary: PHPT compared to control (no multivariate analysis) (Stratum – Pre-surgery)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Monitoring long term outcomes - PHPT (95% CI)
Renal stones (definitive calculi)	771 (1 study)	VERY LOW[a] due to risk of bias	RR 4.38 (1.94 to 9.88)	Moderate	
				16 per 1000	54 more per 1000 (from 15 more to 142 more) PHPT group: 19/271 Control: 8/500

[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 (used the non-randomised studies checklist on evibase for assessment for risk of bias).

Table 6: Clinical evidence summary: PHPT compared to control (no multivariate analysis) – Mixed strata – Pre and post-operative (surgery and non-surgery patients)

Outcomes	No of	Quality of the	Relative	Anticipated absolute effects

	Participants (studies) Follow up	evidence (GRADE)	effect (95% CI)	Risk with control	Risk difference with Monitoring long term outcomes - PHPT (95% CI)
Fractures (all)	180 (1 study)	VERY LOW[a] due to risk of bias,	RR 0.96 (0.74 to 1.24)	Moderate	
				414 per 1000	17 fewer per 1000 (from 108 fewer to 99 more)
Lumbar spine BMD (mg/cm^2) (mild PHPT vs healthy women)	114 (1 study)	VERY LOW[a] due to risk of bias	-	-	The mean lumbar spine BMD (mg/cm^2) (mild PHPT vs healthy women) in the intervention groups was 77.1 higher (31.61 to 122.59 higher)
Femoral neck- BMD (mg/cm^2) (mild PHPT vs healthy women)	114 (1 study)	VERY LOW[a,c] due to risk of bias, imprecision	-	-	The mean femoral neck - BMD (mg/cm^2) (mild PHPT vs healthy women) in the intervention groups was 18.4 higher (24.43 lower to 61.23 higher)
Total femur - BMD (mg/cm^2) (mild PHPT vs healthy women)	114 (1 study)	VERY LOW[a,c] due to risk of bias, imprecision	-	-	The mean total femur - BMD (mg/cm^2) (mild PHPT vs healthy women) in the intervention groups was 21.2 higher (30.33 lower to 72.73 higher)
Vertebral fractures (mild PHPT vs healthy controls)	343 (2 studies)	VERY LOW[a,b] due to risk of bias, inconsistency	RR 2.31 (1.26 to 4.21)	Moderate	
				72 per 1000	94 more per 1000 (from 19 more to 231 more)
Vertebral fractures (non-mild PHPT vs healthy women)	162 (1 study)	VERY LOW[a] due to risk of bias	RR 5.33 (2.64 to 10.77)	Moderate	
				90 per 1000	390 more per 1000 (from 148 more to 879 more)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with control	Risk difference with Monitoring long term outcomes - PHPT (95% CI)
Non-vertebral fractures (mild PHPT vs healthy women)	114 (1 study)	VERY LOW[a,c] due to risk of bias, imprecision	Peto OR 0.22 (0.07 to 0.78)	Moderate	
				191 per 1000	190 fewer per 1000 (from 290 to 90 fewer)
Non-vertebral fractures (non-mild PHPT vs healthy women)	162 (1 study)	VERY LOW[a,c] due to risk of bias, imprecision	RR 0.93 (0.49 to 1.79)	Moderate	
				191 per 1000	13 fewer per 1000 (from 97 fewer to 151 more)
Lumbar spine-BMD (mg/cm^2) (non-mild PHPT vs healthy women)	162 (1 study)	VERY LOW[a] due to risk of bias	-	-	The mean lumbar spine - BMD (mg/cm^2) (non-mild PHPT vs healthy women) in the intervention groups was 73.6 lower (116.15 to 31.05 lower)
Femoral neck-BMD (mg/cm^2) (non-mild PHPT vs healthy women)	162 (1 study)	VERY LOW[a] due to risk of bias	-	-	The mean femoral neck - BMD (mg/cm^2) in the intervention groups was 89.3 lower (121.96 to 56.64 lower)
Total femur-BMD (mg/cm^2) (non-mild PHPT vs healthy women)	162 (1 study)	VERY LOW[a] due to risk of bias	-	-	The mean total femur - BMD (mg/cm^2) in the intervention groups was 102.5 lower (140.13 to 64.87 lower)

[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 (used the non-randomised studies checklist on evibase for assessment for risk of bias).
[b] Heterogeneity, I^2=90%
[c] Downgraded by 1 increment if the confidence interval crossed 1 MID, and downgraded by 2 increments if the confidence interval crossed both MIDs.

Table 7: Clinical evidence summary (modified GRADE table): PHPT compared to control (no multivariate analysis) – Mixed strata- Pre and post-operative (surgery and surgery patients)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative survival rate *for the PHPT group compared to control	Anticipated absolute effects
Relative survival	561 (1 study)	VERY LOW[a] due to risk of bias	86.8% (95% CI 84.9-86.2, p<0.001) (10 years)	124/561 patients died between 1961 and 1994[b]
Relative survival	561 (1 study)	VERY LOW[a] due to risk of bias	62.9% (95% CI 58.5- 67.4, P<0.001). (20 years)	-

[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 (used the non-randomised studies checklist on evibase for assessment of risk of bias).

[b] Control group event rate not reported

* Relative survival rate - is a way of comparing the survival of people who have a specific disease with those who do not, over a certain period of time. It is calculated by dividing the percentage of patients with the disease who are still alive at the end of the period of time by the percentage of people in the general population/control who are alive at the end of the same time period.

Table 8: Clinical evidence summary (modified GRADE table): PHPT (observed) versus expected numbers in the general population (no multivariate analysis) – Mixed strata – Pre and post-operative (surgery and non-surgery patients)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Standardised Incidence Ratio* (SIR) 95% CI and Risk ratio (RR)	Anticipated absolute effects
Any Fracture (at all skeletal sites)[b]	407 (1 study)	VERY LOW[a] due to risk of bias	SIR 1.3 (95% CI 1.1-1.5)*	observed: n=202; expected: n=154.6
Vertebral fracture	407 (1 study)	VERY LOW[a] due to risk of bias	SIR 3.2 (95% CI 2.5-4.0)	observed: n=79; expected: n= 24.6
Hip fractures (women)	1373 (1 study)	VERY LOW[a] due to risk of bias	RR 0.93 (95% CI 0.72-1.19)	observed: 67/1373; expected: 71.76
Hip fractures (men)	551 (1 study)	VERY LOW[a] due to risk of bias	RR 1.39 (95% CI 0.69-2.50)	observed 11/551; expected 7.9

[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 (used the non-randomised studies checklist on evibase for assessment of risk of bias).
[b] A total of 471 fractures occurred in 202 patients after the index date. The majority of fractures of the vertebrae (92%), distal forearm (91%), pelvis (75%), and proximal femur (90%) were due to mild or moderate trauma.

* Predictors of the risk of mild/moderate trauma, vertebral, distal forearm or proximal femur fractures:
Multivariate model
Age (per 10 year increase): relative hazard 1.6 (95% 1.4-1.9)
Female gender: relative hazard 2.3 (95% CI 1.2-4.1)
By multivariate analysis, only age and female gender were significant independent predictors of fracture risk.
**Fracture risk was assessed by comparing new fractures at each site to the number expected from gender and age specific fracture incidence rates for the general population (SIRs).

Table 9: Clinical evidence summary: PHPT patients compared to control (no multivariate analysis) – Stratum post-operative

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Monitoring - no MV analysis- strata post-operative (95% CI)
Mortality	668 (1 study)	VERY LOW[a,b] due to risk of bias, imprecision	RR 1.62 (0.96 to 2.73)	Moderate	
				63 per 1000	39 more per 1000 (from 3 fewer to 109 more) PHPT: 34/334 Control: 21/334
mortality[c]	4461 (1 study)	VERY LOW[a] due to risk of bias	Male: RR 1.30 (95% CI 1.07- 1.57) Female: RR 1.61 (95% CI 1.46-1.78)	-	Male: observed: n=107; expected: n=82.2 Female: observed: n=396; expected: n=245.6 Altogether: observed: n-=503; expected: n=327.8
Fracture[d]	90 (1 study)	VERY LOW[a,b] due to risk of bias, imprecision	RR 1.91 (1.06 to 3.47)	Moderate	
				250 per 1000	227 more (from 15 more to 618 more) PHPT: 22/46 Control: 11/44

[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 (used the non-randomised studies checklist on evibase for assessment of 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] The cause of increased risk of death was cardiovascular disease

Men: RR 1.71 (95% CI 1.34-2.15)

Outcomes	No of Participants (studies) Follow up	Quality of the evidence (GRADE)	Relative effect (95% CI)	Anticipated absolute effects	
				Risk with Control	Risk difference with Monitoring - no MV analysis- strata post-operative (95% CI)
Women: RR 1.85 (955 CI 1.62-2.11) [d]92% of the total fractures and all of the post-menopausal fractures occurred after minor trauma.					

Narrative data:

1. Clifton-Bligh 2015 [21]

n=561

448 had surgery and 113 did not have surgery.

There was no significant difference in the relative survival between surgically and non-surgically treated patients over a 10 year period (figures, no data). The average number of years of life lost by hyperparathyroid patients compared to control population was 7.5 years. There was no significant difference in the death rate between those with an initial serum calcium of >3.00 mmol/l compared with those with an initial serum calcium of <3.00 mmol/l (no data).

In a multivariate analysis in the surgically treated group, the serum calcium did not significantly influence survival (HR 1.57, 95% CI 0.30-8.30, p=0.593). In a multivariate analysis, risk factors associated with death in the surgically treated group were diabetes mellitus (HR 4.09, 95% CI 1.42-6.74, P=0.001), congestive cardiac failure (HR 5.46, 95% CI 1.31-22.87, P=0.002), coronary heart disease (HR 2.16, 95% CI 1.08-0.044). The presence of kidney stones before surgery was associated with reduced mortality (HR 0.364, 95% CI 0.22-0.68, P=0.001).

In the non-surgically treated group, death was significantly associated with a high serum PTH (HR 1.59, 95% CI 1.20-2.11, p=0.001), coronary heart disease (HR 3.10, 95% CI 1.42-6.74, P=0.004), and kidney stones (HR 2.48, 95% CI 1.07-5.76, p=0.035). This difference between the surgically treated and non-surgically treated group with respect to the impact of kidney stones is not clear. 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 (95% 0.38-1.18, p=0.167).

Using a 20 year follow-up for the whole group, multivariate analysis did not show any survival difference between male and female, surgery versus non-surgery (p=0.867), serum calcium >3 mmol/l versus <3 mmol/l (p=0.794), or serum PTH analysed as quartiles (no data).

2. Hedback 1992 [38]

Yearly death reduction

Hyperparathyroid population operated on in 1987-94: mean (range)
Male: 17% (95% 7-26)
Female: 8% (2.00- 13)

Swedish population 1974-1983: male: 0.95% (95% CI 0.81-1.09)
Female: 1.68% (1.53- 1.83)

Swedish population 1987-94: male: 1.51% (1.34-1.67)
Female: 0.88% (0.70-1.05)

3. Melton 1992 [49]

Fracture risk (after diagnosis)

Overall: PHPT 50/90; control 52/90; RR 1.0 (95% CI 0.7-1.4)

Calcium ≥2.74 mmol/l: PHPT 34/90; control 24/90; RR 1.4 (95% CI 0.8-2.4)

Calcium <2.74 mmol/l: PHPT 16/90; control 27/90; RR 0.6 (95% CI 0.3-1.1)

Operated on: PHPT 19/90; control 26/90; RR 0.7 (95% CI 0.4-1.3)

Not operated on: PHPT 31/90; control 26/90; RR 1.2 (95% CI 0.7-2.0)

Comorbid conditions: PHPT 44/90; control 38/90; RR 1.2 (95% CI 0.8-1.8)

No comorbid conditions: PHPT 6/90; control 14/90; RR 0.4 (95% CI 0.2-1.1)

Women: 43/90; 42/90; RR 1.0 (95% CI 0.7-1.6)

Men: 7/90; 10/90; RR 0.7 (95% CI 0.3-1.8)

In a multivariate analysis, only age at diagnosis was an independent predictor of fracture risk in PHPT (P<0.2). A 10 year increase in age corresponded to a 36% increase in fracture risk.

4. Ronni-Sivula [60]

Mortality

PHPT patients 34/334; control 21/334; RR 1.62 (0.96 to 2.73)

The mean age of PHPT patients at death was 65 years, 61 years in men and 66 years in women.

The mean age of control patients at death was 67 years, 62 years in men and 69 years in women.

The deceased patients in the PHPT group had a higher mean value of serum calcium pre-operatively than patients in the entire PHPT group (3.31 mmol/l versus 3.08 mmol/l).

In the deceased patients in the PHPT group, serum creatinine was elevated (>115 mmol/l) pre-operatively in 15 (44%) of the deceased patients. In the entire PHPT group serum creatinine was elevated pre-operatively in 57 patients (17%). In the deceased patients serum creatinine was most often elevated in the groups with hypercalcaemic crises (4/6) and cystic bone changes (3/4) and most rarely in the renal stone group (1/5).

PHPT patients who died had more severe form of disease: 55% had hypercalcaemic crises and 24% had cystic bone changes, 4% had renal stones.

Causes of death:

PHPT: n=18 cardiac disease; n=4 cerebrovascular death; n=1 vascular disease; n=4 uraemia; n=2 malignant tumour; n=2 hypercalcaemic crisis; n=3 other causes

Control: n=8 cardiac disease; n=5 malignant tumour; n=8 other causes

1.5 Economic evidence

1.5.1 Included studies

No relevant health economic studies were identified.

1.5.2 Excluded studies

No health economic studies that were relevant to this question were excluded due to assessment of limited applicability or methodological limitations.

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

1.5.3 Unit costs

Unit costs for common clinical tests for monitoring PHPT were presented to the committee for consideration of cost effectiveness.

Table 10: UK costs of monitoring procedures

Monitoring procedure	Cost (average)	Notes	Source
GP consultation	£37.00	Assumed average duration of 9.22 minutes	PSSRU 2017[24]
Blood tests (adjusted serum calcium, serum creatinine, renal function, lipids)	£1.13	Clinical biochemistry test	NHS Reference Costs 16/17[28]
PTH test	£8.00	Average of 12 test costs sought by the committee from laboratories in their local areas.	Committee estimate
Vitamin D	£16.50	Average of two NHS hospitals[(b)]	Filby 2014[36]
DXA scan	£83.00	Performed in an outpatient setting	NHS Reference Costs 16/17[28]
Ultrasound	£52.00	US with duration <20 minutes, without contrast, outpatient setting	NHS Reference Costs 16/17[28]
X-ray	£30.00	Direct access plain film	NHS Reference Costs 16/17[28]
Blood pressure	£6.00	Calculation based on average cost of 15min contact with community or hospital based nurse	PSSRU 2017[24]
ECG	£ 37.00	ECG not included in 16/17 NHS reference costs. The committee did not consider it is likely to have changed significantly in cost since 2010/11.	NHS Reference costs 10/11[29].

Table 11: UK cost of clinical events associated with PHPT

Event	Cost	Notes	Source
Cardiovascular events			

Event	Cost	Notes	Source
Acute coronary syndrome (including myocardial infarction)	£4,933	Cost of initial 6 months	NICE Hypertension Guideline (update 2011)[52]
Stroke	£10,190		
Heart failure	£2,649		
Renal events			
Renal dysfunction	£803	General renal disorders without interventions, national average, day case, CC score 0–2	NHS Reference costs 16/17[28]
Renal stones – shockwave therapy	£452	Day case	NHS Reference costs 16/17[28]
Renal stones – ureteroscopy	£2,172	50% elective weighted average, and 50% day case weighted average to reflect UK practice	
Renal stones – percutaneous nephrolithotomy	£5,195	Elective weighted average	
Fragility fractures[(a)]			
Hip	£15,722	Total cost of initial hospitalisation and costs in year following fracture.	NICE TA464[54]
Vertebrae	£8,019		
Proximal humerus	£6,625		
Wrist	£4,523		

1.6 Resource costs

The recommendations made by the committee based on this review are not expected to have a substantial impact on resources.

1.7 Evidence statements

1.7.1 Clinical evidence statements

1.7.1.1 Monitoring

No evidence was identified for assessing the optimum type and frequency of monitoring for people with PHPT.

1.7.1.2 Monitoring long term outcomes

1.7.1.2.1 PHPT cases versus matched comparators (adjusted for key confounders) in non-surgical stratum

Evidence from one study (n=8544, median follow up 2.9 years, Low quality) suggested that there was increased risk of all-cause mortality, fatal CVD, non-fatal CVD, hypertension, cerebrovascular disease, renal stones, renal failure, all fractures and osteoporotic fracture associated with PHPT (all patients diagnosed but untreated, mild asymptomatic PHPT).

No evidence was identified for the outcomes of renal tract calcification, pancreatitis, myocardial infarction, number of people meeting the criteria for surgery, serum calcium (>2.85 mmol/litre), 24-hour urine for calcium (>10 mmol/dl) and BMD of proximal femur.

1.7.1.2.2 PHPT cases versus controls (no multivariate analysis) in pre-surgery stratum

Evidence from one study (n=771, Very Low quality) suggested that there was increased risk of renal stones associated with PHPT.

No evidence was identified for the outcomes of mortality, fragility fracture, renal tract calcification, pancreatitis, stroke, hypertension, myocardial infarction, number of people meeting the criteria for surgery, serum calcium (>2.85 mmol/litre), 24-hour urine for calcium (>10 mmol/dl) and BMD of proximal femur.

1.7.1.2.3 PHPT cases versus controls (no multivariate analysis) in mixed pre and post-operative stratum

Evidence for fractures (all) was inconsistent. One study (n=180, Very Low quality) suggested that there was no difference between PHPT and controls for fractures; another study (n=407, Very Low quality) suggested that there was increased risk of fractures associated with PHPT.

Evidence from one study (n=114, Very Low quality) suggested there was a reduced risk of non-vertebral fractures associated with mild PHPT. Evidence from two studies (n=114 for mild PHPT; n=162 for non-mild PHPT, Very Low quality) suggested that there was no difference between PHPT and controls for the outcome lumbar spine BMD. Evidence from one study (n=114, Very Low quality) suggested that there was no difference between PHPT and controls for the outcomes femoral neck BMD, total femur BMD, and non-vertebral fractures associated with non-mild PHPT.

Evidence from 3 studies (n=343; n=407; n=162, Very Low quality) suggested that there was increased risk of vertebral fractures associated with PHPT (both mild and non-mild).

Evidence from one study (n=1373 women; n=551 men, Very Low quality) suggested that there was no difference between PHPT and control for the outcome hip fractures in both men and women.

No evidence was identified for the outcomes of mortality, renal stones, renal tract calcification, pancreatitis, stroke, hypertension, myocardial infarction, number of people meeting the criteria for surgery, serum calcium (>2.85 mmol/litre) and 24-hour urine for calcium (>10 mmol/dl).

1.7.1.2.4 PHPT cases versus controls (no multivariate analysis) in post-operative stratum

The evidence for the outcome mortality was not consistent. One study (n=668, Very Low quality) suggested that there was no difference between those PHPT patients who had surgery and healthy controls; another study (n=4461, Very Low quality) suggested that there was increased mortality in surgery patients compared to healthy controls.

Evidence from one study (n=90, Very Low quality) suggested that there was increased risk of fracture in PHPT patients compared to healthy controls.

No evidence was identified for the outcomes of renal stones, renal tract calcification, pancreatitis, stroke, hypertension, myocardial infarction, number of people meeting the criteria for surgery, serum calcium (>2.85 mmol/litre), 24-hour urine for calcium (>10 mmol/dl) and BMD of proximal femur.

1.7.2 Health economic evidence statements

No relevant economic evaluations were identified.

1.8 The committee's discussion of the evidence

1.8.1 Interpreting the evidence

1.8.1.1 The outcomes that matter most

For the monitoring review the committee considered the outcomes of health-related quality of life and mortality as critical outcomes for decision making. Other important outcomes included deterioration in renal function, fractures (vertebral or long bone), occurrence of kidney stones, persistent hypercalcaemia, bone mass density (BMD) of the distal radius or the lumbar spine, cardiovascular events, adverse events (to include voice change, hypoparathyroidism, hypothyroidism/hyperthyroidism), cancer incidence and re-operation (for post-surgery stratum). No evidence was available for this review.

For the monitoring long-term outcomes review the committee considered the outcomes of mortality, fragility fracture, renal stones, renal tract calcification, pancreatitis, stroke, hypertension, myocardial infarction, number of people who become eligible for surgery/meet the criteria for surgery as critical outcomes for decision making. Other important outcomes included serum calcium (>2.85 mmol/litre), 24-hour urine for calcium (>10 mmol/dl) and BMD of proximal femur (T-score <2.5; Z score <2).

No evidence was identified for the outcomes of pancreatitis, myocardial infarction, number of people meeting the criteria for surgery, serum calcium (>2.85 mmol/litre) and 24-hour urine for calcium (>10 mmol/dl).

1.8.1.2 The quality of the evidence

All the evidence in this review included the incidence of outcomes in people with primary hyperparathyroidism compared with healthy controls.

There were 11 observational studies, of which one was a prospective cohort study and the rest were retrospective cohort studies. The majority of evidence was of Very Low quality due to risk of bias and also, in most cases, imprecision. Only one prospective cohort study adjusted for all key confounders; in five studies the control group was matched for factors such as age and gender but these studies did not adjust for serum calcium level and absence/presence of end organ effects (no multivariate analysis conducted) and the remaining five studies did not have matched controls or adjust for any key confounders (no multivariate analysis conducted). These limitations were taken into account by the committee when interpreting the evidence.

1.8.1.3 Benefits and harms

The studies in this review were stratified as non-surgical, pre-operative, mixed pre- and post-operative and post-operative. No evidence was available for people on calcimimetics, bisphosphonates and normocalcaemic patients.

Non-surgical: The evidence for this group suggested that compared to healthy controls there was increased risk of all-cause mortality, fatal and nonfatal cardiovascular disease (CVD), hypertension, cerebrovascular disease, all fractures, osteoporotic fractures, renal failure and renal stones in patients with asymptomatic primary hyperparathyroidism, with the risk of renal failure and renal stones being the highest.

Pre-operative: There was evidence from only one study for this stratum. The evidence suggested that there was an increased risk of renal stones in people with primary hyperparathyroidism compared to healthy controls.

Mixed pre and post-operative: The studies in this strata included both mild and non-mild primary hyperparathyroidism patients. Data were not available separately for surgical and non-surgical/pre-surgery patients. However, for some outcomes the evidence was reported separately for mild and non-mild patients. The evidence suggested that there was increased risk of vertebral fractures in both mild and non-mild primary hyperparathyroidism patients compared to healthy controls and there was no difference between the primary hyperparathyroidism and healthy controls for lumbar spine BMD, femoral neck BMD and total femur BMD for both mild and non-mild patients. The evidence for all fractures was not consistent; one study suggested that there was no difference and one study suggested that there was increased risk in primary hyperparathyroidism patients compared to healthy controls. The evidence suggested there was a reduced risk for non-vertebral fractures in mild primary hyperparathyroidism patients compared to healthy controls and there was no difference in non-vertebral fractures between non-mild primary hyperparathyroidism patients and healthy controls. The evidence also suggested that there was no difference between the primary hyperparathyroidism and healthy controls for the outcome hip fractures in both men and women.

Post-operative: The evidence for the outcome mortality was not consistent; one study suggested that there was no difference between those primary hyperparathyroidism patients who had surgery and healthy controls and one study suggested that there was increased mortality in surgery patients compared to healthy controls and this was mainly attributed to cardiovascular disease. The evidence also suggested that there was increased risk of fracture in primary hyperparathyroidism patients compared to healthy controls.

Overall the evidence suggested that there was increased risk of mortality, fractures, renal stones, renal failure, cardiovascular disease, low bone density and hypertension associated with untreated asymptomatic primary hyperparathyroidism patients. Due to the low quality of the evidence, the committee also took their clinical experiences into account when making their recommendations.

Based on their experience, the committee agreed that all patients diagnosed with primary hyperparathyroidism will need baseline assessment of their symptoms, BMD by DXA scan, and ultrasound of the renal tract to help determine the optimal management pathway. The committee considered that monitoring serum calcium level and symptoms of hypercalcaemia would support discussion of the most appropriate treatment strategy including surgery. Ultrasound of the kidneys would help in identifying cause for specific interventions or appropriate referral, and DXA scan would help in assessing fracture risk and/or the need for bisphosphonates.

The committee discussed the increased risk of mortality due to cardiovascular causes both before and after parathyroidectomy and hence felt that there is a need for monitoring cardiovascular risk in this group of patients. The committee also discussed the increased risk of renal stones and fractures in people with primary hyperparathyroidism (both before and after surgery) and therefore agreed that these people need to be monitored accordingly and consideration given to adjunctive treatments.

Based on their knowledge and experience, the committee agreed that people who have had parathyroid surgery can be considered biochemically cured if their albumin-adjusted serum calcium level is within the reference range 3 to 6 months after surgery. The committee considered that the risk of recurrent disease following successful removal of a solitary adenoma is very low and that, after the 6-month check, it is sufficient for calcium to be checked as part of routine blood testing to a maximum of once a year. The committee highlighted that for people with multigland disease there is a higher risk of recurrence than in those who had a single adenoma and in monitoring of such patients specialist opinion should be sought. However, the committee noted that the risk is still very low if the person has normal adjusted calcium at 3 to 6 months after surgery. The committee agreed that for people with osteoporosis, although bone density improves after surgery, skeletal recovery

can take some time and would need specialist monitoring. Based on their experience, the committee discussed that risk of kidney stones decreases after successful surgery, but the residual risk persists and hence specialist opinion should be sought for monitoring of such patients.

The committee noted that in patients with multigland disease, a specialist will be aware of associated syndromes (for example multiple endocrine neoplasia type 1 [MEN1], MEN2A, familial isolated hyperparathyroidism, autosomal dominant mild hyperparathyroidism, familial hypocalciuric hypercalcaemia), and hence would be in a better position to make individualised assessment and determine the frequency of monitoring. The committee stated that for those patients with multigland disease discharged back to primary care, serum calcium tests will need to be conducted annually as part of their routine biochemical testing. Current practice is to conduct biochemical tests annually if there is no end organ damage.

The committee stated that in patients with genetic diseases such as MEN-1, primary hyperparathyroidism could be the first presentation so early detection of the disease, correct treatment, and continued care are of great importance; but noted that such cases are infrequent and beyond the scope of this guideline.

The committee discussed from clinical experience that there are no clinical factors that would predict the prognosis of patients with asymptomatic primary hyperparathyroidism. Evidence from the review suggested that around 35% of asymptomatic patients develop indications for surgery during follow-up. Hence the committee agreed that long-term medical monitoring for asymptomatic patients was essential to assess progression to meeting eligibility criteria for surgery and/or any evidence of end organ damage. The committee recommended the following monitoring strategies including assessment of symptoms and comorbidities annually; annual measurement of serum calcium test, estimated glomerular filtration rate (eGFR) or serum creatinine test; DXA scan every 2 to 3 years; ultrasound of the renal tract if renal stones are suspected (see NICE's guideline on renal and ureteric stones). The committee recognised that measurement of renal function is important in assessing calcium and PTH levels. Most patients will have eGFR measured with serum calcium. An elevated serum calcium should be investigated irrespective of eGFR and the proposed algorithms are designed to ensure that if eGFR has not been checked early in the diagnostic odyssey, it is done so as part of the investigation and assessment of patients with hypercalcaemia. Such monitoring may detect clinically relevant changes that may necessitate reconsideration of surgery and/or adjunctive medical therapies. The committee discussed that assessment of symptoms will be annually or when the patients presents with any of the symptoms of primary hyperparathyroidism such as fatigue, depression, abdominal pain, constipation, muscle weakness, loss of concentration, mild confusion etc. The committee discussed that for suspected renal stones patients could present with colic/severe pain, asymptomatic haematuria, passing grit, discomfort etc. Based on the evidence and their experience, the committee agreed that there was an increased risk of fracture associated with primary hyperparathyroidism and hence agreed that DXA scan should be done every 2–3 years in these patients.

The committee agreed that these recommendations could also apply to those people who have refused surgery and in people after failed primary surgery to assess progression of disease in these patients.

The committee also discussed the current National Institutes of Health (NIH) criteria[16] for monitoring in patients with asymptomatic primary hyperparathyroidism who do not undergo parathyroid surgery (2013). The current NIH criteria include the following monitoring strategies: serum calcium annually; skeletal – every 1–2 years (3 sites), X-ray or VFA of spine if clinically indicated (for example height loss, back pain); renal ─ eGFR, annually; serum creatinine, annually. If renal stones are suspected: 24-hour biochemical stone profile, renal imaging by X-ray, ultrasound, or CT.

The committee noted that there are no established guidelines/definitions of cure for primary hyperparathyroidism. The committee from their experience discussed that patients are considered to be biochemically cured if their PTH is in the reference range immediately following surgery and their serum calcium is within the reference range 3–6 months after surgery. The committee stated that post-operative PTH would still be performed if an intraoperative PTH was taken as a very small proportion of patients would show a change from intraoperative PTH level. Overall the committee did not think that a PTH test at 3–6 months would offer any additional clinical value. The committee noted that persistently high calcium at 3–6 months would trigger testing of plasma PTH (as per the recommendations on diagnosis). The committee considered that a 3–6 month post-operative calcium test could be done in secondary care.

1.8.2 Cost effectiveness and resource use

No previously published economic evaluations were identified for the cost-effectiveness of monitoring people with primary hyperparathyroidism. Unit costs of monitoring procedures were presented to the committee for consideration. Costs of clinical events associated with primary hyperparathyroidism – including cardiovascular events, renal events and fragility fractures – were also presented to provide a more comprehensive picture of potential healthcare resource use of primary hyperparathyroidism if the condition is left unchecked. However, as there is no clinical evidence for the extent to which monitoring will prevent such events, cost effectiveness of monitoring could not be evaluated and therefore is highly uncertain.

For people who have had parathyroid surgery, the committee noted that a PTH test immediately following surgery provides a timely indication of whether a patient has been cured of primary hyperparathyroidism due to the short half-life of PTH compared to calcium in the blood and is the most clinically relevant indication of cure. The committee also highlighted that further confirmation of cure at 3 to 6 months is necessary to assess recurrent disease. However, they agreed that this can be achieved with a lower cost test for albumin-adjusted serum calcium rather than a repeated PTH test.

The committee discussed that those with successful parathyroid surgery are generally considered to return to general population risk levels for end organ disease such as renal stones and fractures and therefore do not require further monitoring. However, the committee considered that there may be cases where specialist endocrine opinion should be sought with regards to monitoring due to more complex issues such as multi-gland disease and recurrent disease, or due to comorbidities such as osteoporosis and renal stones. In these cases decisions on monitoring should be made on a case-by-case basis.

In addition, the committee considered that in the event that people who have been cured after parathyroid surgery have a routine blood test for another cause, incidental testing for serum calcium as part of these blood tests could be cost effective. The committee discussed that there is minimal added expense to such testing as this does not require additional time in taking blood, only in analysing the sample (estimated around an additional £0.30 to additionally analyse calcium). The committee highlighted that such incidental testing should be limited to once a year to avoid unnecessary testing for those who may have frequent routine blood tests. The committee discussed that incidental calcium testing could help identify recurrent disease prior to the onset of symptoms or potential consequential end organ damage as a result of hypercalcaemia and therefore avoid potential decrements in quality of life and associated costs of such events. The committee was of the consensus that this practice could therefore be cost effective. However, as there is no clinical evidence available to assess this, this is highly uncertain.

For patients who are either not eligible for surgery, or have chosen not to undergo surgery, consensus from the committee was that monitoring should occur. However, the committee noted that there is some variation in current practice with respect to some of the items tested

as part of the monitoring regime. While most practitioners adopt a fairly standard practice of including tests as specified in the NIH guidance – including annual tests for serum calcium and serum creatinine – some practitioners also test for PTH as part of routine practice. The committee indicated that testing for serum calcium and serum creatinine as part of routine monitoring is sufficient to detect any signs of change in a patient's condition, and are also less costly. It was noted that healthcare providers should use their judgement in determining whether a patient will require a further PTH test based on the results of their tests for serum calcium. Hence, routine testing of PTH levels has not been recommended. This is a potential area for some cost savings.

The committee considered that these recommendations are generally in line with current practice and therefore are not expected to have a significant impact on healthcare resource use.

1.8.3 Other factors the committee took into account

The committee noted that the pre-operative population awaiting surgery are not considered to be in a monitoring setting (see recommendations on surgery).

The committee was aware of two studies [61, 76] assessing long-term outcomes in patients with and without parathyroid surgery which were included in the indications for surgery evidence review.. The study [76] reported that the risks of mortality, fractures and gastric ulcers were lower in patients treated surgically than those treated conservatively. However there was a higher risk of kidney or urinary tract stones in patients treated surgically than patients treated conservatively. Another study [61] was a long-term prospective cohort study of asymptomatic primary hyperparathyroidism patients. The study reported that at 10 years, 25% of the asymptomatic primary hyperparathyroidism patients did show evidence of progressive disease with worsening hypercalcemia, hypercalciuria, and reductions in BMD being the most common complications. The study reported that 37% of asymptomatic patients developed new surgical criteria at any time point over the 15 years of observation. Meeting surgical criteria at study baseline did not predict who would have progressive disease. BMD did not change at any measurement site during the first 8 years of follow-up in the asymptomatic patients. The lumbar spine BMD was stable for the entire 15 years of follow-up. Overall, 59% of the asymptomatic patients had more than a 10% decline in BMD at one or more sites over the 15-year period. The study also reported that 15% of the patients who underwent surgery were symptomatic with kidney stones. At 15 years, serum calcium, PTH, and urinary calcium excretion were all significantly lower in comparison with the individual subjects' baseline values and all well within normal limits. Post-operative increases in BMD were sustained with BMD remaining significantly above baseline for the entire 15 years of follow-up at all three skeletal sites. The committee noted that the findings of these studies were consistent with their clinical experience.

References

1. Abdulkader R, Dharmapalaiah C, Stephenson S, Clunie G. The incidence of previously undiagnosed conditions in patients attending fracture liaison service. Annals of the Rheumatic Disease. 2012; 71(Suppl 3):AB1040

2. Agarwal A, George RK, Gupta SK, Mishra SK. Pancreatitis in patients with primary hyperparathyroidism. Indian Journal of Gastroenterology. 2003; 22(6):224-5

3. Ahsan T, Erum U, Inam Pal KM, Jabeen R, Qureeshi SG, Rehman UL et al. The many guises of primary hyperparathyroidism: An unchanged scenario. Journal of the Pakistan Medical Association. 2017; 67(4):580-5

4. 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

5. Amaral LM, Queiroz DC, Marques TF, Mendes M, Bandeira F. Normocalcemic versus hypercalcemic primary hyperparathyroidism: More stone than bone? Journal of Osteoporosis. 2012; 2012:128352

6. Antonelli R, Falcone S, Scillitani A, Salcuni AS, Carnevale V, Battista C et al. Normocalcemic hyperparathyroidism: Studies on bone loss over a ten-year follow-up time. Endocrine Reviews. 2011; 32(3 Suppl):P3-105

7. Babey M, Arampatzis S, Popp A, Schuematschek-Kainth J, Kopp PA, Lippuner K. Normocalcemic primary hyperparathyroidism in patients with low bone mass: Biochemical and clinical characteristics. Journal of Bone and Mineral Research. 2010; 25(S1):S259

8. Bai HX, Giefer M, Patel M, Orabi AI, Husain SZ. The association of primary hyperparathyroidism with pancreatitis. Journal of Clinical Gastroenterology. 2012; 46(8):656-61

9. Bailey RR, Dann E, Greenslade NF, Little PJ, McRae CU, Utley WL. Urinary stones: a prospective study of 350 patients. New Zealand Medical Journal. 1974; 79(516):961-5

10. Bandeira F, Griz LH, Bandeira C, Pinho J, Lucena CS, Alencar C et al. Prevalence of cortical osteoporosis in mild and severe primary hyperparathyroidism and its relationship with bone markers and vitamin D status. Journal of Clinical Densitometry. 2009; 12(2):195-9

11. Bandeira L, Cozadd D, Bucovsky M, McMahon DJ, Lee JA, Silverberg SJ et al. Occult nephrolithiasis in primary hyperparathyroidism. Endocrine Reviews. 2016; 37(2 Suppl 1):FRI-333

12. Bao L, Li Y, Lin H. Effect of parathyroidectomy and pharmacotherapy in primary hyperthyroidism on bone metabolism. Osteoporosis International. 2013; 24(Suppl 1):S125

13. Battersby C, Burnett W, Winch J. Pancreatitis associated with hyperparathyroidism. Medical Journal of Australia. 1969; 2(25):1268-70

14. Beard DE, Goodyear WE. Hyperparathyroidism and urolithiasis. Journal of Urology. 1950; 64(5):638-42

15. Bhadada SK, Arya AK, Mukhopadhyay S, Khadgawat R, Sukumar S, Lodha S et al. Primary hyperparathyroidism: insights from the Indian PHPT registry. Journal of Bone and Mineral Metabolism. 2018; 36(2):238-45

16. 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

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

18. Cannon J, Lew JI, Solorzano CC. Parathyroidectomy for hypercalcemic crisis: 40 years' experience and long-term outcomes. Surgery. 2010; 148(4):807-12; discussion 812-3

19. Carnaille B, Oudar C, Pattou F, Combemale F, Rocha J, Proye C. Pancreatitis and primary hyperparathyroidism: forty cases. Australian and New Zealand Journal of Surgery. 1998; 68(2):117-9

20. Cassibba S, Pellegrino M, Gianotti L, Baffoni C, Baralis E, Attanasio R et al. Silent renal stones in primary hyperparathyroidism: prevalence and clinical features. Endocrine Practice. 2014; 20(11):1137-42

21. 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

22. Corlew DS, Bryda SL, Bradley EL, 3rd, DiGirolamo M. Observations on the course of untreated primary hyperparathyroidism. Surgery. 1985; 98(6):1064-71

23. Csupor E, Toth E, Meszaros S, Ferencz V, Szucs J, Lakatos P et al. Is there any connection between the presence of kidney stones in primary hyperparathyroidism and the location of an underlying adenoma? Experimental and Clinical Endocrinology and Diabetes. 2005; 113(5):257-61

24. Curtis L, Burns A. Unit costs of health and social care 2017. Canterbury. Personal Social Services Research Unit University of Kent, 2017. Available from: https://www.pssru.ac.uk/project-pages/unit-costs/unit-costs-2017/

25. Danzi JT, Farmer RG, Esselstyn CB, Jr. Recurrent pancreatitis associated with normocalcemia, parathyroid hyperplasia, and increased serum parathormone. Cleveland Clinic Quarterly. 1974; 41(1):39-43

26. De Geronimo S, Romagnoli E, Diacinti D, D'Erasmo E, Minisola S. The risk of fractures in postmenopausal women with primary hyperparathyroidism. European Journal of Endocrinology. 2006; 155(3):415-20

27. Deaconson TF, Wilson SD, Lemann J, Jr. The effect of parathyroidectomy on the recurrence of nephrolithiasis. Surgery. 1987; 102(6):910-3

28. Department of Health. NHS reference costs 2016/2017. Available from: https://improvement.nhs.uk/resources/reference-costs/ Last accessed: 17/01/2018.

29. Department of Health. NHS reference costs 2010-11. 2012. Available from: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140 Last accessed: 05/10/2018.

30. Diaz De La Guardia FV, Martin MA, Arrabal Polo MA, Flores SQ, Ortiz JLM, Gomez AZ. Renal lithiasis in patients with primary hyperparathyroidism. Evolution and treatment. Archivos Españoles de Urología. 2010; 63(1):32-40

31. Dimkovic NB, Wallele AA, Oreopoulos DG. Renal stone disease, elevated iPTH level and normocalcemia. International Urology and Nephrology. 2002; 34(1):135-41

32. Dolgin C, Lo Gerfo P, LiVolsi V, Feind C. Twenty-five year experience with primary hyperparathyroidism at Columbia Presbyterian Medical Center. Head and Neck Surgery. 1979; 2(2):92-8

33. Dumitrescu B, van Helden S, ten Broeke R, Nieuwenhuijzen-Kruseman A, Wyers C, Udrea G et al. Evaluation of patients with a recent clinical fracture and osteoporosis, a multidisciplinary approach. BMC Musculoskeletal Disorders. 2008; 9:109

34. Eufrazino C, Veras A, Bandeira F. Epidemiology of primary hyperparathyroidism and its non-classical manifestations in the city of Recife, Brazil. Clinical Medicine Insights: Endocrinology and Diabetes. 2013; 6:69-74

35. Falko JM, Maeder MC, Conway C, Mazzaferri EL, Skillman TG. Primary hyperparathyroidism: Analysis of 220 patients with special emphasis on familial hypocalciuric hypercalcemia. Heart and Lung. 1984; 13(2):124-31

36. Filby A, Lewis L, Taylor M. An economic evaluation of interventions to improve the uptake of vitamin D supplements in England and Wales. London. National Institute for Health and Care Excellence, 2014. Available from: https://www.nice.org.uk/guidance/ph56/documents/economic-evaluation-report2

37. Heath H, 3rd. Clinical spectrum of primary hyperparathyroidism: evolution with changes in medical practice and technology. Journal of Bone and Mineral Research. 1991; 6 (Suppl 2):S63-70; discussion S83-4

38. Hedback G, Oden A. Increased risk of death from primary hyperparathyroidism--an update. European Journal of Clinical Investigation. 1998; 28(4):271-6

39. Hedback GM, Oden AS. Cardiovascular disease, hypertension and renal function in primary hyperparathyroidism. Journal of Internal Medicine. 2002; 251(6):476-83

40. Jha S, Jayaraman M, Jha A, Jha R, Modi KD, Kelwadee JV. Primary hyperparathyroidism: A changing scenario in India. Indian Journal of Endocrinology and Metabolism. 2016; 20(1):80-3

41. Kenny AM, MacGillivray DC, Pilbeam CC, Crombie HD, Raisz LG. Fracture incidence in postmenopausal women with primary hyperparathyroidism. Surgery. 1995; 118(1):109-14

42. 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

43. Kobayashi T, Sugimoto T, Chihara K. Clinical and biochemical presentation of primary hyperparathyroidism in Kansai district of Japan. Endocrine Journal. 1997; 44(4):595-601

44. Larsson K, Lindh E, Lind L, Persson I, Ljunghall S. Increased fracture risk in hypercalcemia. Bone mineral content measured in hyperparathyroidism. Acta Orthopaedica Scandinavica. 1989; 60(3):268-70

45. 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

46. Lowe H, McMahon DJ, Rubin MR, Bilezikian JP, Silverberg SJ. Normocalcemic primary hyperparathyroidism: further characterization of a new clinical phenotype. Journal of Clinical Endocrinology and Metabolism. 2007; 92(8):3001-5

47. Lueg MC. Hypertension and primary hyperparathyroidism: a five-year case review. Southern Medical Journal. 1982; 75(11):1371-4

48. Marques TF, Vasconcelos R, Diniz E, Rego D, Griz L, Bandeira F. Normocalcemic primary hyperparathyroidism in clinical practice: an indolent condition or a silent threat? Arquivos Brasileiros de Endocrinologia e Metabologia. 2011; 55(5):314-7

49. 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

50. Misiorowski W, Zgliczyski W. Prevalence of primary hyperparathyroidism among patients with low bone mass. Advances in Medical Sciences. 2012; 57(2):308-13

51. Mollerup CL, Lindewald H. Renal stones and primary hyperparathyroidism: natural history of renal stone disease after successful parathyroidectomy. World Journal of Surgery. 1999; 23(2):173-5; discussion 176

52. National Clinical Guideline Centre. Hypertension: the clinical managment of primary hypertension in adults: update of clinical guidelines 18 and 34. NICE clinical guideline 127. London. National Clinical Guideline Centre, 2011. Available from: http://guidance.nice.org.uk/CG127

53. 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

54. National Institute for Health and Care Excellence. Bisphosphonates for treating osteoporosis. NICE technology appraisal guidance 464. London. National Institute for Health and Care Excellence, 2017. Available from: http://guidance.nice.org.uk/TA464

55. Nilsson IL, Aberg J, Rastad J, Lind L. Maintained normalization of cardiovascular dysfunction 5 years after parathyroidectomy in primary hyperparathyroidism. Surgery. 2005; 137(6):632-8

56. Pradeep PV, Mishra A, Agarwal G, Agarwal A, Verma AK, Mishra SK. Long-term outcome after parathyroidectomy in patients with advanced primary hyperparathyroidism and associated vitamin D deficiency. World Journal of Surgery. 2008; 32(5):829-35

57. Pratley SK, Posen S, Reeve TS. Primary hyperparathyroidism. Experiences with 60 patients. Medical Journal of Australia. 1973; 1(9):421-6

58. Purnell DC, Smith LH, Scholz DA, Elveback LR, Arnaud CD. Primary hyperparathyroidism: a prospective clinical study. American Journal of Medicine. 1971; 50(5):670-8

59. Rajeevan T, Cunning C, Abdulla A. 26 management of primary hyperparathyroidism (PHPT) in older people: A series review. Age and Ageing. 2014; 43(Suppl_1):i6-i6

60. Ronni-Sivula H. Causes of death in patients previously operated on for primary hyperparathyroidism. Annales Chirurgiae et Gynaecologiae. 1985; 74(1):13-8

61. 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

62. Scholz DA, Purnell DC. Asymptomatic primary hyperparathyroidism. 10-year prospective study. Mayo Clinic Proceedings. 1981; 56(8):473-8

63. Siilin H, Lundgren E, Mallmin H, Mellstrom D, Ohlsson C, Karlsson M et al. Prevalence of primary hyperparathyroidism and impact on bone mineral density in elderly men: MrOs Sweden. World Journal of Surgery. 2011; 35(6):1266-72

64. Silverberg SJ, Gartenberg F, Jacobs TP, Shane E, Siris E, Staron RB et al. Longitudinal measurements of bone density and biochemical indices in untreated primary hyperparathyroidism. Journal of Clinical Endocrinology and Metabolism. 1995; 80(3):723-8

65. Silverberg SJ, Shane E, Jacobs TP, Siris ES, Gartenberg F, Seldin D et al. Nephrolithiasis and bone involvement in primary hyperparathyroidism. American Journal of Medicine. 1990; 89(3):327-34

66. Siminovitch JM, James RE, Esselstyn CB, Jr., Straffon RA, Banowsky LH. The effect of parathyroidectomy in patients with normocalcemic calcium stones. Journal of Urology. 1980; 123(3):335-7

67. 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

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

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

70. Turchi JJ, Flandreau RH, Forte AL, French GN, Ludwig GD. Hyperparathyroidism and pancreatitis. JAMA. 1962; 180(10):799-804

71. 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

72. 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

73. Vestergaard P, Mollerup CL, Frokjaer VG, Christiansen P, Blichert-Toft M, Mosekilde L. Cohort study of risk of fracture before and after surgery for primary hyperparathyroidism. BMJ. 2000; 321(7261):598-602

74. Vestergaard P, Mollerup CL, Frokjaer VG, Christiansen P, Blichert-Toft M, Mosekilde L. Cardiovascular events before and after surgery for primary hyperparathyroidism. World Journal of Surgery. 2003; 27(2):216-22

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

76. 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

77. Vestergaard P, Mosekilde L. Parathyroid surgery is associated with a decreased risk of hip and upper arm fractures in primary hyperparathyroidism: a controlled cohort study. Journal of Internal Medicine. 2004; 255(1):108-14

78. 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

79. Wilson RJ, Rao S, Ellis B, Kleerekoper M, Parfitt AM. Mild asymptomatic primary hyperparathyroidism is not a risk factor for vertebral fractures. Annals of Internal Medicine. 1988; 109(12):959-62

80. Yu N, Donnan PT, Leese GP. A record linkage study of outcomes in patients with mild primary hyperparathyroidism: the Parathyroid Epidemiology and Audit Research Study (PEARS). Clinical Endocrinology. 2011; 75(2):169-76

81. Yu N, Donnan PT, Murphy MJ, Leese GP. Epidemiology of primary hyperparathyroidism in Tayside, Scotland, UK. Clinical Endocrinology. 2009; 71(4):485-93

82. Yu N, Leese GP, Donnan PT. What predicts adverse outcomes in untreated primary hyperparathyroidism? the Parathyroid Epidemiology and Audit Research Study (PEARS). Clinical Endocrinology. 2013; 79(1):27-34

83. Yu N, Leese GP, Smith D, Donnan PT. The natural history of treated and untreated primary hyperparathyroidism: The parathyroid epidemiology and audit research study. QJM: An International Journal of Medicine. 2011; 104(6):513-21

Appendices

Appendix A: Review protocols

Table 12: Review protocol: Monitoring strategies

Field	Content
Review question	6.1 What is the optimum type and frequency of monitoring for people with PHPT (for example, pre-operative, postoperative, non-surgical)?
Type of review question	Intervention The optimum use of monitoring for people with PHPT – looking at patient outcomes for different strategies or frequencies of monitoring. This will cover pre-operative, post-operative and non-surgical monitoring.
Objective of the review	To determine the clinical and cost effectiveness of different strategies or frequencies of monitoring, at the pre-operative and postoperative stage as well as in people with PHPT not undergoing surgery.
Eligibility criteria – population	Adults (18 years or over) with primary hyperparathyroidism Strata: • Pre-operative • Post-operative • Non-surgical • Previous surgery • Pregnant women Exclude people: • with secondary and tertiary HPT • with multiple endocrine neoplasia • with familial hyperparathyroidism • with parathyroid carcinoma • on medications interfering with calcium metabolism (lithium). 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)	Different techniques/ tools/ variables being monitored: • Blood tests (adjusted serum calcium, serum creatinine) • Imaging (DXA for bone disease) (±VFA) • US for kidney stones, renal tract calcification • X-ray for fragility fracture for vertebral fracture • Other (24-hour urinary calcium, creatinine clearance • CV variables (BP, lipids, ECG) • Vitamin D (for post-operative monitoring) • Renal function Different frequencies of monitoring (acceptable frequencies – this may differ for the different tests – e.g. bloods 6–12 months, DXA 2–3 years) Different durations of monitoring (for post-surgery stratum, e.g. for 3 months or 6 months – to include the optimum timing of serum calcium assessment post-surgery to determine cure)
Eligibility criteria	• Comparing types of strategies to each other

– comparator(s)	• Comparing different frequencies of the same strategy
Outcomes and prioritisation	**Report all outcomes separately for <6 months and ≥6 months** **Critical outcomes:** • HRQOL (continuous outcome) • Mortality (dichotomous outcome) **Important outcomes:** • Deterioration in renal function (continuous – study may also report renal replacement) • Fractures (vertebral or long bone) (dichotomous outcome) • Occurrence of kidney stones (dichotomous outcome) • Persistent hypercalcaemia (dichotomous outcome) • BMD of the distal radius or the lumbar spine (continuous) • Cardiovascular events (dichotomous outcome) • Adverse events (to include voice change, hypoparathyroidism, hypothyroidism/hyperthyroidism; dichotomous outcome) • Cancer incidence (dichotomous outcome) • Reoperation (for post-surgery stratum)
Eligibility criteria – study design	RCTs and systematic reviews of RCTs In the absence of RCT evidence for the critical outcomes, NRSs will be included (only if the following key confounders are matched for or adjusted for in the analysis) Key confounders: • Age • Absence/presence of end-organ effects • Adjusted serum calcium level
Other inclusion exclusion criteria	• Non-English language articles • Conference abstracts
Proposed sensitivity / subgroup analysis, or meta-regression	N/A
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). • GRADEpro was used to assess the quality of evidence for each outcome. • Endnote for bibliography, citations, sifting and reference management. • 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 Date: all years Health economics search databases to be used: Medline, Embase, NHSEED, HTA

	Date: Medline, Embase from 2002 NHSEED, HTA – all years Language: Restrict to English only Supplementary search techniques: backward citation searching Key papers: Not known
Identify if an update	N/A
Author contacts	https://www.nice.org.uk/guidance/indevelopment/gid-ng10051
Highlight if amendment to previous protocol	N/A
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 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.

	Staff from the 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	The NGC is funded by NICE and hosted by the Royal College of Physicians.
Name of sponsor	The NGC is funded by NICE and hosted by the Royal College of Physicians.
Roles of sponsor	NICE funds the NGC to develop guidelines for those working in the NHS, public health and social care in England.
PROSPERO registration number	Not registered

Table 13: Monitoring long term outcomes protocol

Field	Content
Review question	What are the long-term outcomes in people with PHPT?
Type of review question	Prognostic
Objective of the review	To determine whether PHPT is associated with poor long-term outcomes to determine what monitoring they need to undergo
Eligibility criteria – population / disease / condition / issue / domain	Adults (18 years or over) with primary hyperparathyroidism Strata: • Preoperative • Non-surgical • Post-operative • People on calcimimetics • People on bisphosphonates • Previous surgery • Normocalcaemic patients Subgroup: • People on HRT Exclude people: • with secondary and tertiary HPT • with multiple endocrine neoplasia • with familial hyperparathyroidism • with parathyroid carcinoma • people on medications interfering with calcium metabolism (lithium).
Outcomes and prioritisation	• Mortality (dichotomous) • Fragility fracture (dichotomous) • Renal stones (dichotomous) • Renal tract calcification (dichotomous) • Pancreatitis (dichotomous)

	• Stroke (dichotomous) • Hypertension (dichotomous) • Myocardial infarction (dichotomous) • Number of people who become eligible for surgery/meet the criteria for surgery (dichotomous) • Serum calcium (>2.85 mmol/L) (dichotomous) (continuous if dichotomous not available) • 24-hour urine for calcium (>10 mmol/dl) (dichotomous) (continuous if dichotomous not available) • BMD of proximal femur (T-score <2.5; Z score <2) (dichotomous) (continuous if dichotomous not available) Report age interaction (this will be a narrative – how age interacts with the outcome of interest) Follow-up: minimum 2 years
Eligibility criteria – study design	Prospective cohort studies Retrospective cohort studies will be included only if insufficient prospective cohort studies are identified. Key confounders: • Age • Absence/presence of end-organ effects • Adjusted serum calcium level
Other inclusion exclusion criteria	Exclusions: • Non-English language papers • Conference abstracts • Not compared to healthy controls • Studies <50 participants Note: This review will be looking at the incidence of outcomes in people with PHPT compared with healthy controls. We will not be looking at studies reporting the incidence in people with PHPT as this is less informative.
Proposed sensitivity / subgroup analysis, or meta-regression	N/A
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). • GRADEpro was used to assess the quality of evidence for each outcome. • Endnote for bibliography, citations, sifting and reference management • Data extractions performed using EviBase, a platform designed and maintained by the National Guideline Centre (NGC).
Information sources – databases and	Clinical search databases to be used: Medline, Embase, Cochrane Library Date: all years

dates	Health economics search databases to be used: Medline, Embase, NHSEED, HTA Date: Medline, Embase from 2014 NHSEED, HTA – all years Language: Restrict to English only Supplementary search techniques: backward citation searching Key papers: Not known
Identify if an update	N/A
Author contacts	https://www.nice.org.uk/guidance/indevelopment/gid-ng10051
Highlight if amendment to previous protocol	For details please see section 4.5 of Developing NICE guidelines: the manual.
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. 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	A multidisciplinary committee developed the evidence review. The committee

contributions of authors and guarantor	was convened by the National Guideline Centre (NGC) and chaired by Jonathan Mant in line with section 3 of Developing NICE guidelines: the manual. Staff from the 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	The NGC is funded by NICE and hosted by the Royal College of Physicians.
Name of sponsor	The NGC is funded by NICE and hosted by the Royal College of Physicians.
Roles of sponsor	NICE funds the NGC to develop guidelines for those working in the NHS, public health and social care in England.
PROSPERO registration number	Not registered

Table 14: 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. • 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). • 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). • Unpublished reports will not be considered unless submitted as part of a call for evidence. 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. 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).[53] **Inclusion and exclusion criteria** • 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. • 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. • If a study is rated as 'Partially applicable', with 'Potentially serious limitations' or both then there is discretion over whether it should be included.

Review question	All questions – health economic evidence
	Where there is discretion 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. The health economist will be guided by the following hierarchies. *Setting:* • UK NHS (most applicable). • OECD countries with predominantly public health insurance systems (for example, France, Germany, Sweden). • OECD countries with predominantly private health insurance systems (for example, Switzerland). • Studies set in non-OECD countries or in the USA will be excluded before being assessed for applicability and methodological limitations. *Health economic study type:* • Cost–utility analysis (most applicable). • Other type of full economic evaluation (cost–benefit analysis, cost-effectiveness analysis, cost–consequences analysis). • Comparative cost analysis. • Non-comparative cost analyses including cost-of-illness studies will be excluded before being assessed for applicability and methodological limitations. *Year of analysis:* • The more recent the study, the more applicable it will be. • 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'. • Studies published before 2002 will be excluded before being assessed for applicability and methodological limitations. *Quality and relevance of effectiveness data used in the health economic analysis:* • 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 are therefore difficult to retrieve. Search filters were applied to the search where appropriate.

Table 15: 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
The Cochrane Library (Wiley)	Cochrane Reviews to 2018 Issue 8 of 12 CENTRAL to 2018 Issue 7 of 12 DARE, and NHSEED to 2015 Issue 2 of 4 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

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

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

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)

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 16: Database date parameters and filters used

Database	Dates searched	Search filter used
Medline	2002 – 06 August 2018	Exclusions Health economics studies
Embase	2002 – 06 August 2018	Exclusions Health economics studies
Centre for Research and Dissemination (CRD)	HTA - Inception – 06 August 2018 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

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

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 Monitoring strategies

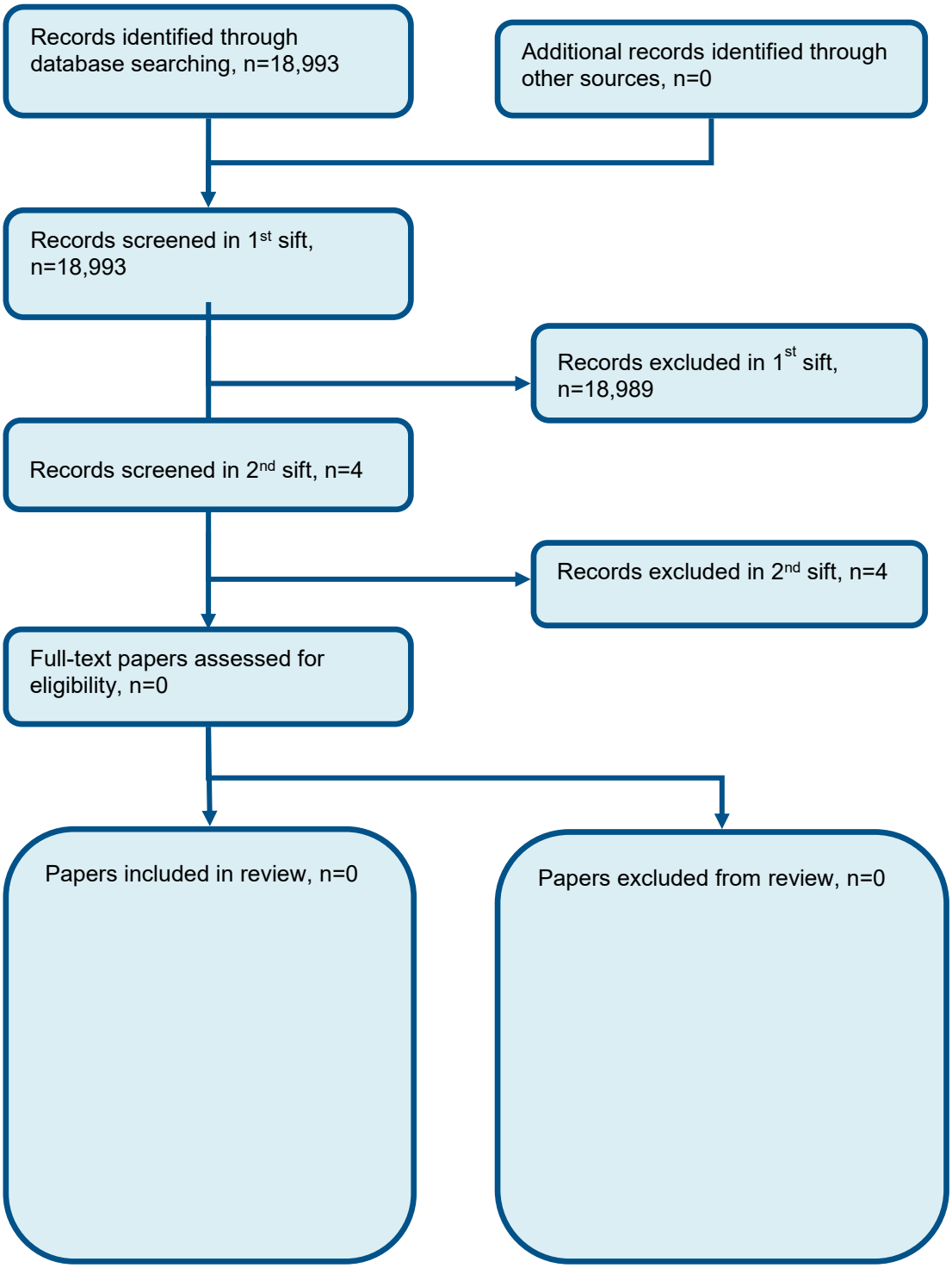
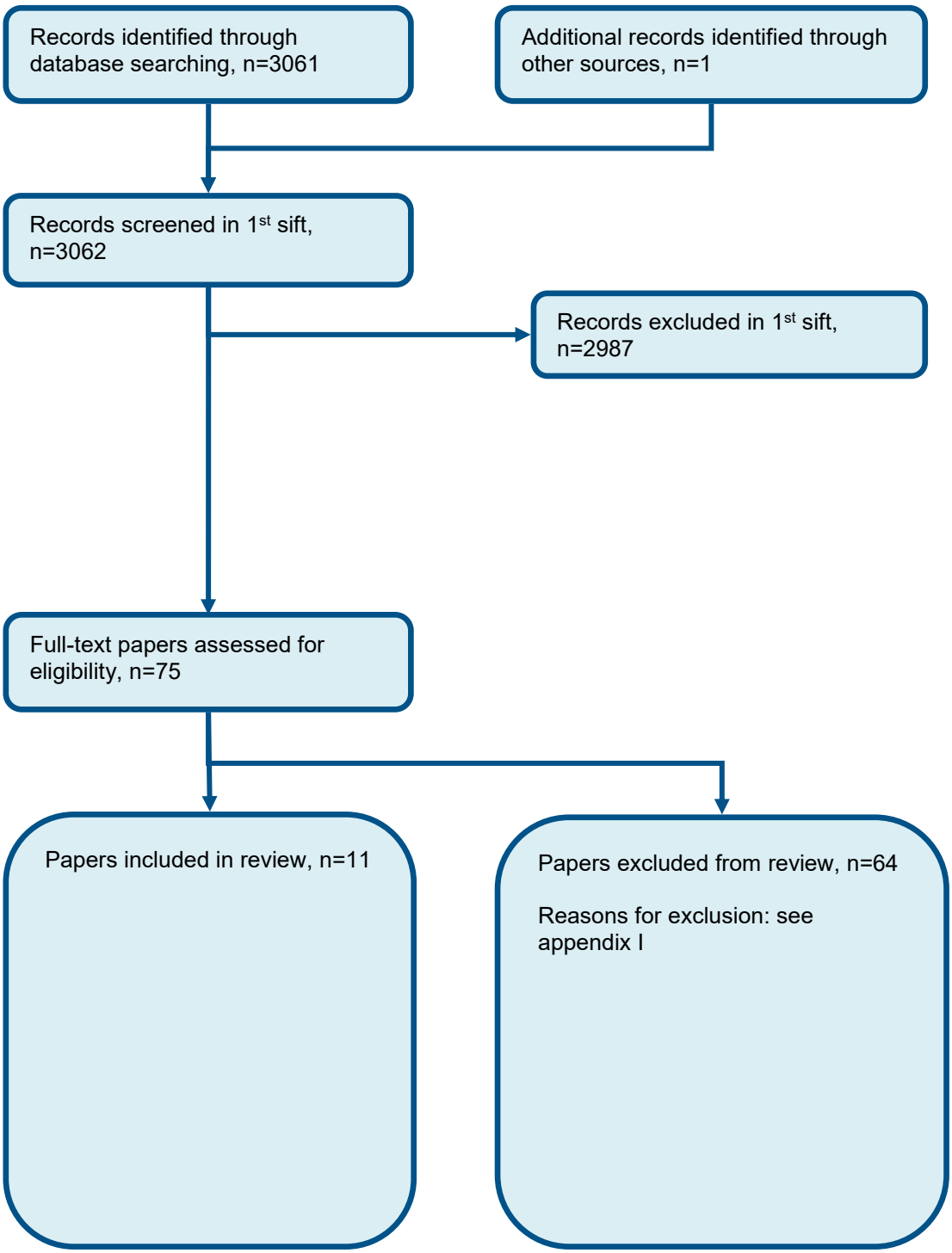


Figure 2: Flow chart of clinical study selection for the review of Monitoring long-term outcomes



Appendix D: Clinical evidence tables

Study	**Yu 2011** [80]
Study type	Prospective cohort study
Number of studies (number of participants)	1 (n=8544) n=1424 cases; n=7120 comparators
Countries and setting	Conducted in UK; Setting: Tayside, Scotland, 1997–2006
Line of therapy	N/A
Duration of study	9 years (1997–2006)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Pre-operative
Subgroup analysis within study	Not applicable
Inclusion criteria	First two raised calcium concentrations lower than 2.90 mm; no calcium exceeded 3 mm during the study period; not treated by parathyroidectomy; not treated with Cinacalcet; absence of pre-existing renal stones and renal failure as indicated in both the hospital records and the serum creatinine records (with baseline serum creatinine ≤150 lm) and absence of osteoporotic fractures, also as indicated in the hospital records. Mild untreated group were defined as untreated PHPT patients whose serum calcium concentrations were <2.9 mmol/L within the first 6 months after a positive diagnosis with absence of previous (prior to PHPT diagnosis) fracture fragility, renal stones and renal failure and not treated with cinacalcet.
Exclusion criteria	Further exclusion criteria were applied to the mild untreated group: serum calcium was follow-up for >6 months; less than two serum calcium measurements within the first 6 months.

Recruitment/selection of patients	All Tayside residents over 20 years of age were considered as potential cases. In conjunction with hospital admission records, nuclear medicine and histology data, the PHPT cohort was primarily diagnosed if they met either of the following biochemical criteria: albumin-corrected serum calcium >2.55 mm (10.22 mg/dl, reference range 2.10–2.55 mm) on at least two occasions, with plasma parathyroid hormone (PTH) concentration >3 pm (13.5 ng/l, reference range 1.0–6.9 pm); or, albumin-corrected serum calcium >2.55 mm (10.22 mg/dl) on a single occasion, with plasma PTH concentration >6.9 pm (31.05 ng/l). A subgroup of 'mild PHPT' patients was further selected for this study, these being patients with untreated PHPT, low concentrations of hypercalcaemia and an absence of renal stones, renal failure and osteoporotic fractures, at diagnosis. Each of the selected patients was then matched with five individuals, or comparators, by age, gender and calendar year of PHPT diagnosis, from the general Tayside population, with either no calcium records or normal serum calcium concentration during the study period. The calendar year of the matching was the index date for each comparator.
Age, gender and ethnicity	Age: cases – 68.3 (13.6); comparators – 68.3 (13.6) Females (%): case 1001 (70.3); comparators – 5005 (70.3)
Further population details	All patients with diagnosed but untreated, mild PHPT. Baseline characteristics: Previous bisphosphonate – cases – 105 (7.4); comparators – 220 (3.1) History of previous condition: cardiovascular: cases 213 (15%); comparators – 445 (6.3%) Cerebrovascular disease: cases – 59 (4.1); comparators – 153 (2.1) Renal failure: cases – 0; comparators – 29 (0.4) Renal stone: cases – 0; comparators – 28 (0.4) Fractures: cases – 25 (1.8); comparators – 250 (3.5) Osteoporotic fractures: cases – 0; comparators – 185 (2.6)
Extra comments	Data were modelled using the Cox proportional hazards models. Confounding covariates considered were multiple deprivation index (SIMD), history of bisphosphonates prescription, history of hospital admitted CVD, cerebrovascular disease, hypertension, renal failure, renal stones, psychiatric disease, fractures, cancer and

	diabetes. The study used individual data for both cases and comparators.
Indirectness of population	No indirectness
Funding	Authors

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT CASES versus MATCHED COMPARATORS

5005 years of follow-up [the median follow-up was 1042 days (2.9 years)]

Protocol outcome 1: Mortality at end of follow-up
-Actual outcome: All-cause mortality (Propensity score adjusted*) HR 1.64 955 CI 1.43-1.87; P<0.001

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

Protocol outcome 2: Myocardial infarction at end of follow-up
-Actual outcome: Fatal CVD (Propensity score adjusted*) HR 1.64 955 CI 1.32-2.04; P<0.001

-Actual outcome: Non-fatal CVD (Propensity score adjusted*) HR 2.48 95% CI 2.13-2.89; P<0.001
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

Protocol outcome 3: Stroke at end of follow-up
- Actual outcome: cerebrovascular (Propensity score adjusted*) HR 2.51 95% CI 1.95-3.22; P<0.001
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

Protocol outcome 4: Hypertension at end of follow-up
- Actual outcome: hypertension (Propensity score adjusted*); HR 2.60 95% CI 2.04-3.31; P<0.001
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

Protocol outcome 5: Renal failure at end of follow-up
- Actual outcome: Renal failure (Propensity score adjusted) HR 13.83 95% CI 10.41-18.37; P<0.001
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

Protocol outcome 6: Renal stones at end of follow-up - Actual outcome: Renal stones (Propensity score adjusted*) HR 5.15 95% CI 2.69-9.83; P<0.001 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 Protocol outcome 7: Fragility fracture at end of follow-up - Actual outcome: All fractures (Propensity score adjusted*) HR 1.75 95% CI 1.36-2.26; P<0.001 - Actual outcome: Osteoporotic fractures (Propensity score adjusted*); HR 1.63 95% CI 1.41-2.18; P<0.001 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 The matching term, age and gender was used as a strata variable in all analyses. Confounding covariates considered were multiple deprivation index, history of bisphosphonates prescription, history of hospital admitted CVD, cerebrovascular disease, hypertension, renal failure, renal stones, psychiatric disease, fractures, cancer and diabetes. *Propensity score is the predicted probability of having calcium checked derived from multiple logistic regression.	
Protocol outcomes not reported by the study	Renal tract calcification at end of follow-up; Pancreatitis at end of follow-up

Study	**Clifton-Bligh, 2015** [21]
Study type	Retrospective cohort study
Number of studies (number of participants)	1 (n=561)
Countries and setting	Conducted in Australia; Setting: Hospital, Department of Endocrinology
Line of therapy	N/A
Duration of study	10 years and 20 years follow-up
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Surgery and non-surgery

Subgroup analysis within study	Not applicable
Inclusion criteria	Patients diagnosed with PHPT.
Exclusion criteria	Not stated
Recruitment/selection of patients	Patients diagnosed with PHPT between 1961 and 1994 were identified, medical records were obtained and examined, and a determination was made as to whether or not they were alive at the end of 1994. Survival data of individuals with PHPT compared with expected survival in the general Australian population was obtained from the Life Tables. Control: The control population was the Australian population at large for whom Life Tables from 1961-1994 existed at the time (Life Tables are published by the Australian Government Actuary).
Age, gender and ethnicity	Age: not stated Females (%):not stated
Further population details	Control population matched for age, sex, the year observation began, and the duration of the observation.
Extra comments	The relative survival over a 20 year time interval was calculated for the patients studied between 1972 and 2011. The group was divided into 2 cohorts: those diagnosed between 1972 and 1981 and those diagnosed between 1982 and 1991, and 20 year relative survival was calculated for each cohort. Before 1972 diagnosis of PHPT was made if surgical removal of a parathyroid gland restored eucalcaemia, or if a full investigation failed to find another cause for hypercalcaemia. After 1972, the diagnosis of PHPT was made if the serum calcium and serum PTH was above the upper limit of the reference range. Because of the concept that a person with mild PHPT might not require surgery, 113 of the patients with mild PHPT were not subjected to neck exploration and 448 patients had parathyroid surgery.
Indirectness of population	No indirectness
Funding	Authors
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL	

Protocol outcome 1: Mortality at end of follow-up
- Actual outcome: Relative survival rate 86.8% (95% CI 84.9-86.2, p<0.001) (10 years)

124/561 patients died between 1961 and 1994*

Relative survival rate 62.9% (95% CI 58.5- 67.4, P<0.001) at (20 years)

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; other - not adjusted for key confounders

Narrative data:

448 had surgery and 113 did not have surgery.

There was no significant difference in the relative survival between surgically and non-surgically treated patients over a 10 year period (figures, no data). The average number of years of life lost by hyperparathyroid patients compared to control population was 7.5 years. There was no significant difference in the death rate between those with an initial serum calcium of >3.00 mmol/L compared with those with an initial serum calcium of <3.00 mmol/L (no data).

In a multivariate analysis in the surgically treated group, the serum calcium did not significantly influence survival (HR 1.57, 95% CI 0.30-8.30, p=0.593). In a multivariate analysis, risk factors associated with death in the surgically treated group were diabetes mellitus (HR 4.09, 95% CI 1.42-6.74, P=0.001), congestive cardiac failure (HR 5.46, 95% CI 1.31-22.87, P=0.002), coronary heart disease (HR 2.16, 95% CI 1.08-0.044). The presence of kidney stones before surgery was associated with reduced mortality (HR 0.364, 95% CI 0.22-0.68, P=0.001).

In the non-surgically treated group, death was significantly associated with a high serum PTH (HR 1.59, 95% CI 1.20-2.11, p=0.001), coronary heart disease (HR 3.10, 95% CI 1.42-6.74, P=0.004), and kidney stones (HR 2.48, 95% CI 1.07-5.76, p=0.035). This difference between the surgically treated and non-surgically treated group with respect to the impact of kidney stones is not clear. 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 (95% 0.38-1.18, p=0.167).

Using a 20 year follow-up for the whole group, multivariate analysis did not show any survival difference between male and female, surgery vs non-surgery (p=0.867), serum calcium >3 mmol/l versus <3 mmol/L (p=0.794), or serum PTH analysed as quartiles (no data).

Protocol outcomes not reported by the study	Renal stones, fragility fracture, renal tract calcification, stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/L), 24-h urine for calcium (>10 mmol/dl), BMD of proximal femur (T-score <2.5; Z score <2)

Study	De Geronimo, 2006 [26]
Study type	Cohort study
Number of studies (number of participants)	1 (n=196): n= 98 PHPT (n=25 mild PHPT; n= 73 non-mild PHPT); n=98 control
Countries and setting	Conducted in Italy; Setting: Mineral metabolic centre
Line of therapy	N/A
Duration of study	N/A
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Surgery and no surgery patients
Subgroup analysis within study	Not applicable
Inclusion criteria	98 consecutive post-menopausal patients with PHPT. They were grouped as mild or non-mild according to criteria established by the Consensus Development conference on the management of asymptomatic primary hyperparathyroidism: Serum calcium greater than 1m g/dl above the upper limits of normal; 24-h total urine calcium excretion of more than 400 mg; creatinine clearance reduced by more than 30% compared to age-matched persons; bone density at lumbar spine, hip or distal radius that is more than 2.5 SD below peak bone mass; patients under 50 years of age.
Exclusion criteria	Not reported
Recruitment/selection of patients	PHPT: 98 consecutive postmenopausal patients with PHPT from Mineral metabolism centre. Control: 89 healthy postmenopausal women, matched for age, years since menopause, BMI. They were randomly selected from ambulatory post-menopausal women sent by their GPs to the hospital as part of a menopause-screening programme.
Age, gender and ethnicity	Age: mild PHPT 60.84 (6.82); non-mild PHPT 61.5 (8.40); control 60.65 (6.92)

	Females (%): all women -100%
Further population details	25 of the 98 patients were considered as suffering from mild disease. In this group of patients hypercalcaemia was occasionally detected in the course of the standard biochemical evaluation performed on all subjects undergoing BMD measurement. In the remaining 73 patients, a severe disease was present; 13 of them had a history of non-vertebral fractures; 7 referred by GPs for nephrolithiasis; 21 referred by GPs for osteoporosis; 15 patients renal stones shown by ultrasonography; 3 were referred for hypercalcaemia; 2 had life threatening episode of pancreatitis. The remaining 12 patients complained of bone pain and/or neuromuscular symptoms.
Extra comments	Comparison: Mild or non-mild PHPT post-menopausal women vs Healthy subjects Outcome assessment: All patients and normal subjects had standardised lateral radiographs in anterior-posterior and left lateral projections of the thoracic and LS, centred at T8 and L3 respectively. The radiographs were examined first for quality and then for fractures by an experienced skeletal radiologist. Vertebral deformity was defined according to the semi-quantitative method, when anterior, middle, or posterior height loss was more than 20% with respect to the adjacent vertebra. This criteria for defining vertebral fracture had a relatively high true-positive rate based on the classifications from previous reports.
Indirectness of population	No indirectness
Funding	Authors

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL

Protocol outcome 1: BMD at end of follow-up (Mean (SD))
-Actual outcome: Lumbar spine BMD (mg/cm^2) (mild PHPT vs healthy women) mean (SD)

mild PHPT 916.1 (100.4); control 839.0 (109.8)

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 ; other-not adjusted for key confounders

Protocol outcome 2: BMD at end of follow-up
-Actual outcome: Femoral neck - BMD (mg/cm^2) (mild PHPT vs healthy women) (Mean (SD))

mild PHPT 709.0 (92.7); control 690.6 (109.1)

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 ; other-not adjusted for key confounders

Protocol outcome 3: BMD at end of follow-up
-Actual outcome: Total femur - BMD (mg/cm^2) (mild PHPT vs healthy women) (Mean (SD))

mild PHPT 823.4 (116.2); control 802.2(116.0)

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; other-not adjusted for key confounders

Protocol outcome 4: Fractures at end of follow-up
-Actual outcome: Vertebral fractures (mild PHPT vs healthy women)

mild PHPT n=11 (44%); control n=8 (9%)

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; other-not adjusted for key confounders

Protocol outcome 5: Fractures at end of follow-up
-Actual outcome: vertebral fractures (non-mild PHPT vs healthy women)

non-mild PHPT n=35 (47%); control n=8 (9%)

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; other-not adjusted for key confounders

Protocol outcome 6: Fractures at end of follow-up
-Actual outcome: Non-vertebral fractures (mild PHPT vs healthy women)

mild PHPT n=0; control n=17 (19.1%)

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 ; other-not adjusted for key confounders Protocol outcome 7: Fractures at end of follow-up -Actual outcome: Non-vertebral fractures (non-mild PHPT vs healthy women) non-mild PHPT n=13 (17.8%); control n=17 (19.1%) 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; other-not adjusted for key confounders Protocol outcome 8: BMD at end of follow-up -Actual outcome: Lumbar spine-BMD (mg/cm^2) (non-mild PHPT vs healthy women) (Mean (SD)) non-mild PHPT 765.4 (156.6); control 839.0 (109.8) (Mean (SD)) 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; other-not adjusted for key confounders Protocol outcome 9: BMD at end of follow-up -Actual outcome: Femoral neck-BMD (mg/cm^2) (non-mild PHPT vs healthy women) (Mean (SD)) non-mild PHPT 601.3 (102.5); control 690.6 (109.1) 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; other-not adjusted for key confounders Protocol outcome 10: BMD at end of follow-up -Actual outcome: Total femur-BMD (mg/cm^2) (non-mild PHPT vs healthy women) (Mean (SD)) non-mild PHPT 699.7 (126.0); control 802.2 (116.0) 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; other-not adjusted for key confounders Note: Includes both surgery and non-surgery patients but does not report exact number of patients who underwent surgery.	
Protocol outcomes not reported by the study	Mortality, stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/l), 24-h urine for calcium (>10 mmol/dl)

Study	Hedback 1998 [38]
Study type	Retrospective cohort study
Number of studies (number of participants)	1 (n= 4461 [915 men and 3546 women])
Countries and setting	Conducted in Sweden; Setting: hospital/community
Line of therapy	N/A
Duration of study	Total number of observation years was 3205 giving mean follow-up times of 3.6 and 3.5 years respectively (range 0-8 years)
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Post-operative
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients diagnosed as having PHPT or parathyroid adenoma
Exclusion criteria	Not stated
Recruitment/selection of patients	Patients diagnosed as having PHPT or parathyroid adenoma according to the International Classification of Diseases during 1987-94 from Swedish national patient registry. The patient series studied consisted of those individuals who at the same time were reported to have undergone removal of a parathyroid gland or adenoma. The inclusion date of a patient was the date of arrival at the hospital before surgery. Control-whole Swedish population matched for age, sex and calendar year.
Age, gender and ethnicity	Age: mean age at surgery was 61.3 years (SD 14.4 years) for men; 64.7 years (SD 12.7 years) for women.

	Females: 3546/4461
Further population details	The number of deaths in the patient series was compared with the number of expected deaths, estimated on the basis of the annual official reports published by the Swedish Central Bureau of Statistics, which give the total number of deaths from different causes, separately for men and women, in 5 year age cohorts.
Extra comments	-
Indirectness of population	No indirectness
Funding	Funds of Sahlgrenska Hospital for tumour disease research.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL

Protocol outcome 1: Mortality at end of follow-up
-Actual outcome: Mortality: Male: RR 1.30 (95% CI 1.07- 1.57); observed: n=107; expected: n=82.2

- Female: RR 1.61 (95% CI 1.46-1.78); observed: n=396; expected: n=245.6

Altogether: observed: n=503; expected: n=327.8

Narrative data:

Yearly death reduction

Hyperparathyroid population operated on in 1987-94:

Male: 17% (95% 7-26)

Female: 8% (2.00-13)

Swedish population 1974-1983: male: 0.95% (95% CI 0.81-1.09)

Female: 1.68% (1.53-1.83)

Swedish population 1987-94: male: 1.51% (1.34-1.67)

Female: 0.88% (0.70-1.05) Note from authors: the number of patients operated on annually were fairly constant, and no great change in attitude concerning management of these patients occurred in Sweden during the study period. Another explanation could be frequent occurrence of cardiovascular disease among these patients, together with the improvement in the treatment of cardiovascular diseases that has taken place in recent years. Also, treatment of PHPT would in itself be treatment of cardiovascular disease and might reduce cardiovascular mortality among PHPT patients. 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; other-not adjusted for key confounders	
Protocol outcomes not reported by the study	Fractures, Stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/L), BMD, 24-h urine for calcium (>10 mmol/dl)

Study	**Kenny 1995** [41]
Study type	Retrospective cohort study
Number of studies (number of participants)	1 (n=90) n=46 PHPT; n= 44 control
Countries and setting	Conducted in USA; Setting: hospital and community
Line of therapy	N/A
Duration of study	Follow-up 5 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Post-operative
Subgroup analysis within study	Not applicable
Inclusion criteria	PHPT group- Post-menopausal women who had undergone parathyroidectomy for hyperparathyroidism during

	a 5 year period (1986 to 1991) Control - postmenopausal women without hyperparathyroidism contacted by random digit dialling and interviewed as controls.
Exclusion criteria	Not stated
Recruitment/selection of patients	Medical records at the University of Connecticut Health Centre and Hartford Hospital were reviewed for the diagnosis code for hyperparathyroidectomy. Surgeons involved in their care agreed to send introductory letters informing the women that they would be invited to answer a health survey by telephone during the upcoming month. All surgeons contacted to aid in recruitment agreed to participate. Fifty seven post-menopausal women who had undergone successful parathyroidectomy for pathologically confirmed hyperparathyroidism at the University of Connecticut Health Centre or Hartford Hospital between 1987 and 1992 were contacted to assess their interest in participating in the study: 46 women agreed to participate, 5 were unreachable, 4 declined participation, and 2 had died of complications of cerebral vascular accident and pneumonia. Controls were obtained by random digit dialling with the two most common prefixes of the case subjects; 512 households were contacted and asked whether there was a 50 years old or older who would be interested in answering a health survey. Forty-four postmenopausal women consented to participate.
Age, gender and ethnicity	Age: PHPT- 68.9 (10.8); control- 67.4 (10.8) Females (%): all women
Further population details	The women in the control were group were similar to those with PHPT in regard to age, weight, and height. Medical conditions reported by 46 patients with PHPT and 44 control subjects: Hypothyroidism- PHPT 13 (28%); control 5 (11%) Cancer- PHPT 5 (11%); control 2 (5%) Diabetes mellitus- PHPT 3 (7%); control 2 (5%) Hypertension- PHPT 10 (22%); control 15 (34%)
Extra comments	Not adjusted for key confounders
Indirectness of population	No indirectness

Funding	Not reported
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL Protocol outcome 1: Fracture at end of follow-up -Actual outcome: Fracture incidence at 5 years follow-up: PHPT - 22/46; control- 11/44 92% of the total fractures and all of the post-menopausal fractures occurred after minor trauma. When women who presented because of bone disease were excluded from the analysis, reports of fracture remained significantly higher in the PHPT group (40%, 16/40) than in controls (p=0.05). Multiple fractures also occurred more commonly in those with PHPT. 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 ; other-not adjusted for key confounders	
Protocol outcomes not reported by the study	Mortality, stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/l), BMD, 24-h urine for calcium (>10 mmol/dl)

Study	**Khosla 1999** [42]
Study type	Retrospective cohort study
Number of studies (number of participants)	1 (n=407 patients with PHPT)
Countries and setting	Conducted in USA; Setting: community
Line of therapy	N/A
Duration of study	28 years (1965–1992)

Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Surgery and non-surgery
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients with PHPT
Exclusion criteria	Not stated
Recruitment/selection of patients	The database the Rochester Epidemiology Project was used to identify 435 Rochester residents with PHPT during period, 1965-1992. Seven of these patients refused subsequent authorisation for chart review, and 21 had no follow-up after age 35 years, resulting in a final cohort of 407 subjects. These subjects were then followed forward in time through their linked medical records in the community (retrospective cohort study) until death or the most recent clinical contact and backward in time to the first medical record entry in the community.
Age, gender and ethnicity	Age: NR Females (%): NR
Further population details	The majority were women (344, 775) and most were 45 years of age or older at the diagnosis of HPT (335, 82%). The mean age at diagnosis was 57.8 years. The average maximum serum calcium level was 10.9 (0.6) mg/dl. Median serum PTH, measured by a C-terminal assay, was 47 µl.eq/ml (25-75% interval, 33-71 µl.eq/ml; normal ≤50 µl.eq/ml). The majority of the patients were asymptomatic and the relative proportion of patients presenting with a symptom or complication of primary PHPT (urolithiasis, fracture, hypercalcaemic crisis, peptic ulcer disease, pseudogout, or band keratopathy) declined as biochemical screening was introduced in this population: 21% in the prescreening era (1965-June 1974), 9% following the introduction of auto-mated serum calcium determinations (July 1974-1982), and 2% in the postscreening era (1983-1992). Most of these patients were managed conservatively, with parathyroid surgery ultimately performed on only 93 patients (23%).
Extra comments	For consistency, the index date for the diagnosis of primary PHPT was the date when hypercalcaemia was first evident, and not when the clinician recorded the diagnosis of primary PHPT. For each subject, all in-patient and outpatient medical records at any local provider of medical care were searched for the occurrence of specific fractures. Emphasis was on fractures at the skeletal sites usually associated with osteoporosis and

	these were recorded regardless of whether they occurred before or after the recognition of primary PHPT. The diagnosis of vertebral fracture was accepted on the basis of a radiologists report of compression or collapse of one or more thoracic or lumbar vertebrae. All fractures were classified according to the circumstances of the injury. By convention, falls from standing height or less were considered moderate trauma, while motor vehicle accidents and falls from heights were deemed severe trauma.
Indirectness of population	No indirectness
Funding	Grants from the National Institutes of Health, United States Public Health Service.

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL

Protocol outcome 1: Fracture at end of follow-up
-Actual outcome: any fracture: observed: 202; expected: 154.6; SIR 1.3 (95% CI 1.1-1.5)*

*Predictors of the risk of mild/moderate trauma, vertebral, distal forearm or proximal femur fractures:

Multivariate model

Age (per 10 year increase): relative hazard 1.6 (95% 1.4-1.9)

Female gender: relative hazard 2.3 (95% CI 1.2-4.1)

By multivariate analysis, only age and female gender were significant independent predictors of fracture risk.

Actual outcome: vertebral fracture: observed: 79; expected: 24.6; SIR 3.2, 95% CI 2.5-4.0

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 ; other-not adjusted for key confounders

Narrative data:

The cohort of 407 patients with primary PHPT was followed for 5766 person-years and 295 patients (73%) were still alive at last follow-up.

23% of PHPT patients underwent parathyroid surgery

Protocol outcomes not reported by the study	Stroke, hypertension, pancreatitis, myocardial infarction, serum calcium (>2.85 mmol/l), BMD, 24-h urine for calcium (>10 mmol/dl)

Study	**Larsson, 1993** [45]
Study type	Retrospective cohort study
Number of studies (number of participants)	1 (n=1373 patients with PHPT)
Countries and setting	Conducted in Sweden; Setting: hospital
Line of therapy	N/A
Duration of study	19 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Surgery and non-surgery
Subgroup analysis within study	Not applicable
Inclusion criteria	All patients who were admitted to hospital with the diagnosis of PHPT.
Exclusion criteria	Not stated
Recruitment/selection of patients	During the period 1965-1983, all admissions to hospitals in the Uppsala Health care region were reported to the inpatient register, a population-based computerised register maintained by the National Board of Health and Welfare. The six counties in the region comprised both rural and urban areas, and the population in 1983 was 1.48 million i.e. approximately 17% of the total Swedish population. From 1965 to 1983, a total of 3,486 million hospital admissions were included in the In-patient Register. In the study analysis, data on admissions from 5 of the 6 counties in Uppsala Health care region were utilised.
Age, gender and ethnicity	Age: mean age at diagnosis: 62.6 (11.7); mean age at end of follow-up: 67.8 (10.8) Females: males: 1373/551

Further population details	Included in the final cohort were 1924 persons, 1373 females, and 551 males, 30 years and older, who, during the years 1965-1983 were admitted to hospital with the diagnosis of HPT. Of these, 1318 patients (69%), 975 females and 343 males had been subject to parathyroidectomy at the same or a subsequent admission. The annual number of patients with the diagnosis of PHPT increased during the study period, the increment being most prominent in women above the age of 60. In the general population, the prevalence of PHPT is estimated to be about 1% in females above the age of 50, but considerably lower in males and in other age groups. In only 3 cases was PHPT detected at the admission for hip fracture, 2 because of hypercalcaemic symptoms. Mean serum calcium at the time of diagnosis for women with PHPT was 2.91 (0.22), and for men 2.79 (0.18) mmol/l; p<0.01. The serum calcium levels at diagnosis were stable during the study period.
Extra comments	Comparison: Patients with PHPT vs Entire background population Follow-up: For each patient the observation period started in 1965 or at the age of 30 for those below this age in 1965 and ended at the date of hip fracture diagnosis, death, or at the end of 1983. Expected outcome: In calculation of the expected numbers of fractures, the calendar year and age-specific incidence rates of hip fractures in the entire background population in the region were used. The respective number of accumulated person-years of observation was multiplied by the year and age-specific incidence rates of hip fractures, yielding the expected number of fractures during the entire study period.
Indirectness of population	No indirectness
Funding	Financial support received from the Swedish Medical Research council and the Faculty of Medicine, Uppsala University.
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL Protocol outcome 1: Hip fractures at end of follow-up -Actual outcome: Hip fractures women: observed: 67/1373; expected: 71.76; RR 0.93 (95% CI 0.72-1.19) Actual outcome: Hip fractures men: observed: observed 11/551; expected 7.9; RR 1.39 (95% CI 0.69-2.50)	

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 ; other-not adjusted for key confounders Narrative data: 1318 patients (69%) underwent parathyroid surgery	
Protocol outcomes not reported by the study	Mortality, Stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/l) , BMD, 24-h urine for calcium (>10 mmol/dl)

Study	**Melton, 1992** [49]
Study type	Population based retrospective cohort study
Number of studies (number of participants)	1 (n=180) n=90 patients with PHPT; n=90 matched control subjects
Countries and setting	Conducted in USA; Setting: community
Line of therapy	N/A
Duration of study	1965–1976
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Surgery and non-surgery
Subgroup analysis within study	Not applicable
Inclusion criteria	PHPT newly diagnosed during the life among Rochester during 1965 through 1976
Exclusion criteria	Patients with an incidental autopsy diagnosis of parathyroid adenoma or hyperplasia were excluded.

Recruitment/selection of patients	Records of patients first diagnosed with PHPT retrieved form the medical record linkage system (the Rochester Epidemiology Project) as a result of inpatient or outpatient care, death certification or autopsy. 90 cases of PHPT newly diagnosed during the life among Rochester during 1965 through 1976; 83 subjects had histopathologic proof of parathyroid adenoma or hyperplasia or had hypercalcaemia with pathognomonic radiographic signs and/or elevated serum parathyroid hormone concentrations. The remainder had hypercalcaemia for more than a year without another cause being found after careful evaluation. Control: The 90 patients were matched by age and gender to control subjects from the local population who had no evidence of PHPT. The control pool consisted of Rochester residents who were medically attended at Mayo clinic during year (±2 years) in which each patient was initially diagnosed.
Age, gender and ethnicity	Age (mean) years: PHPT-58.5; control-58.7
Further population details	The matched patients and control subjects were followed up in the time through their linked medical records in the community until death or the most recent clinical contact.
Extra comments	Not adjusted for serum calcium level and absence/presence of end organ effects. The medical records contained the clinical history and the radiologists report of each fracture, but the original roentgenograms were not available for review. Fractures were also classified according to the circumstances of the injury: falls from standing height or less were considered moderate trauma; motor vehicle accidents and falls from heights were deemed severe trauma. Fracture ascertainment is believed to be complete except for vertebral fractures, some of which are never diagnosed.
Indirectness of population	No indirectness
Funding	Grants from the National Institutes of Health, Bethesda, Md.
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL Protocol outcome 1: Fractures at end of follow-up -Actual outcome: all fractures (after diagnosis) during 1072 persons-years follow-up: PHPT: 50/90; control: 52/90 -Actual outcome: vertebral fractures: PHPT: 9/90; control: 5/90 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; other-not adjusted for key confounders

Narrative data:

Surgery: 19/50

No surgery: 31/50

Fractures were more common among the patients with PHPT at most skeletal sites, but none of the individual differences reached statistical significance. Altogether, 10% of the patients and 7% of the control subjects had experienced a fracture at age 35 years or older, which was attributed to one of the three mechanisms usually associated with osteoporotic fractures (simple falls, spontaneous fractures, or in the case of some vertebral fractures, diagnosis as an incidental finding).

Fracture risk after index date (after diagnosis):

Overall: PHPT 50/90; control 52/90; RR 1.0 (95% CI 0.7-1.4) (not significant)

Calcium ≥2.74 mmol/l: PHPT 34/90; control 24/90; RR 1.4 (95% CI 0.8-2.4) (not significant)

Calcium <2.74 mmol/l: PHPT 16/90; control 27/90; RR 0.6 (95% CI 0.3-1.1) (not significant)

Operated on: PHPT 19/90; control 26/90; RR 0.7 (95% CI 0.4-1.3) (not significant)

Not operated on: PHPT 31/90; control 26/90; RR 1.2 (95% CI 0.7-2.0) (not significant)

Comorbid conditions: PHPT 44/90; control 38/90; RR 1.2 (95% CI 0.8-1.8) (not significant)

No comorbid conditions: PHPT 6/90; control 14/90; RR 0.4 (95% CI 0.2-1.1) (not significant)

Women: 43/90; 42/90; RR 1.0 (95% CI 0.7-1.6) (not significant)

Men: 7/90; 10/90; RR 0.7 (95% CI 0.3-1.8) (not significant)

In a multivariate analysis, only age at diagnosis was an independent predictor of fracture risk in PHPT (P<0.2). A 10-year increase in age corresponded to a 36% increase in fracture risk.

Protocol outcomes not reported by the study	Mortality, stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/l), BMD, 24-h urine for calcium (>10 mmol/dl)

Study	Ronni-Sivula, 1985 [60]
Study type	Retrospective cohort study
Number of studies (number of participants)	1 (n=668) n=334 (PHPT); n=334 healthy controls
Countries and setting	Conducted in Finland; Setting: department of surgery
Line of therapy	N/A
Duration of study	Operated 1956–1979 and follow-up 1980–1982
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Post-operative
Subgroup analysis within study	Not applicable
Inclusion criteria	Patients operated on for PHPT in the years 1956-79.
Exclusion criteria	Not stated
Recruitment/selection of patients	334 patients (83 men and 251 women) were operated on for PHPT in the years 1956-79. A follow-up study of this material was performed in the years 1980-82. 34 patients had died before the end of the year 1980. All clinical data as well as causes of death of these patients were collected and examined. Control group: Each PHPT patient in the original material was given a pair who was sex and age matched and operated on for appendicitis, varicose veins or haemorrhoids in the same year. The control group consisted of 334 patients with the same sex and age distribution and a follow-up time of the same length as the PHPT patients. The cause of death and age at the time of death were checked in those controls who had died before the end of the year 1980.
Age, gender and ethnicity	Age (mean) years: PHPT-53 (46 years in men and 55 years in women)

	Females: males: 3:1
Further population details	-
Extra comments	No multivariate analysis. Not adjusted for serum calcium level and absence/presence of end organ effects.
Indirectness of population	No indirectness
Funding	Not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL

Protocol outcome 1: mortality at end of follow-up
-Actual outcome: mortality at end of follow-up: PHPT patients 34/334; control 21/334

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; other-not adjusted for key confounders

Narrative data:

Of the 334 PHPT patients 34 (10.2%) had died, 11 men (13.3%) and 23 women (9.2%). The mean age at death was 65 years, 61 years in men and 66 years in women. Of the dead PHPT patients 5 patients had lived less than one year, 23 patients 1-5 years and 6 patients 5-16 years.

21 of the 334 control patients (6.3%) had died during the same follow-up time, 7 men (8.4%) and 14 women (5.6%).

The mean age of the control patients at death was 67 years, 62 years in men and 69 years in women. Of the control patients who died 6 lived less than one year, 7 lived 1-5 years and 8 over 5 years. The difference in mortality between the PHPT patients and the controls was statistically significant ($p<0.05$).

Characteristics of the deceased PHPT patients:

The deceased patients in PHPT group had higher mean value of serum calcium pre-operatively than patients in the entire PHPT group (3.31 mmol/l vs 3.08 mmol/l).

In the deceased patients in the PHPT group, serum creatinine was elevated (>115 mmol/l) pre-operatively in 15 (44%) of the deceased patients. In the entire PHPT group serum creatinine was elevated pre-operatively in 57 patients (17%). In the deceased patients serum creatinine was most often elevated in the groups with hypercalcaemic crises (4/6) and cystic bone changes (3/4) and most rarely in the renal stone group (1/5).

PHPT patients who died had more severe form of disease: 55% had hypercalcaemic crises and 24% had cystic bone changes, 4% had renal stones.

Protocol outcomes not reported by the study	Fracture, stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/l) , BMD, 24-h urine for calcium (>10 mmol/dl)

Study	**Suh 2008 [69]**
Study type	Retrospective cohort study
Number of studies (number of participants)	1 (n=571) n=271 PHPT; n=500 control
Countries and setting	Conducted in USA; Setting: hospital
Line of therapy	N/A
Duration of study	2001–2006
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Pre-operative
Subgroup analysis within study	Not applicable
Inclusion criteria	All patients who had undergone renal imaging within 6 months before parathyroid surgery for PHPT.
Exclusion criteria	Not stated
Recruitment/selection of patients	PHPT: Renal sonography scans of 271 patients with surgically proven parathyroid adenomas were reviewed. Each had a renal imaging study completed at Rhode Island hospital or other sites in the Rhode Island Medical Imaging network within 6 months before parathyroidectomy performed between January 1, 2001 and January 1, 2006. Renal imaging was routinely requested for all patients with a suspected parathyroid adenoma. Control: The records and images of a control group were collected to assess the prevalence of renal calculi in patients not being evaluated for PHPT. This group included age matched subjects who had right upper

	quadrant sonograms obtained from July 1, 2006 to Sep 29, 2006 for various reasons (e.g. upper abdominal pain, cholelithiasis, abnormal liver function studies) in 500 patients.
Age, gender and ethnicity	Age (mean): PHPT-62 years; control- 59 years Females/males: PHPT- 226/45; control-288/212
Further population details	-
Extra comments	The sonography images were initially interpreted by reviewing the original radiology report for each examination. The findings in the reports were classified as positive or as negative for renal calculi. Positive cases were re-evaluated by a board certified reviewer to confirm stones because it is well documented in the literature that blood vessels may be misinterpreted as calculi on sonography. For definitive diagnosis stones had to show echogenicity (i.e. echo difference), posterior acoustic shadowing or a positive twinkle sign. The bladder was not evaluated in either study group.
Indirectness of population	No indirectness
Funding	Not stated
RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL Protocol outcome 1: Renal stones at end of follow-up -Actual outcome: renal stones (definitive calculi): PHPT group: 19/271; Control: 8/500 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; other-not adjusted for key confounders <u>Narrative data:</u> PHPT: In each patient, stone disease was unilateral. None of these kidneys had hydronephrosis, and none of the patients had symptoms of renal stones. Control group: Two kidneys with mild hydronephrosis. In neither case of hydronephrosis was a stone detected. In both study groups, stones varied in size from 3 to 20 mm.	
Protocol outcomes not reported by the study	Mortality, fracture, pancreatitis, stroke, hypertension, myocardial infarction, serum calcium (>2.85 mmol/l) ,

	BMD, 24-h urine for calcium (>10 mmol/dl)

Study	**Wilson, 1988** [79]
Study type	Retrospective cohort study (Prospectively collected data retrospectively analysed)
Number of studies (number of participants)	1 (n=174)
Countries and setting	Conducted in USA; Setting: Outpatient department of a bone and mineral metabolism clinic.
Line of therapy	N/A
Duration of study	10 years
Method of assessment of guideline condition	Adequate method of assessment/diagnosis
Stratum	Non-surgical
Subgroup analysis within study	Not applicable
Inclusion criteria	All patients with mild asymptomatic PHPT. Absence of symptoms due to PHPT, no current kidney stone disease, a plasma calcium level of less than 3.00 mmol/L, a plasma creatinine level of less than 133µmol/L, no radiographic evidence of osteitis fibrosa, and a forearm bone densitometry not more than 2.5 standard deviations below the mean value expected for age, sex and race.
Exclusion criteria	Not stated
Recruitment/selection of patients	All patients meeting inclusion criteria in the clinic between Jan 1976 and Dec 1985. During the 10 year period 174 patients were seen. All had been referred by physicians practicing in and around the Detroit metropolitan area. In almost every patient hypercalcaemia was an incidental finding. The mean plasma calcium level was 2.77 (0.09) mmol/L (reference value, 2.40 (0.08).

Age, gender and ethnicity	Age (years): PHPT- 62 (12); control (range): 45-74
Further population details	Comparison: PHPT group vs Healthy white women. These data had been obtained in 200 ambulatory white female patients having routine annual physical examination.
Extra comments	Diagnostic criteria for PHPT: persistent hypercalcaemia (calcium ≥2.65 mmol/L), with no clinical indication of another cause, and evidence of parathyroid hormone hypersecretion, with increased or non-suppressed values of radioimmunoassay of parathyroid hormone or nephrogenous cyclic adenosine monophosphate excretion per unit of glomerular filtrate. The diagnosis was made at the clinic between 1 January 1976 and December 1985. For the outcome vertebral fractures: the presence or absence of vertebral fractures was determined independently by 2 observers, and any doubtful vertebrae were measured. A reduction in anterior height of more than 20% compared with an adjacent vertebra was classified as a fracture; the study criteria for vertebral fracture included both wedge and compression fractures. The results were compared with data previously reported from the institution.
Indirectness of population	No indirectness
Funding	Not stated

RESULTS (NUMBERS ANALYSED) AND RISK OF BIAS FOR COMPARISON: PHPT versus CONTROL

Protocol outcome 1: Fractures at end of follow-up
-Actual outcome: Vertebral fractures: PHPT: 3/174; control: 3/55

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; other-not adjusted for key confounders

Narrative data:

Only 3 patients had a vertebral fracture, 2 at T12 and one at L1; the overall prevalence was 1.7%. In white women, the group at greatest risk, the prevalence was 2.8%; the 95% CI for the proportion in this group was 0 to 5.9. This prevalence was not significantly different from that previously observed in white women without acute illness. In the 3 patients with vertebral fractures, no new fractures occurred during more than 10 years of observation in two and during

more than 5 years in one. In one patient, a vertebral fracture had been present for at least 11 years before PHPT was diagnosed; the total observation period was 21 years; and no worsening deformation or new fracture occurred. Ten cases of wrist fractures occurred in 8 patients, but no patient had any history of other fractures.	
Protocol outcomes not reported by the study	Mortality, stroke, pancreatitis, hypertension, myocardial infarction, serum calcium (>2.85 mmol/l), BMD, 24-h urine for calcium (>10 mmol/dl)

Appendix E: Forest plots

E.1 PHPT cases versus matched comparators (adjusted for key confounders) Strata – Non-surgical

Figure 3: All-cause mortality

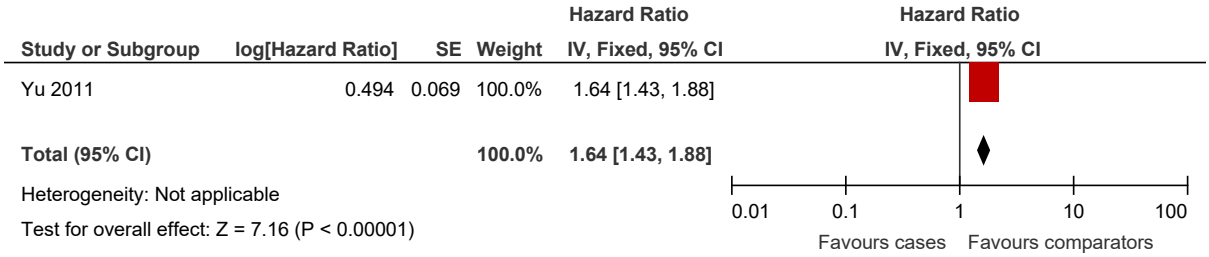


Figure 4: Fatal CVD

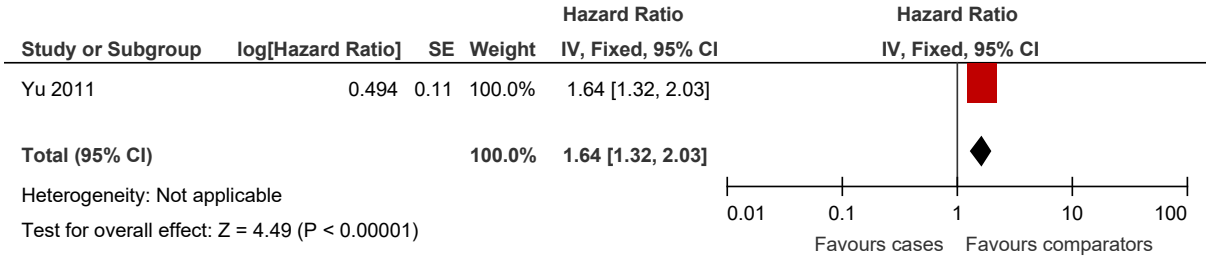


Figure 5: Non-fatal CVD

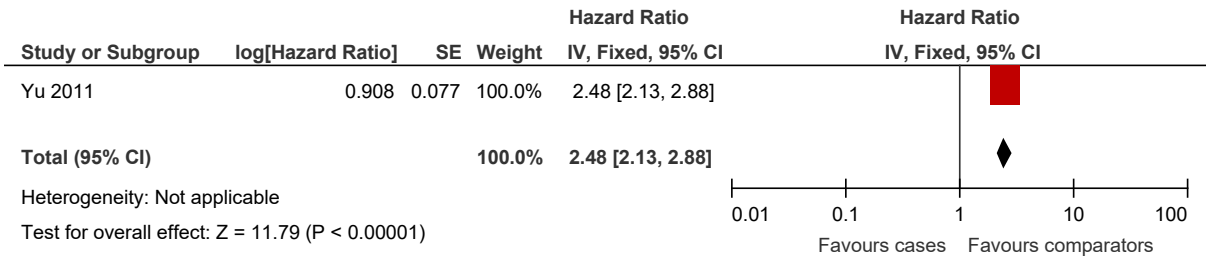


Figure 6: Cerebrovascular disease

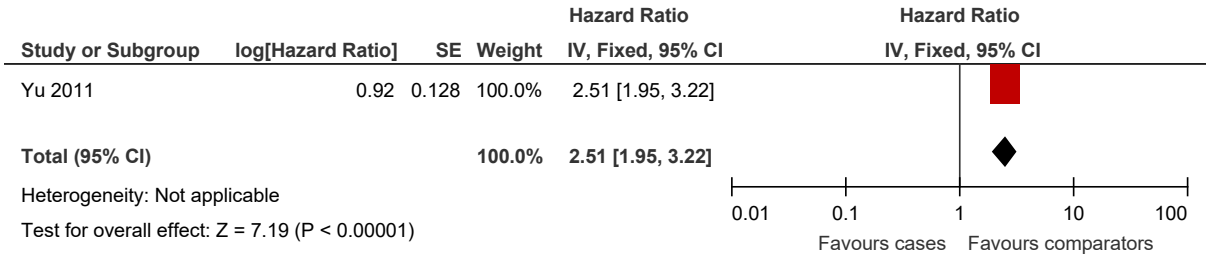


Figure 7: Hypertension

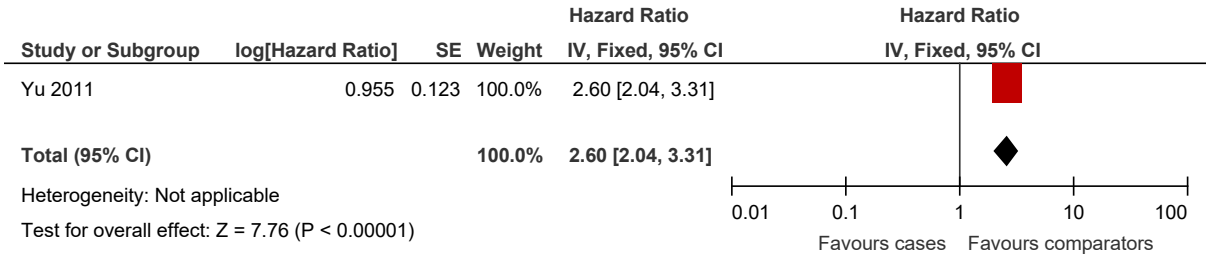


Figure 8: Renal failure

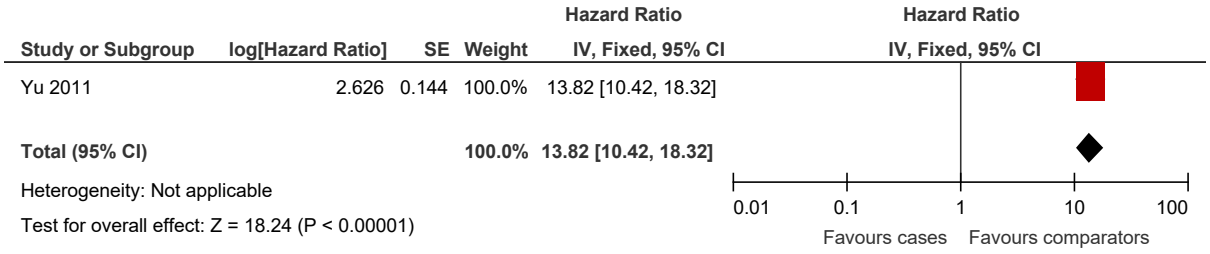


Figure 9: Renal stones

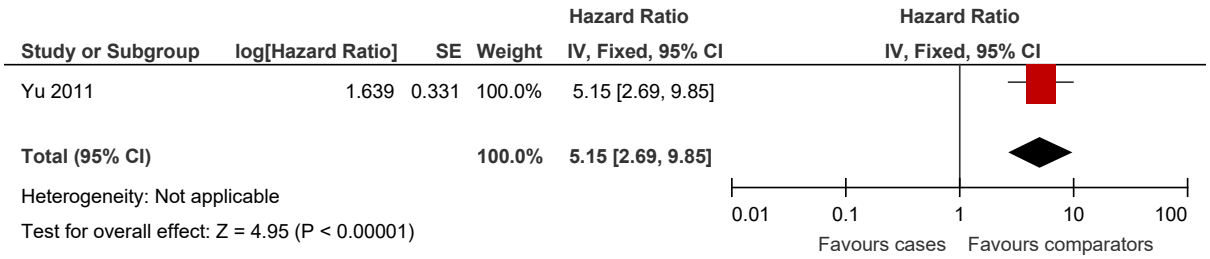


Figure 10: All fractures

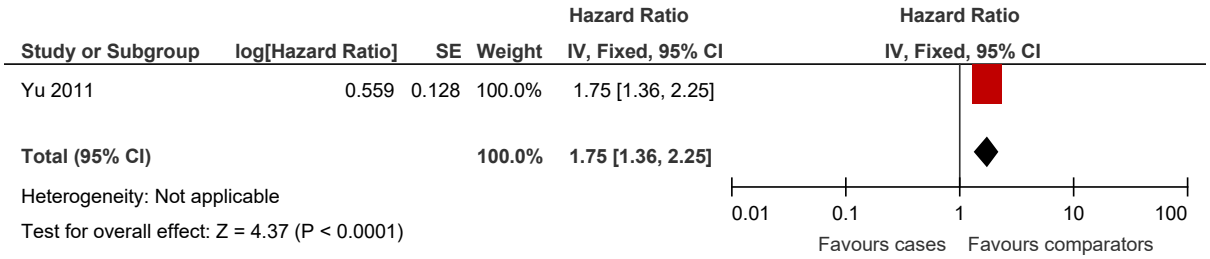
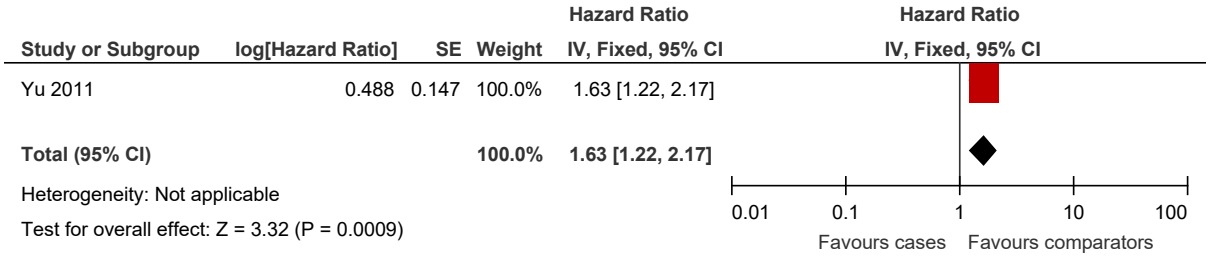
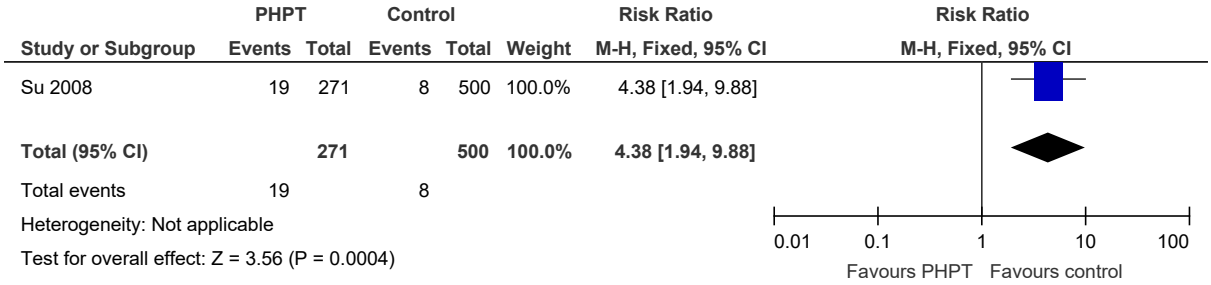


Figure 11: Osteoporotic fractures



E.2 PHPT versus control (no multivariate analysis) Stratum-Pre-surgery

Figure 12: Renal stones



E.3 PHPT cases versus control (no multivariate analysis)- Strata – Pre-operative and post-operative (surgery and non-surgery patients)

Figure 13: Fractures (all)

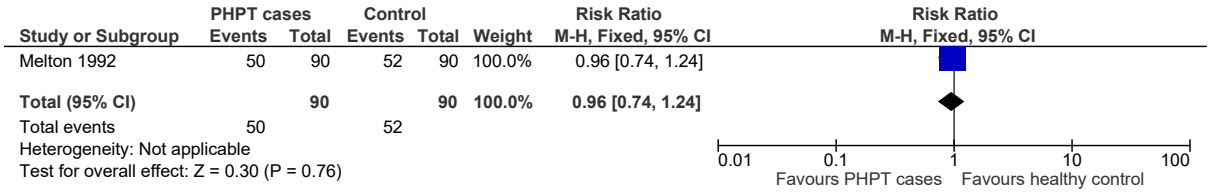


Figure 14: Lumbar spine-BMD (gm/cm^2) (mild PHPT versus healthy women)

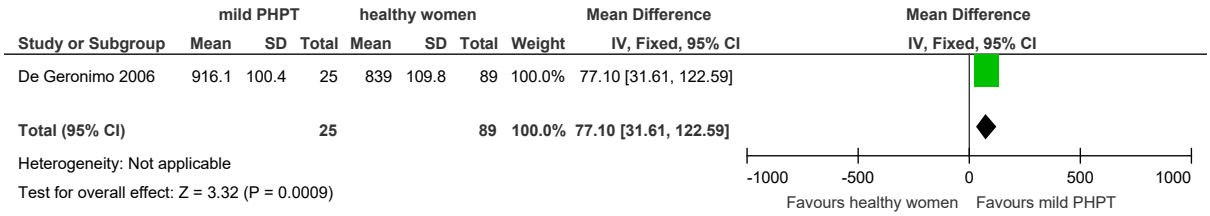


Figure 15: FN BMD (mg/cm^2) (mild PHPT versus healthy women)

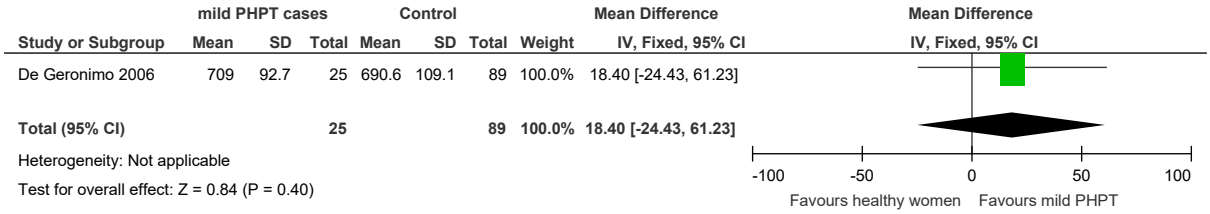


Figure 16: FT BMD (mg/cm^2) (mild PHPT versus healthy women)

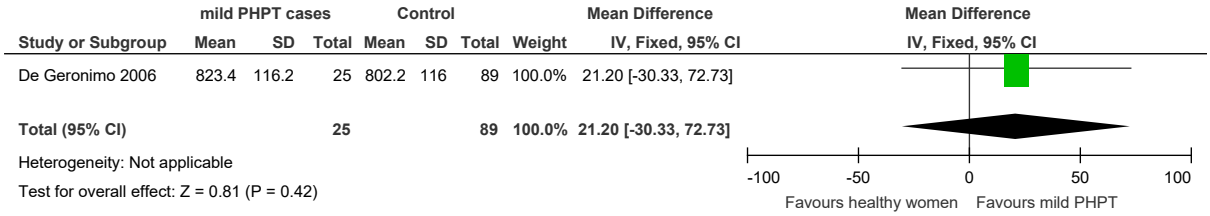


Figure 17: Vertebral fractures (mild PHPT versus control)

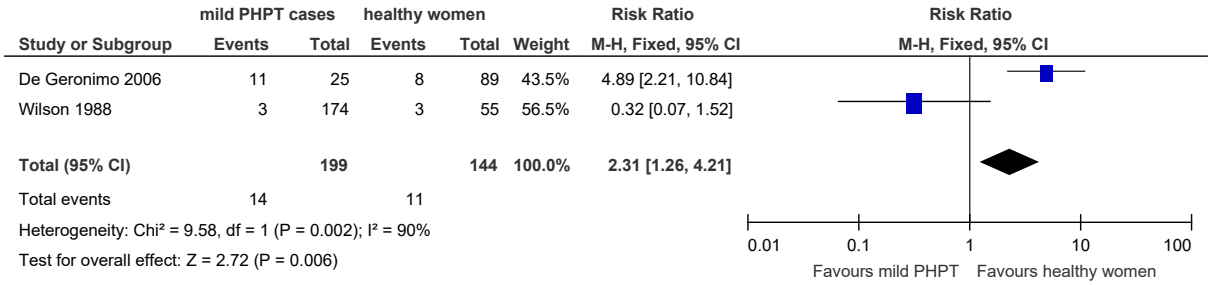


Figure 18: Vertebral fractures (non-mild PHPT versus healthy women)

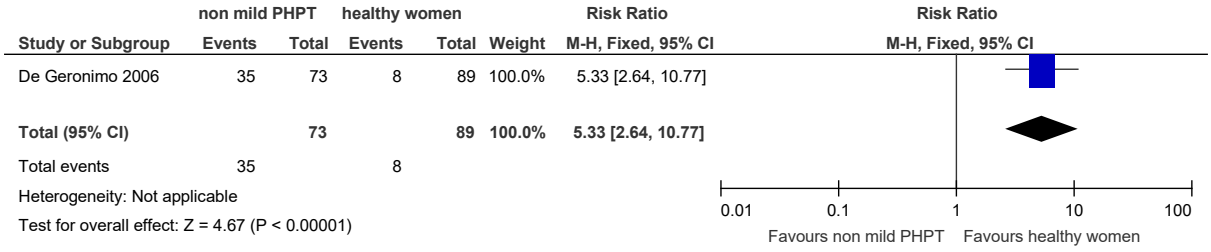


Figure 19: Non-vertebral fractures (mild PHPT versus healthy women)

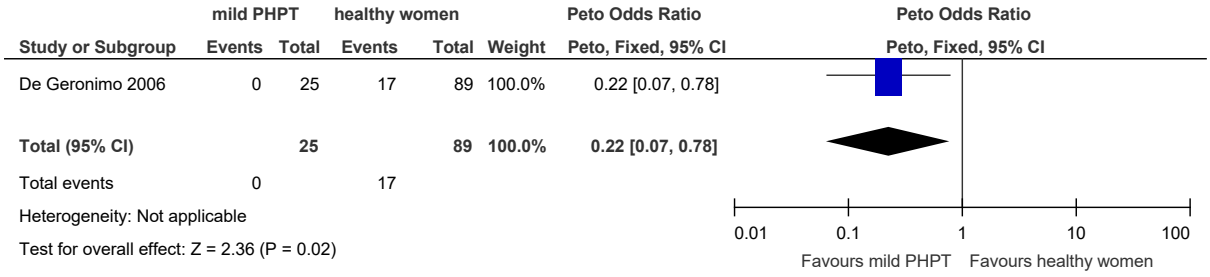


Figure 20: Non-vertebral fractures (non-mild PHPT versus healthy women)

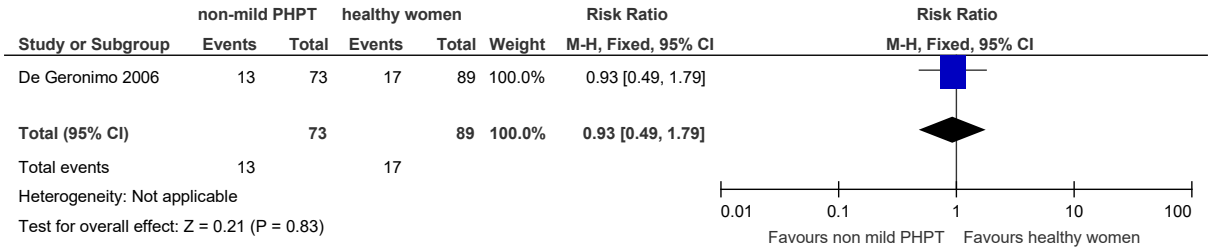


Figure 21: LS-BMD (mg/cm^2) (non-mild PHPT versus healthy women)

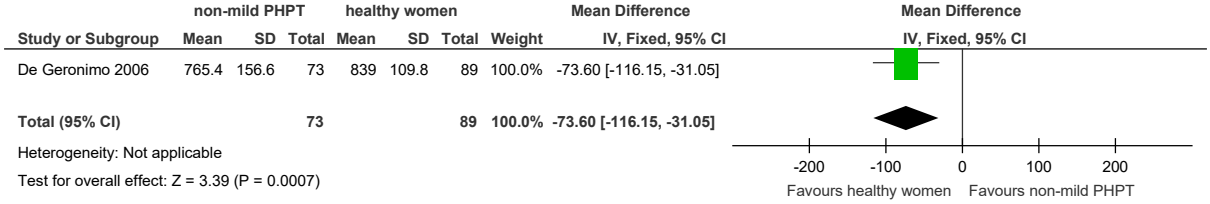


Figure 22: Femur neck-BMD (mg/cm^2) (non-mild PHPT versus healthy women)

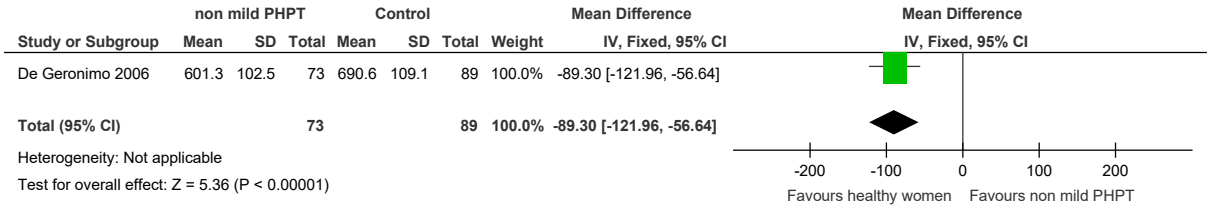
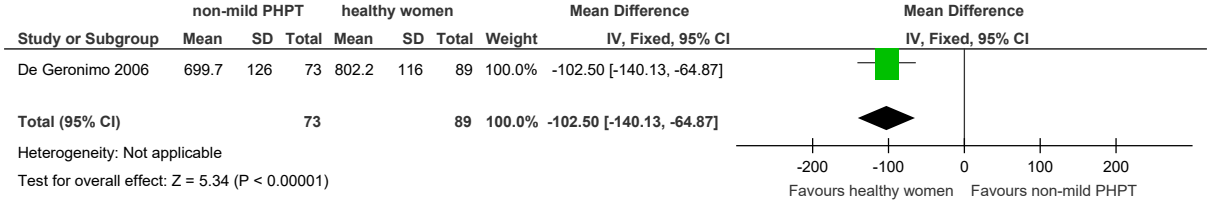


Figure 23: Total femur-BMD (mg/cm^2) (non-mild PHPT versus healthy women)



E.4 PHPT cases versus control (no multivariate analysis) – strata post-operative

Figure 24: Mortality

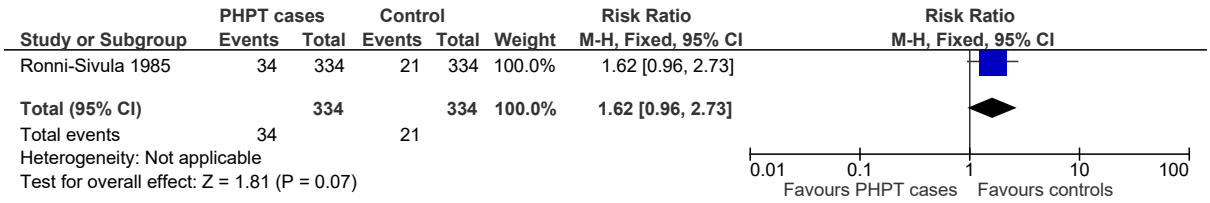
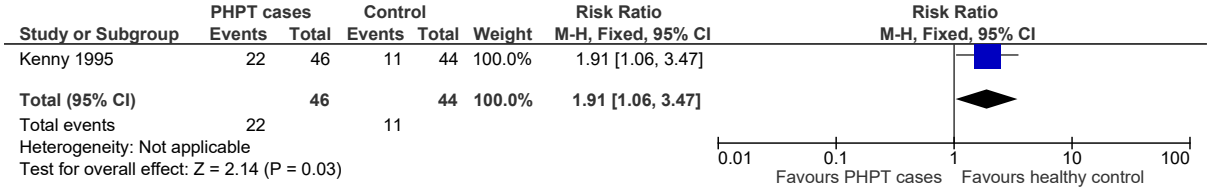


Figure 25: Fractures



Appendix F: GRADE tables

Table 17: Clinical evidence profile: PHPT cases versus matched comparators (adjusted for key confounders) – non-surgical

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Cases	Control	Relative (95% CI)	Absolute		
All-cause mortality (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 1.64 (1.43 to 1.87)[b]	-[c]	⊕⊕OO LOW	CRITICAL
Fatal CVD (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 1.64 (1.32 to 2.04)[b]	-[c]	⊕⊕OO LOW	CRITICAL
Non-fatal CVD (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 2.48 (2.13 to 2.89)[b]	-[c]	⊕⊕OO LOW	CRITICAL
Hypertension (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 2.60 (2.04 to 3.31)[b]	-[c]	⊕⊕OO LOW	CRITICAL
Cerebrovascular disease (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 2.51 (1.95 to 3.22)[b]	-[c]	⊕⊕OO LOW	CRITICAL
Renal failure (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 13.83 (10.41 to 18.37)[b]	-[c]	⊕⊕OO LOW	CRITICAL

Renal stones (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 5.15 (2.69 to 9.83)[b]	-[c]	⊕⊕OO LOW	CRITICAL
All fractures (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 1.75 (1.36 to 2.26)[b]	-[c]	⊕⊕OO LOW	CRITICAL
Osteoporotic fractures (follow-up median 2.9 years)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	NR	NR	HR 1.63 (1.22 to 2.19)[b]	-[c]	⊕⊕OO LOW	CRITICAL

[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] Confounding covariates considered were multiple deprivation index, history of bisphosphonates prescription, history of hospital admitted CVD, cerebrovascular disease, hypertension, renal failure, renal stones, psychiatric disease, fractures, cancer and diabetes.
[c] Absolute effect cannot be calculated as adjusted control group event rate not reported.
* All patients with diagnosed but untreated, mild asymptomatic PHPT.

Table 18: Clinical evidence profile: PHPT versus control (no multivariate analysis) – Pre-operative

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**PHPT**	**Control**	**Relative (95% CI)**	**Absolute**		
Renal stones												
1	observational studies	very serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	19/271 (7%)	1.6%	RR 4.38 (1.94 to 9.88)	54 more per 1000 (from 15 more to 142 more)	⊕OOO VERY LOW	CRITICAL

[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.

Table 19: Clinical evidence profile: PHPT versus control (no multivariate analysis) – Mixed pre-operative and post-operative (surgical and non-surgical patients)

Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PHPT	CONTROL	Relative (95% CI)	Absolute		
Fracture												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	50/90 (55.6%)	41.4%	RR 0.96 (0.74 to 1.24)	17 fewer per 1000 (from 108 fewer to 99 more)	⊕OOO VERY LOW	CRITICAL
Lumbar spine BMD (mg/cm^2) (mild PHPT vs healthy women) (Better indicated by higher values)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	25	89	-	MD 77.1 higher (31.61 to 122.59 higher)	⊕OOO VERY LOW	CRITICAL
Femoral neck- BMD (mg/cm^2) (mild PHPT vs healthy women) (Better indicated by higher values)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	None	25	89	-	MD 18.4 higher (24.43 lower to 61.23 higher)	⊕OOO VERY LOW	CRITICAL
Total femur- BMD (mg/cm^2) (mild PHPT vs healthy women) (Better indicated by higher values)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	Serious[c]	None	25	89	-	MD 21.2 higher (30.33 lower to 72.73 higher)	⊕OOO VERY LOW	CRITICAL
Vertebral fractures (mild PHPT vs control)												
2	observational studies	Serious[a]	very serious[b]	no serious indirectness	no serious imprecision	None	14/199 (7%)	7.2%	RR 2.31 (1.26 to 4.21)	94 more per 1000 (from 19 more to 231 more)	⊕OOO VERY LOW	CRITICAL

Vertebral fractures (non-mild PHPT vs healthy women)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	35/73 (47.9%)	9%	RR 5.33 (2.64 to 10.77)	390 more per 1000 (from 148 more to 879 more)	⊕OOO VERY LOW	CRITICAL
Non-vertebral fractures (mild PHPT vs healthy women)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	None	0/25 (0%)	19.1%	Peto OR 0.22 (0.07 to 0.78)	190 fewer per 1000 (from 290 to 90 fewer)	⊕OOO VERY LOW	CRITICAL
Non-vertebral fractures (non-mild PHPT vs healthy women)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	very serious[c]	None	13/73 (17.8%)	19.1%	RR 0.93 (0.49 to 1.79)	13 fewer per 1000 (from 97 fewer to 151 more)	⊕OOO VERY LOW	CRITICAL
Lumbar spine-BMD (mg/cm^2) (non-mild PHPT vs healthy women) (Better indicated by higher values)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	73	89	-	MD 73.6 lower (116.15 to 31.05 lower)	⊕OOO VERY LOW	CRITICAL
Femoral neck-BMD (mg/cm^2) (non-mild PHPT vs healthy women) (Better indicated by higher values)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	73	89	-	MD 89.3 lower (121.96 to 56.64 lower)	⊕OOO VERY LOW	CRITICAL
Total femur-BMD (mg/cm^2) (non-mild PHPT vs healthy women) (Better indicated by higher values)												
1	observational studies	Serious[a]	no serious inconsistency	no serious indirectness	no serious imprecision	None	73	89	-	MD 102.5 lower (140.13 to 64.87 lower)	⊕OOO VERY LOW	CRITICAL

[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] Heterogeneity, I^2=90%
[c] Downgraded by 1 increment if the confidence interval crossed 1 MID, and downgraded by 2 increments if the confidence interval crossed both MIDs.

Table 20: Clinical evidence profile: PHPT versus control (no multivariate analysis) – strata post-operative

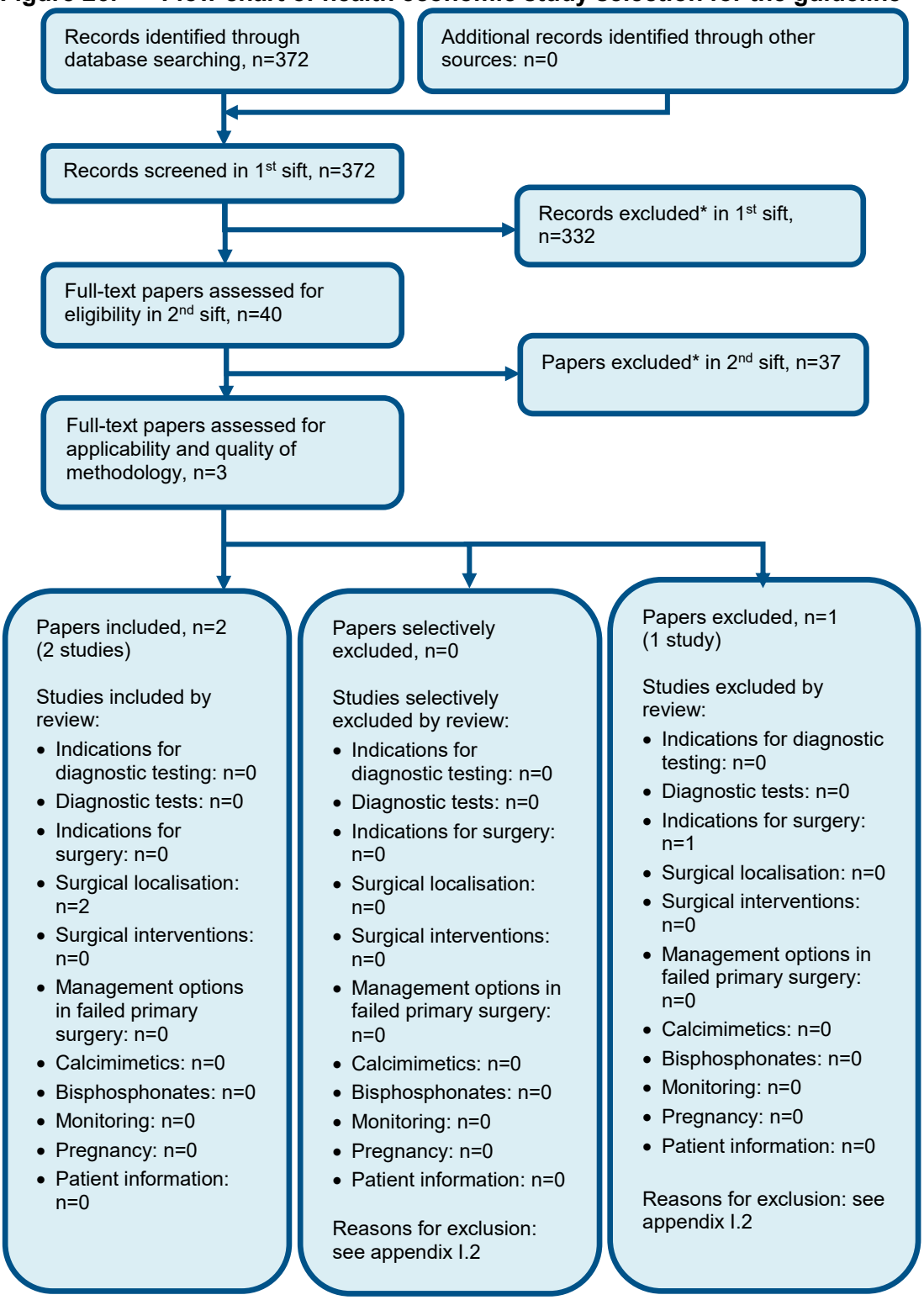
Quality assessment							No of patients		Effect		Quality	Importance
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	PHPT	Control	Relative (95% CI)	Absolute		
Mortality												
1	observational studies	very serious[a]	no serious inconsistency	no serious indirectness	serious imprecision[b]	None	34/334 (10.2%)	6.3%	RR 1.62 (0.96 to 2.73)	39 more per 1000 (from 3 fewer to 109 more)	⊕OOO VERY LOW	CRITICAL
Fracture												
1	observational studies	very serious[a]	no serious inconsistency	no serious indirectness	serious imprecision[b]	None	22/46 (47.8%)	25%	RR 1.91 (1.06 to 3.47)	227 more (from 15 more to 618 more)	⊕OOO VERY LOW	CRITICAL

[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.

Appendix G: Health economic evidence selection

Figure 26: 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

No economic studies were included in this review.

Appendix I: Excluded studies

I.1 Excluded clinical studies

Table 21: Studies excluded from the clinical review

Study	Exclusion reason
Abdulkader 2012[1]	Conference abstract
Agarwal 2003[2]	Incorrect study design. Case report.
Ahsan 2017[3]	n=25. Excluding studies less than 50 participants.
Alvarez-Allende 2014[4]	Conference abstract
Amaral 2012[5]	Inappropriate comparison. Study compared the clinical and laboratory data between the normocalcaemic and mild hypercalcaemic patients.
Antonelli 2011[6]	Conference abstract
Babey 2010[7]	Conference abstract
Bai 2012[8]	Incorrect study design – literature review to explore association between primary hyperparathyroidism (PHPT) and acute or chronic pancreatitis.
Bailey 1974[9]	Incorrect study design
Bandeira 2009[10]	Inappropriate comparison. Study aims to determine the prevalence of cortical osteoporosis in patients with symptomatic PHPT and compare it with the asymptomatic form.
Bandeira 2016[11]	Conference abstract
Bao 2013[12]	Conference abstract
Battersby 1969[13]	Incorrect study design – case report (of pancreatitis with PHPT)
Beard 1950[14]	Incorrect study design – case series
Bhadada 2018[15]	Non-comparative study
Bonzelaar 2016[17]	Conference abstract
Cannon 2010[18]	Inappropriate comparison. Study describes the surgical outcome and long term results of hypercalcaemic crisis patients after parathyroidectomy compared to non-crisis patients.
Carnaille 1998[19]	Incorrect comparison. Study looked at association of pancreatitis with PHPT.
Cassibba 2014[20]	Incorrect study design – retrospective analysis of a case series
Corlew 1985[22]	n=47. Excluding studies less than 50 participants.
Csupor 2005[23]	Inappropriate comparison. Study aimed to assess the potential association between the surgically confirmed location of the disease and the presence of kidney stone.
Danzi 1974[25]	Incorrect study design – case report.
Deaconson 1987[27]	Inappropriate population group. Study reports the influence of parathyroidectomy on the natural history of nephrolithiasis and changes in the rates of new stone formation.
Diaz de la Guardia 2010[30]	Not in English
Dimkovic 2002[31]	Inappropriate population. Study aimed to examine patients with kidney stone disease, elevated iPTH, but normal serum calcium level and normal urinary excretion of calcium.
Dolgin 1979[32]	Study analysed the effect of routine screening of calcium and phosphate levels on the incidence and spectrum of PHPT. No useable outcomes.

Study	Exclusion reason
Dumitrescu 2008[33]	Incorrect population. Study aimed to determine the prevalence of contributors to secondary osteoporosis in patients presenting with a clinical vertebral or non-vertebral fracture.
Eufrazino 2013[34]	Incorrect study design – cross-sectional study
Falko 1984[35]	No comparison group. Study assessed clinical and biochemical spectrum of patients with PHPT who had surgery.
Heath 1991[37]	Incorrect study design – case series
Hedback 2002[39]	Incorrect study design – case series
Jha 2016 [40]	Non-comparative study
Kobayashi 1997 [43]	Non-comparative study
Larsson 1989[44]	No useable outcomes
Lowe 2007[46]	No comparison group. Study described the clinical course of 37 patients with normocalcaemic PHPT who were followed for up to 8 years.
Lueg 1982[47]	Incorrect study design – case series
Marques 2011[48]	Incorrect study design. Retrospective review of medical records to describe the characteristics of normocalcaemic primary hyperparathyroidism (NPHPT) in patients seen for osteoporosis evaluation.
Misiorowski 2012[50]	No useable outcomes. The aim of the study was to evaluate the diagnostic power of the bone densitometry in diagnosis of PHPT.
Mollerup 1999[51]	Inappropriate comparison – before and after surgery. The study aimed to evaluate the risk of renal stone recurrence after successful surgical treatment of PHPT.
Nilsson 2005[55]	Inappropriate population and outcomes. Study explored long-term effects of parathyroidectomy on cardiovascular functions in PHPT.
Pradeep 2008[56]	Non-comparative study
Pratley 1973[57]	Incorrect study design – case series.
Purnell 1971 [58]	Non-comparative study
Rajeevan 2014[59]	Incorrect study design – series review
Rubin 2008[61]	Inappropriate comparison. Study compared PHPT patients who had undergone surgery vs those without surgery.
Scholz 1981[62]	Non-comparative study
Siilin 2011[63]	Study assessed BMD between PHPT and men without PHPT. No clinical outcomes.
Silverberg 1990[65]	No comparison group
Silverberg 1995[64]	Non-comparative study
Siminovitch 1980[66]	Study assessed the effect of parathyroidectomy in patients with normocalcaemic calcium stones. No useable outcomes.
Soreide 1997[67]	Inappropriate comparison. The study evaluated survival after surgical treatment for primary hyperparathyroidism.
Strewler 1995[68]	Literature review. Screened for references.
Turchi 1962[70]	Incorrect study design – case report.
Vanderwalde 2006[71]	Study aimed to determine the effect of parathyroidectomy on fracture risk in patients with PHPT. Inappropriate comparison-comparison groups were parathyroidectomy vs observation.
Vanderwalde 2009[72]	Inappropriate comparison – comparison groups were parathyroidectomy vs observation
Vestergaard 2000[73]	Inappropriate comparison
Vestergaard 2003[75]	Study included in surgery review

Study	Exclusion reason
Vestergaard 2003[76]	Study included in surgery review
Vestergaard 2003[74]	Inappropriate comparison. The aim of this study was to evaluate cardiovascular morbidity before and after surgery for PHPT.
Vestergaard 2004[77]	Inappropriate comparison
Wermers 1998 [78]	Non-comparative study
Yu 2009[81]	Study did not meet protocol criteria. Study evaluated prevalence and incidence of PHPT.
Yu 2011[83]	No protocol outcomes. Study provided information on the natural history of asymptomatic 'mild' PHPT patients with a long follow-up period, in terms of the biochemical progression of the disease.
Yu 2013[82]	No useable outcomes. Study aimed to identify the best biochemical risk factors for predicting adverse outcomes in untreated PHPT.

I.2 Excluded health economic studies

None.

Appendix J: Research recommendations

J.1 Long-term consequences of management strategies for PHPT

Research question: What are the long-term outcomes of different management strategies for primary hyperparathyroidism? What strategies are most cost-effective?

Why this is important:

There is limited evidence on the long-term outcomes of the different management strategies such as surgery, calcimimetics and bisphosphonates (see evidence report C, evidence report G and evidence report H) . In order for people to make an informed choice regarding their treatment research is needed on this topic.

Criteria for selecting high-priority research recommendations:

PICO question	Population: People diagnosed with PHPT Intervention(s): Surgery, calcimimetics and bisphosphonates (or any combination, no treatment (surveillance/conservative management) Comparison: Compared to each other Outcome(s): Health-related quality of life, mortality, preservation of end organ function (bone mineral density, fractures, renal stones and renal function), persistent hypercalcaemia, cardiovascular events, adverse events, cancer incidence.
Importance to patients or the population	There is limited evidence on the long-term outcomes of the different management strategies. In order for people to make an informed choice regarding their treatment research is needed on this topic.
Relevance to NICE guidance	Limited recommendations were made on surgery, calcimimetics and bisphosphonates.
Relevance to the NHS	Altered guidance may have a financial impact on the NHS depending on whether the evidence either supports or does not support certain management strategies.
National priorities	None
Current evidence base	The majority of randomised evidence compared parathyroidectomy versus conservative management. The longest follow up period was 5 years. One very small study reported at 17 years but only a very limited number of outcomes were reported.
Equality	None
Study design	RCT
Feasibility	20-year follow up period to ensure all patient outcomes are captured.
Other comments	None
Importance	• High: the research is essential to inform future updates of key recommendations in the guideline.