AMERICAN UNIVERSITY OF BEIRUT

EFFECTS OF GROWTH HORMONE THERAPY ON BONE IN ADULTS WITH OSTEOPENIA OR OSTEOPOROSIS AND WITHOUT GROWTH HORMONE DEFICIENCY: A SYSTEMATIC REVIEW AND META-ANALYSIS

by

MAYA GEORGES BARAKE

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Sciences in Health Research
to the Scholars in HeAlth Research Program (SHARP)
of the Faculty of Health Sciences and the Faculty of Medicine
at the American University of Beirut

Beirut, Lebanon
April 2016

AMERICAN UNIVERSITY OF BEIRUT

EFFECTS OF GROWTH HORMONE THERAPY ON BONE IN ADULTS WITH OSTEOPENIA OR OSTEOPOROSIS AND WITHOUT GROWTH HORMONE DEFICIENCY: A SYSTEMATIC REVIEW AND META-ANALYSIS

by

MAYA BARAKE

Approved by:

Dr. Asma Arabi, Associate Professor
Internal Medicine, FM — Advisor

AA on behalf of Dr Elie AKL

Dr. Elie Akl, Associate Professor
Internal medicine, FM — Member of Committee

Dr. Ghada El Hajj Fuleihan, Professor
Internal medicine, FM — Member of Committee

AA on behalf of Dr Tritos

Dr. Nicholas A. Tritos, Assistant Professor
Neuroendocrine Unit, Department of Medicine
Harvard Medical School — Member of Committee

Date of thesis defense: April 19, 2016

AMERICAN UNIVERSITY OF BEIRUT

THESIS, DISSERTATION, PROJECT RELEASE FORM

Student Name: ___BARAKE_____MAYA_____GEORGES_____
Last First Middle

◉ Master's Thesis ○ Master's Project ○ Doctoral Dissertation

☑ I authorize the American University of Beirut to: (a) reproduce hard or electronic copies of my thesis, dissertation, or project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

☐ I authorize the American University of Beirut, **three years after the date of submitting my thesis, dissertation, or project,** to: (a) reproduce hard or electronic copies of it; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes.

MAYA BARAKE May 4, 2016

Signature Date

ACKNOWLEDGMENTS

I would like to offer my sincerest gratitude and recognition to my advisor, Dr. Asma Arabi and my co-advisor Dr. Nicholas Tritos for their continuous guidance and support throughout the preparation of this thesis.

Special thanks to my thesis committee members, Dr Elie Akl and Dr. Ghada El Hajj Fuleihan for their feedback, care and encouragement.

Special thanks to my colleagues, Drs Sarah El Ghandour and Nancy Nakhoul for their help and support.

AN ABSTRACT OF THE THESIS OF

Maya Barake for Master of Sciences
Major: Health Research (SHARP)

Title: Effects of Growth Hormone Therapy on Bone in Adults with Osteopenia or Osteoporosis and Without Growth Hormone Deficiency: A Systematic Review and Meta-Analysis

Background: Osteoporosis is a metabolic bone disease that constitutes both a significant personal burden as well as a major public health concern. Several alternatives are available to treat this disease through decrease in bone resorption. However, options are much more limited with regards to anabolic agents. Growth hormone (GH) is a peptide hormone normally produced by the pituitary gland. Studies have shown that GH plays an important role in bone metabolism. In adults, growth hormone deficiency (GHD) has been associated with low bone density and increased fracture risk, an effect that is counteracted by growth hormone replacement. Whether GH treatment can result in a similar benefit in adults with age-related bone loss, who presumably have age-related decline in GH, is unanswered.

Objectives: The objectives of this systematic review and meta-analysis are: (1) Examine the effect of growth hormone therapy on bone densitometric endpoints, bone turnover markers and fracture risk in adults with osteopenia or osteoporosis and no organic growth hormone deficiency; (2) Evaluate the safety of growth hormone therapy in this population of interest.

Search methodology: A systematic search of the existing literature was conducted using Medline, Embase and the Cochrane Register of Controlled Trials, without any time, language or study size restriction; the search was updated in November 2015. A manual search of the references of both original articles collected and pertinent review articles on the topic was also conducted.

Eligibility criteria: We included prospective controlled trials conducted in postmenopausal women and men above the age of 50 years, with age-related osteopenia or osteoporosis and without organic growth hormone deficiency, in whom treatment with GH was used as compared to placebo or a comparator for at least six months.

Data collection and analysis: Included trials were reviewed and data was collected from them as preplanned by two independent reviewers. We assessed risk of bias in retained randomized trials using the Cochrane risk of bias tool in duplicate. For outcomes covered by at least 2 trials, we synthesized data by meta-analysis. We calculated the weighted mean difference (WMD) and 95% Confidence Interval (CI) for bone mineral density (BMD), bone mineral content (BMC), bone turnover markers levels reached with GH treatment as compared to a comparator; and the risk ratio and 95% CI to develop fractures and vertebral fractures. Analysis was done using RevMan version 5.3. We reviewed in all available studies mortality and reported adverse events.

Results: We included in this systematic review seven studies and one follow-up trial of an earlier study. Outcomes of interest were covered as follows: 3 trials reported BMD by dual x-ray absorptiometry (DXA), 2 BMC by DXA, 2 BMC by single-photon absorptiometry (SPA), 5 bone turnover markers and 4 reported fractures. Included studies only involved women, which were characterized by severe osteoporosis and were treated with variable GH dosing regimens given along with other osteoporosis therapies for a time period extending between 6 to 24 months.

No significant difference was obtained through treatment with GH as compared to control in BMD at the lumbar spine (WMD -0.01 [-0.04, 0.02]), the total hip (WMD 0 [-0.05, 0.06]) and the femoral neck (WMD 0 [-0.03, 0.04]).
Similarly, we found no significant difference with GH therapy versus control in BMC at the lumbar spine (WMD -0.71 [-1.63, 0.22]) and the femoral neck (WMD -0.06 [-0.28, 0.16]). We also found no difference for BMC measured by SPA at the forearm (WMD -0.06 [-0.23, 0.10]).

The effect of GH therapy on 4 bone turnover markers was combined separately in meta-analysis. We obtained a significant increase in the bone formation marker procollagen type-I carboxy-terminal propeptide (PICP) with GH therapy as compared to control (WMD 14.03 [2.68, 25.38]). As for the effect on osteocalcin, another formation marker, and on the bone resorption markers total urinary pyridinolines and carboxy-terminal collagen crosslinks (CTX), a non-significant trend to favor GH versus control was observed.

We obtained a significant decrease in the risk of fractures with GH treatment as compared to control (RR=0.63 [0.46, 0.87]) with no significant change in the occurrence of vertebral fractures specifically.

Adverse events reported during growth hormone therapy included the well-recognized side effects of GH related to fluid retention, such as peripheral edema, musculoskeletal pain and carpal tunnel syndrome. They were usually transient and reversible, often relieved by decreasing the treatment dose. No mortality was reported during the study period in included studies. In the trial with 10-year follow-up no significant difference in mortality was obtained with GH when compared to a control population.

Conclusion: Growth hormone therapy may not be effective in improving bone density in women with age-related bone loss. It may, however, decrease fracture risk, without major adverse events. There is a need to further explore these findings along with the impact of GH on bone quality in randomized controlled trials using GH therapy for extended periods in both men and women with age-related bone loss.

CONTENTS

ILLUSTRATIONS

Page

LIST OF TABLES

Page

LIST OF ABBREVIATIONS

GH	Growth Hormone
GHD	Growth Hormone Deficiency
WMD	Weighted Mean Difference
CI	Confidence Interval
BMD	Bone Mineral Density
BMC	Bone Mineral Content
DXA	Dual-energy X-ray Absorptiometry
SPA	Single Photon Absorptiometry
PICP	Procollagen type-I Carboxy-terminal Propeptide
CTX	C-telopeptide cross-link of type-I collagen
WHO	World Health Organization
SD	Standard Deviation
GHRH	Growth Hormone Releasing Hormone
GHBP	Growth Hormone Binding Protein
GHR	Growth Hormone Replacement
IGF-I	Insulin-like Growth Factor I
IGFBP	IGF-Binding Protein
TNF-∝	Tumor Necrosis Factor Alpha
IL-6	Interleukin-6
IL-1	Interleukin-1
RANK-L	Receptor Activator of Nuclear Factor KB
BMI	Body Mass Index

rhGH	Recombinant Human Growth Hormone
GHR	Growth Hormone Replacement
HR	Hazard Ratio
PRISMA	Preferred Reporting Items for Sytematic reviews and Meta-Analyses
COMET	Core Outcome Measures in Effectiveness Trials
DPA	Dual Photon Absorptiometry
OC	Osteocalcin
CICP	C-terminal Propeptide of type-I Collagen
CT	Calcitonin
Chi^2	Chi-squared test
GRADE	Grading of Recommendations, Assessment, Development, and Evaluation

CHAPTER 1

INTRODUCTION

1.1 Osteoporosis

1.1.1. Definition and disease burden

Osteoporosis is a skeletal disorder characterized by compromised bone strength leading to increased risk of fractures. Osteoporosis is diagnosed by the presence of fragility fractures and/or low bone mineral density (BMD) measured by dual-energy x-ray absorptiometry (DXA) scans (1,2). DXA measures both bone mineral content (BMC) and bone area, and calculates from them areal BMD. According to the World Health Organization (WHO), osteoporosis is diagnosed in postmenopausal women and men aged more than 50 when the BMD measured by DXA is equal to or more than 2.5 standard deviations (SDs) below the BMD of a young adult reference population. When the BMD is between 1 and 2.5 SDs below the reference, the condition is termed osteopenia (3). The presence of low bone mineral density is associated with increased fracture risk (4). The prevalence of the disease increases with age, namely secondary to the decline in bone-protective sex hormones and constitutes both a significant personal burden as well as a major public health concern. After the age of 50, one in three women and one in five men are projected to develop an osteoporosis-related fracture (1,5,6,7).

1.1.2 Available therapies

In adults, bone is constantly being remodeled. Bone remodeling consists of two phases: bone resorption (where mature bone tissue is removed from the skeleton) and

bone formation (where bone is being rebuilt). Both bone formation and bone resorption are normally tightly coupled so that bone mass does not change. Osteoporosis occurs when resorption exceeds formation.

A number of therapeutic agents have been approved for the treatment of osteoporosis and the prevention of fractures. Current FDA-approved pharmacologic therapies include anti-resorptive medications that work by reducing bone resorption, including bisphosphonates, denosumab, calcitonin, estrogen and selective estrogen receptor modulators and one anabolic agent, teriparatide that result in increased bone formation (8). While available drugs have been successful in reducing fracture risk by 20 to 70%, depending on the selected site and the medication used, however, a larger number of fractures are still not prevented due to limitation in the efficacy of the current therapies (9). Bisphosphonates, which form the mainstay of therapy, have a prolonged tissue half-life in bone and a potential for side effects (10). And while several alternatives are available to decrease bone resorption, options are much more limited with regards to bone formation. There is thus a need for new therapeutic options for the treatment of osteoporosis (9).

1.2 Growth hormone physiology

Growth hormone (GH) is a peptide hormone produced by the anterior pituitary. It is secreted in a pulsatile fashion, with the majority of daily secretion occurring during slow wave sleep (11). Growth hormone release is stimulated by growth hormone releasing hormone (GHRH) produced by the hypothalamus and by ghrelin formed mainly in the gastrointestinal tract (stomach), and inhibited by centrally produced somatostatin (12). Growth hormone is available in the circulation partially bound to two

different GH binding proteins (GHBPs), one of which being homologous to the extracellular portion of the GH receptor (GHR) (13,14). The latter is down regulated by abdominal obesity (15).

Growth hormone works directly by binding to GHRs available in various peripheral tissues. It also acts indirectly through production of insulin-like growth factor 1 (IGF-1) in the liver and paracrine IGF-1 secretion in several tissues, including bone. IGF-1 circulates bound to IGF-binding proteins (IGFBPs), IGFBP-3 being the principal binding protein. Serum levels of IGF-1 are maintained relatively stable throughout the day in the circulation. They are thus often used as a reliable marker of overall GH action (16). IGF-1 is also locally produced in various tissues where it exerts paracrine and autocrine actions (17). Gonadal steroids have an important influence on IGF-1 secretion. Studies done in healthy volunteers show that for the same IGF-1 level, 24-hour GH profiles are approximately three times higher in women then in men (18). Oral estrogen administration attenuates IGF-1 production through hepatic effects (19). An inverse relation has also been observed between adiposity and circulating IGF-1 (20).

Growth hormone secretion is maintained throughout life. Circulating levels decrease, however, with aging. GH release is increased 2-3 times during puberty. Levels decrease thereafter with increasing age, with a parallel decrease in IGF-1 levels (21,22,23). The age-related decline in GH release is associated with increased somatostatin levels and is, at least, partially related to a decrease in hypothalamic GHRH production (24).

1.3 Growth hormone and bone

Growth hormone exerts a direct effect on skeletal muscle and bone. It also affects bone indirectly through systemic IGF-1 (produced in the liver) and locally produced IGF-1 in osteoblasts (12, 25). Both GH and IGF-1 receptors are expressed in bone where they can mediate GH action (26).

1.3.1 In-vitro studies

In-vitro, GH addition to cultured bone cells increases activation and differentiation of osteoblasts, cells that are responsible for bone formation (27, 28). Similarly, IGF-1 promotes osteoblastogenesis and reduces osteoblast apoptosis (29). Addition of GH to cultured chondrocytes or cartilage cells has a direct stimulatory effect on their synthesis at the pre-chondrocyte stage. IGF-1 similarly stimulates chondrocytes, albeit at a later stage of maturation (30). Growth hormone and IGF-1 are thus potentially anabolic to bone, both independently and synergistically. The effect of IGF-1 on osteoblasts is increased in the presence of GH and IGFBP-3. IGF-1 alone is unable to substitute for GH action (31).

GH also exerts a stimulatory effect on bone resorption. It promotes production of pro-inflammatory cytokines, such as Tumor necrosis factor alpha (TNF-α), Interleukin 6 (IL-6) and Interleukin 1 beta (IL-1 beta) that incite formation of osteoclasts, cells responsible of bone resorption (32, 33). In-vitro, IGF-1 similarly promotes bone resorption through induction of the ligand of the receptor activator of nuclear factor κB (RANK-L) synthesis (34).

1.3.2 Animal studies

The anabolic action of GH on bone was demonstrated many years ago in dogs, wherein the administration of GH resulted in increased skeletal mass (35). In mice, knockout of the GH receptor resulted in decreased bone turnover, reduced cortical bone and preserved trabecular bone. IGF-1 treatment in these animals almost completely restored their bone structure (36). Similarly, deletion of liver-specific IGF-1 resulted in decreased cortical bone (37). While these studies illustrated the role of systemic IGF-1 in maintaining cortical bone integrity in mice, targeted knockout of local IGF-1 receptor in osteoblasts led to decreased trabecular bone, indicating a critical paracrine role of IGF-1 in the preservation of cancellous bone in mice (38). In-vivo studies in mice also illustrate the importance of both GH and IGF-1 in skeletal growth. Mouth mutants lacking both GHR and IGF-1 show more severe growth retardation as compared to mice with either deficiency alone (39).

1.3.3 Humans: Growth hormone deficiency and bone

In humans, the role of GH in bone physiology is illustrated in patients with growth hormone deficiency (GHD). In children and adolescents, GHD results in short stature and decreased BMD (40). In adulthood, after the attainment of final height, GHD is similarly associated with low BMD, along with a decrease in serum and urinary markers of bone turnover (41). The consequences of GHD may differ, however, depending on the time of onset of the deficiency. In a cohort of patients with childhood-onset GHD, total body and lumbar spine BMC and BMD were lower, as compared to a group of patients with adult-onset GHD, matched for age, gender, body mass index (BMI) and number of anterior pituitary hormone deficiencies (42). Bone loss is also

correlated to the severity of GHD, with lower BMD observed in patients with more severe deficiency in GH, defined based on the diagnostic GHRH/Arginine test for GHD (43). The presence of concurrent central hypogonadism with GHD similarly influences bone integrity, whereby patients with untreated hypogonadism have lower BMD than those who are eugonadal (44). Gender has also been discussed as a factor impacting the effect of GHD on bone in adults. When compared to healthy controls, men with GHD had lower BMC and BMD. Women, however, had similar bone densities to their non-GHD controls (45).

Along with the drop in bone content and bone density, the risk of fractures appears to be increased in GHD patients. When compared to a non-GHD control population, adults with GHD had threefold increased fracture prevalence, as assessed by retrospective questionnaires (46). In a large-scale epidemiological survey, using a radiological approach, adult patients with GHD were similarly found to have higher fracture frequency, independent of the presence of additional pituitary deficiencies (47). Using quantitative morphometric analysis, the frequency of vertebral fractures was also found to be significantly higher in GHD patients versus controls (63.6% versus 37.7%) (48). Although fractures occurred more frequently in untreated patients with GHD and low bone density, BMD was not sufficient to identify patients who fractured, as around one half of fractures occurred in individuals with normal BMD. An effect of GHD on bone quality, beyond the effect on bone content and density is thus hypothesized (48). More recently, in a 6-year prospective follow-up study of GHD patients, 30% experienced incident morphometric vertebral fractures, which was correlated with the decrease in lumbar spine BMD (49).

1.4 Therapeutic use of growth hormone

Historically, GH was first derived from cadaveric pituitary glands and was indicated for the treatment of short stature in children with GHD. In 1985, cadaveric human GH was withdrawn due to its association with Creutzfeldt-Jakob disease. Thereafter, recombinant human growth hormone therapy (rhGH) was introduced and became more readily available for clinical use (16).

1.4.1 Effects of growth hormone replacement on bone in GHD

In children and adolescents with GHD, treatment with GH has been clearly correlated with restoration of linear growth, becoming a routine clinical practice for that indication (50,51). Along with the effect on linear growth, growth hormone replacement (GHR) also improved bone mass with subsequent increase in BMD and a fourfold decrease in fracture frequency as compared to untreated controls (40). The beneficial effect of GHR continues through the transition period to adulthood, wherein twenty-four months of GH treatment is associated with a 3.5% increase in lumbar spine BMD as compared to untreated young controls (52). It is thus advised to retest for persistent GH deficiency after completion of linear growth, unless GHD is highly likely, based on the presence of multiple additional pituitary hormone deficiencies or genetic mutations causing GHD, and to continue replacement therapy through young adulthood if deficiency persists (51).

The role of GHR in adults with growth hormone deficiency has been the subject of several studies. GHR is associated with increased bone turnover. Markers of bone formation, such as osteocalcin and bone-specific alkaline phosphatase and of bone resorption, including urinary pyridinolines and deoxypyridinolines increase with GH

treatment and become significantly different from placebo after 3 months of treatment (53). Increased bone turnover results in increased bone remodeling and subsequent expansion of the bone remodeling space. Bone remodeling is a continuous cycle of bone destruction and renewal, whereby individual bone units undergo bone resorption, followed by bone formation, which requires at least 6 months for completion. Since GH increases bone remodeling, bone measurements taken in the first few months of treatment would show decrease in BMC and BMD, illustrating bone with incomplete formation (54). Expectedly, in our previous meta-analysis on GH replacement in adults with GHD, BMC and BMD at the lumbar spine and BMD at the femoral neck, decreased with GHR in randomized studies extending up to one year. The initial drop in bone density is followed by a subsequent time-dependent gain in bone mass, whereby a significant increase in BMD at the lumbar spine and femoral neck is obtained in studies extending for more than 12 months, along with a non-significant increase in total femur BMD. The gain in BMD ranged between 1% and 7% at the spine and 0.6% and 4% at the femoral neck (55). The benefits of growth hormone therapy with regards to bone density are maintained with prolonged replacement therapy. In a 10-year prospective open label study, continuous treatment led to sustained increase in bone mass and density in hypopituitary adults (56). Similarly, in another prospective observational study conducted for 15 years, GHR induced a sustained increase in lumbar spine BMC and BMD by 9% and 5%, respectively. At the femur neck, a peak increase in BMC and BMD of 7% and 3%, respectively, occurred after 7 years. At 15 years, BMD returned to baseline, whereas BMC remained 5% higher than baseline (57). A comparable result was observed in an observational study conducted in the Netherlands for 15 years, in which case GHR resulted in a progressive increase in bone densitometric endpoints over

10 years, with stabilization thereafter (58). GH administration had a persistent effect on BMD, 18 months after withdrawal of initial replacement (59).

A gender-dimorphic effect of GHR on BMD is possible, but remains unclear in the existing literature. Subgroup analysis in the meta-analysis of GHR in GHD reveals a higher response in BMD in men, as compared to women and results reached statistical significance in men only (55). Similarly, in an individual randomized trial on GHR in adults with GHD, men derived a clear benefit with regards to their bone density, whereas women replaced with GH showed no further benefit in their BMD, as compared to placebo (60). In the 10-year prospective study on GHR, women on estrogen replacement had a greater increase in BMD, when compared to women not on estrogen (56). Women also showed less increase in BMD, as compared to men, in the same cohort on longer follow-up for 15 years, despite a higher dose of GH therapy (57).

Whereas a decrease in fracture risk is anticipated with increased BMD, studies addressing fracture end points with GHR are scarce so far and none are randomized controlled clinical trials (55). A cross-sectional analysis of vertebral fracture rates in GHD patients on GHR revealed a lower prevalence of vertebral deformities in treated adults as compared to those not treated (48). More recently, in a 6-year prospective study on incident vertebral fractures in adults with GHD, morphometric fractures occurred more frequently in patients with untreated GHD as compared to those who were treated. Fracture risk increased progressively throughout the 6-year follow-up and was counteracted at all time points by growth hormone treatment (49). Similarly, in a prospective observational uncontrolled cohort study with a mean follow-up of 4.6 years, annual fracture incidence rates were lower in patients on GHR, as compared to those who were not on treatment (hazard ratio [HR] 0.69, 95% CI 0.54-0.88), an effect size

comparable to that of established osteoporosis therapies. The difference in fracture risk, however, was not seen in the subgroup of patients with pre-existing osteoporosis (61,62).

1.4.2 Effects of GH therapy on bone in non-GHD states

Whether GH is able to increase BMD and improve bone quality in adults with osteoporosis without GHD remains an unanswered question. The rationale for such therapy is that GH secretion by the pituitary, often measured as IGF-1 level in the serum, normally declines with aging (21, 22, 23). When compared to women in pre-menopause, postmenopausal women had significantly lower IGF-1 levels (63,64). The contribution of this age-related relative deficiency in GH/IGF-1 to postmenopausal osteoporosis has been the subject of several studies. In a cross-sectional study in women aged 70 and older, serum IGF-I was found to be an independent predictor of BMD at the femur (65). In another cross-sectional trial in postmenopausal women referred for osteoporosis screening, women with osteoporosis had lower IGF-1 levels. IGF-1 was also lower in patients with vertebral fractures, as compared to those without fractures. IGF-1 correlated positively with BMD, accounting for 8.5% of the variance at lumbar spine BMD and for 4.6% of variance at the femoral neck (66). Similar findings were also observed in men. In men with idiopathic osteoporosis, serum IGF-1 levels were significantly lower than healthy controls, with a positive correlation observed between plasma IGF-1 and bone density at the spine and the forearm (67). Similarly, in another study in Swedish men, IGF-1 correlated negatively with age, while BMD in total body, distal radius, and femoral neck was positively correlated with serum IGF-I (68). Other cross-sectional studies in both men and women with osteoporosis similarly reported

lower circulating levels of IGF-1 and/or IGFBPs, as compared to controls, which were positively correlated with BMD (69,70). Studies investigating a link between serum IGF-1 and osteoporotic fractures are more limited. In a prospective study in older men with mean age of 75, serum IGF-1 was inversely associated with risk of all fractures, hip fractures (HR per SD decrease = 1.45, 95% CI 1.07-1.97) and clinical vertebral fractures (HR per SD decrease = 1.40, 95% CI 1.10-1.78). The association between serum IGF-1 and fracture risk was partly mediated via BMD (71). In a prospective study in postmenopausal women, decreased IGF-1 was associated with increased risk of osteoporotic fractures, however independently of BMD (72).

The role of GH treatment in adult patients with osteoporosis and without GHD has been evaluated in a number of studies with conflicting results. Available studies differed in the number of subjects enrolled, duration of follow-up and concomitant treatment with approved therapies for osteoporosis. In the randomized placebo-controlled trial with the longest follow-up of 3 years, GH resulted in increased bone mineral content of 14% in women with postmenopausal osteoporosis as compared to controls (73). In another trial of similar design, GH therapy did not result in a significant change in BMD (74). Available literature is also not clear with regards to the effect of GH therapy on fracture incidence. A systematic review by Yang et al. addressed the effect of GH therapy on bone in 2012, however the included study population was limited to patients with hip fractures and the change in bone mineral content or bone mineral density was addressed as a secondary outcome (75). Another systematic review and meta-analysis studied the effect of GH in healthy elderly with a BMI less than 35 kg/m2, who were not specifically selected to have osteoporosis at study entry. The review addressed multiple outcomes, including BMD that was not

significantly changed with GH (76). The true effect of GH therapy in patients with age-related osteoporosis has thus not yet been established definitively. It is then important to conduct a systematic review and meta-analysis of available studies examining the effects of GH therapy in adults with age-related osteopenia or osteoporosis, with the aim to clarify the possible anabolic role of GH in the treatment of osteoporosis and its sequelae.

This systematic review addresses prospective controlled studies evaluating the bone effects of GH therapy as compared to no GH therapy (placebo or conventional osteoporosis therapies) in adults with osteopenia or osteoporosis.

1.5 Thesis Objectives

(1) Examine the effect of growth hormone therapy on bone densitometric endpoints, bone turnover markers and fracture risk in adults with osteopenia or osteoporosis and no organic growth hormone deficiency.

(2) Evaluate the safety of growth hormone therapy in the above-mentioned population of interest

1.6 Thesis Hypothesis

Growth hormone therapy maintains or improves bone mineral content and bone mineral density and reduces fracture risk in non-deficient adults with osteopenia or osteoporosis, without significant major adverse events. It is thus a potential anabolic therapy for the treatment of this condition.

CHAPTER 2

DATA AND METHODS

2.1 Protocol

A protocol was developed for this systematic review describing pre-specified objectives, study population, outcomes of interest, search strategy and planned analysis but was not published. Few modifications were made after the writing of the initial protocol aiming at expanding the outcomes of interest, as will be detailed at a later stage. Both the protocol and the systematic review were written in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement. PRISMA is an evidence-based checklist of items that was developed to improve transparent reporting of systematic reviews and meta-analyses. It includes 27 items addressing the preferred content of the title, abstract, introduction, methods, results, discussion and funding of protocols and actual systematic reviews and meta-analyses (77) (Appendix 1).

2.2 Data sources and search strategy

We conducted a systematic search of the literature for controlled prospective and randomized studies on the effect of GH therapy in patients with osteoporosis or osteopenia. Our searches were both computerized and manual. The online search included the databases MEDLINE (1946 to present), EMBASE (1947 to present) and the Cochrane Register of Controlled Trials. No language or year of publication limitation was used. The search was initially conducted in March 2014, and was updated in November 2015. The search was performed using the keywords and MESH

terms relevant to the growth hormone intervention (growth hormone, human growth hormone, somatotropin, somatropin, somatotrophin and somatrophin) and to the study population of adults with age-related osteoporosis or osteopenia (osteoporosis, postmenopausal osteoporosis, osteopenia, bone density, bone loss, bone demineralization, bone fractures, broken bone and low bone mineral density or content) and the Boolean functions AND, OR. Relevant abbreviations for growth hormone (GH, hGH, h-GH, rhGH, r-hGH and rh-GH) and available commercial names for this therapy (Genotropin, Humatrope, Norditropin, Nutropin, Omnitrope, Saizen, Serostim, Tev-tropin and Zorbtive) were also included in the search strategy.

We also performed a manual search of the references of both original articles collected and pertinent review articles on the topic.

The detailed search strategy is available as Appendix 2.

2.3 Eligibility criteria

2.3.1 Type of studies

Inclusion criteria:

- Prospective controlled studies, including randomized and non-randomized controlled trials, placebo-controlled and active comparator (osteoporosis therapy)-controlled trials
- No restriction on date of publication, study language, publication status and study sample size

Exclusion criteria:

- Uncontrolled studies

2.3.2 Type of population

Inclusion criteria:

- Postmenopausal women and men above age 50 years
- Presence of osteopenia or osteoporosis (defined as patients with a T-score on BMD equal to or less than -1 at any skeletal site, and/or as patients with history of fragility fracture)

Exclusion criteria:

- Patients with organic growth hormone deficiency (where the effect of GH therapy on bone has been studied)
- Patients younger than 50 years (where the approach to osteoporosis is different and the rationale for treatment with GH does not apply as a relative age-related decrease in IGF-1 is considered a reason for the potential benefit from GH therapy)

  Studies including participants from all age groups were excluded when the mean patients' age was less than 50 years
- Patients with secondary causes of osteoporosis like end-stage renal disease, glucocorticoid excess (who have a different disease etiology as well as specific indications and modes of therapy)

2.3.3 Type of intervention

Inclusion criteria:

- Treatment with human growth hormone or recombinant growth hormone
- Daily or cyclic therapy

Exclusion criteria:

- Studies where the length of treatment is less than 6 months (as there is no expected benefit from therapy with a short treatment interval, based on individual trials and meta-analysis on the effect of growth hormone replacement on BMD in patients with GHD (55)

2.3.4 Type of outcome

Inclusion criteria:

- Studies reporting bone mineral content, bone mineral density, fractures or bone turnover markers
- No restriction on the method used to measure BMC, BMD, bone markers and fractures
- Studies reporting absolute values for the outcomes of interest after GH therapy

  Studies reporting percent change in the outcome were included, when the absolute value of the outcome could be calculated from the percent change and the baseline value

  When the outcome of interest was not available in text or was only presented graphically, the corresponding author was contacted, by email (initial email and 2 reminders). When the missing data could not be obtained (no reply or unavailability of records), the absolute values were retrieved from the graph, when available. Otherwise, the study was excluded

2.4 Outcome measures

The selection of outcome measures was done after searching the Core Outcome Measures in Effectiveness Trials (COMET) initiative, which includes core outcome measures that are suggested to be measured in clinical trials addressing a specific condition (78). In the case of osteoporosis trials, endpoints include true clinical health status outcomes, namely bone fractures and health status instruments; and intermediate outcomes, namely bone densitometry and biochemical markers (79).

Our initial choice of outcomes included fractures and bone densitometric endpoints (measured by DXA), outcomes that are commonly used in osteoporosis trials. However, due to the small number of studies retrieved and to their heterogeneity, we decided to repeat the initial screen and include studies wherein bone densitometry was measured by older techniques (Single Photon Absorptiometry (SPA) and Double Photon Absorptiometry (DPA)) and studies including bone markers. The validity of using SPA in the measurement of bone density is illustrated in a study showing a tight correlation between SPA and DXA at the forearm site ($r = 0.99$), suggesting that the use of forearm SPA has a diagnostic value close to that of DXA (80,81). Based on this information, we used SPA forearm as an outcome measure. Significant correlations have also been identified between BMC and BMD obtained by DPA and DXA, though DXA had a better precision (82). Bone densitometric results obtained using DPA were thus also used as outcome measures. The advantage of using bone markers as an outcome is that they allow a frequent determination of bone metabolism and they are sensitive for acute changes in bone turnover (83). There are also studies indicating their predictive value in assessing fracture risk and their correlation with changes in bone density (79,84,85). However, they have the disadvantage of potential for variability in

results, as the collection method and time, the type of analytical method used and different patients' characteristics affect the results obtained (83).

The outcomes used can be summarized as follows:

Primary outcome:

- The mean difference in BMD using DXA after treatment with GH v/s comparator

Secondary outcomes:

- The mean difference in BMD using DPA after treatment with GH v/s comparator
- The mean difference in BMC using DXA after treatment with GH v/s comparator
- The mean difference in BMC at the distal forearm using SPA after treatment with GH v/s comparator

Fractures:

- The risk ratio in incidence of fractures after treatment with GH v/s comparator

Bone Markers:

- The mean difference in osteocalcin (OC) after treatment with GH v/s comparator (bone formation marker)
- The mean difference in procollagen type I carboxy-terminal propeptide (PICP) after treatment with GH v/s comparator (bone formation marker)
- The mean difference in C-terminal propeptide of type I collagen (CICP) after treatment with GH v/s comparator (bone formation marker)

- The mean difference in serum C-telopeptide cross-link of type I collagen (CTX) after treatment with GH v/s comparator (bone resorption marker)
- The mean difference in total urinary pyridinolines after treatment with GH v/s comparator (bone resorption marker)

  Adverse events:
- The risk difference in incidence of adverse events and all-cause mortality between groups

2.5 Study selection

Two reviewers (MB and SG) independently screened the title and abstract of all retrieved records with our search strategy. Screening for eligibility was performed based on the inclusion and exclusion criteria detailed previously, with regards to study type, study population and intervention. Initial screening did not include the outcome measured with the aim to increase the sensitivity of the selection process. After the initial screening, we retrieved the full text of articles retained by at least one reviewer. Full texts were then screened independently and in duplicates by two reviewers (MB and NN), following a pre-specified full text screening form (Appendix 3). After screening, reviewers met and reviewed the retained articles. In cases of disagreement (texts included by one reviewer and excluded by another), one reviewer (MB) sought the advice of the thesis committee and the expert in the field (NT). Study selection was summarized using the PRISMA flow diagram.

2.6 Data collection process

Two reviewers (MB and NN) independently and in duplicate extracted data from included articles using a pre-specified data collection form (Appendix 4). The form was developed initially and reviewed by the thesis committee and the expert advisor. The form was then pilot tested on one randomly selected included article for further refinement. Disagreement between reviewers was solved by discussion and, when not resolved, by contacting the thesis advisor and/or the topic expert.

We contacted by email (initial email and 2 reminders) the corresponding authors of two studies that did not include absolute values for the outcomes of interest or who included outcomes only in graphs with no numerical values in their published manuscript (86, 74). We did not get any reply. Data were then extracted from the published graph. We also contacted by email the corresponding authors of three studies in the case of inconsistent data (small standard deviation, possible wrong units of measurement, large value for the outcome as compared to other studies) (87,74,73). Two of the contacted authors answered and corrected for us the results (standard error used instead of SD for BMD in the study by Holloway in 1997 and wrong units published in the study by Landin-Wilhelmsen in 2003 (kilograms instead of grams as unit for BMC and nanogram/liter instead of micrograms/liter for PICP)). Values were then corrected in the extracted data. We did not receive an answer regarding the high values of osteocalcin in the study by Saaf in 1999. The study was then excluded with regards to that specific outcome.

To avoid double counting of publications reporting data on same patient population, we reviewed the authors and centers of all included studies, along with the selection criteria for each study population. Two included studies reported data on the

same patient population, at different time points (73,88). The more recent publication (88) included fracture data while the earlier one (73) reported bone densitometry results, as well as bone markers. Each was then included with regards to its published outcome. Both studies were not included together for the same outcome. The presence of duplicate publications is thus highly unlikely.

Included articles were all published in English and thus translation was not needed.

2.7 Data items

Information was extracted from each included trial, when available, on:

- Study design
- Characteristics of study participants including age, gender, BMI, years since menopause, number of subjects and number of subjects withdrawn
- Characteristics of study participants with regards to their osteopenia or osteoporosis including years since diagnosis, prior therapies, concurrent treatment along with growth hormone such as calcium, vitamin D and others, history of fractures (including number and site) and baseline T-score
- Type of intervention including dose/frequency of growth hormone and duration of treatment, IGF-1 levels and change in IGF-1
- Type of outcome and method of measurement (BMC/BMD using SPA, DPA or DXA, bone markers with the measuring assay, fractures and adverse events)
- Study funding source

2.8 Risk of bias in individual studies

We assessed the risk of bias of the included randomized studies using the Cochrane Collaboration's tool for assessing risk of bias in randomized trials (89). The tool is a checklist of six types of bias and seven domains. It addresses selection bias through evaluation of the random sequence generation (randomization method) and the allocation concealment, performance bias by looking at the blinding of study participants and personnel, detection bias through the blinding of the outcome assessment, attrition bias or incomplete outcome data, reporting bias (selective outcome reporting) and other sources of bias like lack of intention-to-treat analysis. The risk of bias for each item is presented as "low risk of bias", "high risk of bias" or "unclear risk of bias" (89). In this systematic review, the risk of bias was assessed independently and in parallel by two reviewers (MB and NN). Results were compared and discussed after. In the event of disagreement, discussion was pursued with the study advisor. We assessed the risk of bias for the primary outcome i.e. the mean difference in BMD by DXA using GH or a comparator at different skeletal sites, by reviewing the reported items in included studies. In our review, other sources of bias included lack of intention-to-treat analysis. To evaluate for selective outcome reporting, we searched for the availability of a published protocol for the study and compared the specified outcomes of interest in the methods section of each article to those actually retrieved and reported in the results. The results of the bias assessment were presented both in text and graphically.

The risk of bias in the controlled observational study by Krantz et al was assessed using the risk of bias tool proposed by the Agency for Healthcare Research and Quality (90).

The risk of publication bias (resulting from the non publication of studies with negative results) was planned to be assessed by doing a funnel plot of the included studies addressing the primary outcome. Due to the small number of trials (n = 3) addressing the primary outcome, the power of the test was considered low to distinguish chance from true asymmetry (89) and thus publication bias could not be assessed in this manner.

Studies were not excluded from the review or analysis based on the risk of bias.

2.9 Synthesis of results & methods of analysis

2.9.1 Data handling

Data were extracted from individual studies with regards to the selected outcome measures. For continuous outcomes, the mean and standard deviation after treatment were directly obtained from the published manuscript, if available. Otherwise, they were derived using appropriate statistical formulas (89):

- When the standard error was reported, the standard deviation was derived as:

  Standard deviation = Standard error * √n

  (where n is the sample size of the group)
- When the 95% confidence interval was reported, the standard deviation was derived as:

  Standard deviation = √n * (upper limit – lower limit) / 3.92

  (where n is the sample size of the group)
- When the range was available, the SD was estimated as one quarter of the range of values

- When the SD at the end of treatment was not available (91,92), it was assumed to be similar to the baseline SD, since it can be argued that the intervention itself does not alter the variability in the outcome measure

BMC and BMD data were all reported using the same scale (g/cm^2). Laboratory markers were sometimes reported using different units of measurement. In that event, they were appropriately converted, according to international conversion measures, to allow their combination in meta-analysis.

For studies presenting outcomes at different time intervals, data were collected at the different time points but were combined in meta-analysis at the most distal controlled time point available.

For trials that randomized participants to one of several intervention groups (studies with more than two arms), data handling was done as follows:

- In the study by Holloway et al (87), participants were randomly assigned to four treatment groups: GH plus Calcitonin (CT), GH plus placebo, placebo plus CT and placebo plus placebo. Results were reported in the published manuscript for the four treatment groups. Each pair-wise comparison was included separately in our meta-analysis (GH plus placebo v/s placebo plus placebo as one pair and GH plus CT v/s placebo plus CT as another pair). This allowed including all treatment groups while respecting the inclusion criteria (GH as an intervention in one arm v/s placebo or osteoporosis therapy in the other arm) and without double counting.
- In the study by Landin-Wilhelmsen et al (73), participants were randomized to one of two growth hormone intervention groups using different treatment doses or to a placebo control group. Due to the presence of a common control group,

the study outcomes (BMC, BMD, bone markers) for the two GH groups were pooled together

Calculation of the pooled mean and the pooled standard deviation was done using the following formula:

Pooled mean = $(n_1 m_1 + n_2 m_2) / (n_1 + n_2)$ (93)

Pooled standard deviation = $[(n_1-1) (Sd_1)^2 + (n_2-1) (Sd_2)^2] / (n_1 + n_2 - 2)$ (formula used for independent, random samples) (94)

Where n_1 is the sample size of treatment group 1, n_2 is the sample size of treatment group 2, m_1 is the mean of treatment group 1, m_2 is the mean of treatment group 2, sd_1 is the standard deviation of treatment group 1 and sd_2 is the standard deviation of treatment group 2

2.9.2 Data synthesis

Data retrieved for selected outcomes were summarized in text and tables. When at least 2 articles were available for an outcome, meta-analysis was used for data synthesis. Eligible results for meta-analysis were all in the form of continuous data, except for fracture risk. The effect summary size was then the weighted mean difference (WMD), presented with a 95% CI.

For the primary outcome, we calculated the WMD and 95% CI in BMD obtained by DXA after intervention with GH or comparator at each skeletal site (lumbar spine, total femur and femoral neck).

For the secondary outcomes, we retrieved the WMD and 95% CI in BMC by DXA after intervention at the lumbar spine and the femoral neck and in BMC by SPA at the distal forearm. We also calculated the WMD and 95% CI in each bone marker after

intervention (osteocalcin, PICP, CTX, CICP, total urinary pyridinolines). As for the fracture incidence, we obtained the risk ratio and 95% CI between compared groups.

Outcomes were combined in meta-analysis using a random-effect model. This model is a conservative estimate of the overall effect size, whereby it assumes that combined studies have different effect sizes that follow a certain distribution. The pooled effect size in that case is the center of this distribution of study outcomes. Choice of this model over a fixed-effect model was made based on the included studies that presumably show a certain level of heterogeneity and within study variability. The fixed-effect model assumes a similar effect size in included studies and low variability in between, which is theoretically due to chance, and was thus considered less applicable in our case (89).

Using the random-effect model, a weighted mean difference was obtained for each of the primary and secondary continuous outcomes, representing the summary effect. The weight is dependent on both the within study variability (equal to the inverse variance) and the between study variability (represented by Tau^2). Our null hypothesis is that the mean difference is zero (no difference between the GH intervention and the comparator). A 95% CI, obtained with each WMD, was used to establish the statistical significance of the obtained results (89). For the dichotomous outcome of fracture incidence, a risk ratio was derived as the summary effect with a ratio of 1 representing no difference in risk of fracture between groups.

Analysis was conducted using RevMan (version 5.3).

2.9.3 Assessment of heterogeneity

Evaluation for heterogeneity or inconsistency among included studies was done using the chi-squared test (Chi^2). In this test, a large test statistic or a low p value is a sign of heterogeneity i.e. it indicates that the difference between effect sizes in the different studies is not due to chance alone. Because of the presence of a small number of studies and a small sample size that potentially decrease the power of the test to detect heterogeneity, a p value < 0.1 was considered significant (89).

Heterogeneity was quantified using I^2. I^2 is derived from the test statistic *Q* and the degree of freedom *df* of the chi-squared test and gives the percentage of variability in the effect size that is due to inconsistency rather than chance.

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

Values of 30%, 50% and 75% were considered as indicative of moderate, substantial and considerable amount of heterogeneity, respectively (89).

2.10 Additional analysis

A pre-planned subgroup analysis was planned to evaluate the effect of gender on the observed primary outcome. However, we could not perform such analysis as retained studies only included women.

Similarly, the effect of age, severity of osteoporosis, treatment duration and growth hormone dose on the main outcomes could not be analyzed using meta-regression, as pre-planned, due to the small number of studies included. Meta-regression is usually not advised when included studies are less than ten (89).

2.11 Assessment of the quality of evidence

The quality of evidence was assessed for the primary outcome, the mean difference in BMD measured by DXA at different skeletal sites. Evaluation was carried using the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) methodology. This tool permits to assess the confidence in the obtained effect estimate through three steps. The initial step involves classification into an initial level of confidence based on the study design: randomized trials will lead to a high confidence while observational studies result in low confidence. The next step will lead to lowering or raising confidence based on the risk of bias (already conducted), the inconsistency (i.e. heterogeneity), indirectness and imprecision (i.e. wide confidence interval) of results. It also takes into account publication bias, effect size, presence of dose response and possible confounders. Based on this assessment, a final level of confidence is derived (high, moderate, low or very low) (95).

CHAPTER 3

RESULTS

3.1 Study selection

A total of 7664 citations were identified through the online search of the databases Medline, Embase and the Cochrane Register of Controlled Trials. After duplicate removal, we retained 6025 citations. We did a title and abstract screen for these citations, which left us with 58 citations. Reasons for exclusion at this initial screen included abstracts of articles clearly alluding to the wrong study population, including children and adolescents, patients with growth hormone deficiency, patients with renal failure or secondary osteoporosis and healthy adults. It also included titles of animal studies, abstracts indicating a different intervention such as growth hormone releasing hormone and IGF-1 and review articles. We retrieved the full text articles for 57 citations out of 58 for further screening. 1 article was excluded due to unavailability of a full text, despite exhaustive search by a specialized librarian. 49 articles were then excluded and we were left with 8 articles for inclusion, 1 of which being the follow-up of an earlier study with new outcomes. Reasons for exclusion after detailed review included: review and commentary articles (n=6), animal studies (n=4), trials whereby treatment is with IGF-1 or GHRH or where is no treatment (n=5), studies in premenopausal women or men less than 50 (n=4), trial in a population with short bowel syndrome (n=1), studies in GHD (n=2), studies in healthy elderly (n=3) or healthy adults (n=1), trials with a treatment period less than 6 months ranging between 7 days and 3 months (n=5), short-term treatment for fracture healing (n=8), trials that lacked any of our outcomes of interest (n=3) and uncontrolled trials (n=7). Among the

remaining 8 studies, 3 studies included data on BMD measured by DXA, 2 studies had BMC by DXA, 3 studies reported BMD or BMC by SPA, 1 study utilized DPA for bone densitometric endpoints, 5 articles published values of bone markers and 4 studies reported fractures (one having a 10 year follow up report) (figure 1).

3.2 Study characteristics

Five out of 8 included studies were randomized, placebo-controlled, with 4 being double blind. Among the remaining 3 studies, 2 were randomized controlled and 1 was controlled. Studies reporting BMD by DXA were all randomized, double blind, placebo-controlled.

Baseline demographic and clinical characteristics of study participants are detailed in table 1. A total of 272 subjects were included in 7 studies, 22 of whom withdrew during follow-up. There were 210 subjects in studies reporting BMD, 180 of whom had BMD measured by DXA and 30 by DPA. There were 165 subjects in studies reporting BMC, including 103 checked by DXA and 62 by SPA. 219 participants were included in trials reporting data on bone markers and 135 were part of studies actively seeking fractures.

The mean age of study participants ranged from 60.7 to 69.2 years and their mean body mass index (BMI) from 24.4 to 25.3 kg/m^2. All included studies involved women. No study reporting on men was eligible for inclusion. 3 studies reported on duration since menopause. One of them (74) included women who were at least 5 years after menopause. The remaining two studies involved women who were at a mean of 11.8 and 17.3 years from menopause, respectively (96,86). In 4 out of 7 studies, baseline IGF-1 levels were reported (86,87,74,73). In 3 trials, they were described as

being normal for age (86,74,73) and in one trial as being low for age (87). In all studies, a significant increase from baseline was observed with GH therapy, in parallel to the GH intake with a quick return to baseline levels with treatment cessation. The increase in IGF-1 was described as being within age standards in 3 studies (87, 74, 73).

One trial was conducted in women with newly diagnosed osteopenia, defined as lumbar spine BMD less than 1 SD below the mean value for young women (WHO disease definition) (87). The remaining studies included participants with osteoporosis. The condition was defined based on the WHO criteria of BMD equal to or lower than 2.5 SD of young adults in 1 study (73). In 5 studies, osteoporosis was defined by the presence of a vertebral fracture (clinically defined in one trial (86) and radiologically detected in the other 4 (74,91,92,96). When reported, the lumbar spine SD from the mean young reference population (T-score) ranged between - 2.7 and – 2.8. The presence of at least one vertebral fracture was an inclusion criteria in 5 out of 7 studies (86,74,91,92,96). In one study looking at fracture endpoints, 56% of women initially had an osteoporotic fracture (73). One trial did not have osteoporotic fractures at baseline (including patients with osteopenia) (87). In 2 studies where the duration of osteoporosis was reported, 1 trial was conducted in patients with newly diagnosed osteoporosis (96) and one study involved patients with a mean of 1.9 years since disease establishment (73). Four included studies reported on prior therapies before study initiation. In one study, patients were treatment naïve (96). In one trial, patients were maintained on calcium and a multivitamin before study initiation (74) while in two studies patients were on hormone replacement therapy (41 % in the trial by Holloway and 100% in the Landin-Wilhelmsen study).

Along with the growth hormone or placebo intervention, 6 studies out of 7 maintained participants on calcium supplementation, with daily doses ranging between 500 and 1200 mg (86,87,73,74,91,92). One study reported the use of vitamin D at a dose of 400 IU daily (73). Concurrent therapies for osteoporosis were utilized in study populations (in both the GH and the control group) in all included trials, except for one study (74). In the study by Erdstieck et al., patients were maintained on the bisphosphonate pamidronate at the treatment dose of 150 mg per day (86). Two studies had patients on hormone replacement therapy (41% of study subjects in the trial by Holloway et al. and all women in the Landin-Wilhemsen study) (87,73). In the 10-year follow up report on the Landin-Wilhelmsen initial study, 41% of participants were still on hormone replacement, while 23% started bisphosphonates and 3% teriparatide (88). Salmon calcitonin was used in 4 studies. It was used either in all study subjects (91,92) or in subgroups (87,96); in repeated cycles of 5 days (87), 21 days (96) or 3 months (92), or during the whole study period for three days weekly (91). There were a wide range of GH treatment regimens. The two most recent studies used regular daily dosing with GH over the whole study period (0.3 - 0.83 mg/d) (74,73). Two studies had patients on GH three times weekly over the study period (86,91). The remaining 3 studies used cycles of 7 days or 2 months of GH (87,92,96). Two studies used weight-based dosing regimens (86,87), while the rest used fixed dosing. Three studies had cycles with particularly high GH doses (2 – 4 mg/d) (91,92,96). With the exception of the 2 oldest studies that prescribed human growth hormone (91,92), all included trials used recombinant human GH. Treatment duration for controlled studies ranged between 6 months and 24 months. One extension study examined fracture data after 10 years (88).

3.3 Risk of bias assessment

Assessment of risk of bias is summarized in figure 2 and detailed in table 2. For the primary outcome BMD, measured at both the lumbar spine and the femoral neck, all studies were considered at an unclear risk for reporting bias (no protocol registration available). Selection bias from randomization was unclear in two trials (87,74) and low in one (73). In contrast, there was a low risk for selection bias from allocation concealment in the studies by Holloway and Landin-Wilhelmsen (87,73) and an unclear risk in the Saaf study (74). A high risk of performance bias was found due to unclear blinding of study personnel, except in the Landin-Wilhelmsen trial (low risk) (73). Detection bias was low in the three trials. A low risk of attrition bias was noted in the studies except for the one by Holloway (high risk) (87). Both studies by Holloway and Saaf (87,74) had different types of risks for different types of biases, however the study by Landin-Wilhelmsen was at low risk for all biases except for selective outcome reporting (unclear risk) (73).

3.4 Results of individual studies and synthesis of results

3.4.1 Effect of GH therapy v/s comparator on BMD measured by DXA

3.4.1.1 Lumbar spine

Three randomized placebo-controlled studies examined the effect of GH therapy on BMD at the lumbar spine (87,74,73).

In the Holloway trial, using intention-to-treat analysis, GH cyclical treatment for 24 months in women with osteopenia resulted in a significant increase in LS BMD from baseline in both the GH/Placebo group ($1.72 \pm 0.74\%$, $p<0.05$) and the GH/CT group ($2.70 \pm 0.81\%$, $p<0.01$) and a non-significant change in the placebo groups. Observed

changes were not significant at 12 months. At the end of therapy, there was a significant difference in the change in BMD from baseline between the GH/Placebo group and the Placebo/Placebo group. Results did not change when subtracting the non-adherent participants (22%) or when accounting for estrogen status. Significant changes form baseline observed at the end of randomization no longer achieved significance after one year of stopping GH (87).

In the Landin-Wilhelmsen trial, no significant change from baseline and no significant difference was observed between groups treated with placebo, low dose GH (0.33mg/d) or high dose GH (0.83 mg/d) on a daily basis, along with hormone replacement therapy, started on average 4 years before study initiation with a range between 1 to 25 years. An increase from baseline in BMD became significant in all study groups at 3 years of therapy (controlled non-randomized phase), even in the placebo group; however with no significant difference between GH and placebo groups. Change from baseline was maintained significant after 1 and 2 years of study termination and GH cessation (73).

In the Saaf study, continuous daily therapy with GH (mean dose 0.5 mg/d) over one year did not result in any significant change in LS BMD and at the end of the trial BMD results were comparable to those achieved with placebo. Maintenance of GH therapy for an additional year resulted in a significant increase of $6.0 \pm 2.2\%$ from baseline (74).

The results of the three studies were combined in meta-analysis, with the Holloway study including 2 pair-wise comparisons (GH plus placebo v/s placebo plus placebo as one pair and GH plus CT v/s placebo plus CT as another pair), each included separately in result synthesis. A total of 93 women were prescribed GH for a duration

ranging between 12 months and 24 months and compared to 71 women on placebo. A non-significant difference was observed between the two treatment groups (WMD = -0.01, 95% CI [-0.04, 0.02], P = 0.55). There was no heterogeneity among included studies (Chi^2= 2.25, df = 3, P = 0.52, I^2 = 0) (figure 6).

3.4.1.2 Total hip

The effect of GH therapy on BMD at the total hip was only studied in one included trial (87). In this trial, GH was given in two subgroups, one alone and one concurrent with calcitonin, each treatment subgroup compared to a placebo group with or without calcitonin. No significant difference was observed in total hip BMD between the groups treated with GH alone or placebo, with a mean difference of 0.03 favoring placebo and a 95% confidence interval [-0.10, 0.05] at the end of the study. Similarly, a non-significant mean difference favoring GH of 0.03, 95% CI [-0.04, 0.10] was observed between subgroups treated with GH and calcitonin as compared to placebo and calcitonin. When examining change from baseline, both the GH/Placebo group and the Placebo/CT group showed a significant increase of 1.26 ± 0.60% and 1.27 ± 0.53%, respectively.

Combining the two subgroups in meta-analysis resulted similarly in a non-significant difference in BMD after treatment with GH or placebo (WMD = 0, 95% CI [-0.05, 0.06], P = 0.92). There was no evidence of heterogeneity between the two subgroups (Chi^2= 1.14, df = 1, P = 0.29, I^2 = 12) (figure 7).

3.4.1.3 Femoral neck

Three randomized placebo-controlled studies examined the effect of GH therapy on BMD at the femoral neck (87,74,73).

In the Holloway trial where the femoral neck BMD was a secondary outcome, no significant difference was observed at the end of the study between the two pair-wise comparison groups (87).

In the Landin-Wilhelmsen trial, no significant change from baseline and no significant difference were observed between GH treatment groups and placebo, given along with sex steroid replacement therapy for 18 months. The high dose GH group reached statistically significant increase from baseline at 1 and 2 years from stopping GH, however these changes were not significant when compared to placebo (73).

In the Saaf study, daily therapy with GH (mean dose 0.5 mg/d) over one year led to a significant decrease in FN BMD (3.4 ± 1.6%), as compared to baseline. This drop was regained when therapy was continued in an uncontrolled manner for an additional year (74).

The results of the three studies were again combined in meta-analysis, with the Holloway study included as 2 pair-wise comparisons, resulting in 4 included comparisons. One comparison included treatment with calcitonin along with GH and placebo (87), and another included sex steroid replacement therapy (73).

A total of 87 women were prescribed GH for a duration ranging between 12 months and 24 months and compared to 64 women on placebo. A non-significant difference was observed between the two treatment groups (WMD = 0, 95% CI [-0.03, 0.04], P = 0.77). There was no evidence of heterogeneity among included studies (Chi^2= 1.09, df = 3, P = 0.78, I^2 = 0) (figure 8).

3.4.2 Effect of GH therapy v/s comparator on BMD measured by DPA

Dual-photon absorptiometry was only used in the measurement of BMD in one included study (96).

In a 24 months trial, 30 patients with osteoporosis and vertebral fractures were randomized into one of 3 treatment groups of cyclic GH (7 days), calcitonin (21 days) and rest (61 days) or cyclic placebo/calcitonin/rest or cyclic GH/placebo/rest. At the end of the trial, no significant difference in BMD was found between the 3 groups at the three studied sites (lumbar spine, femoral shaft and distal radius). When evaluating the change from baseline in individual groups, an increase in lumbar spine BMD was observed in the combined GH/calcitonin group as compared to a significant decrease in the 2 other groups, a significant decrease in femoral shaft BMD was noted in the GH groups as compared to no change in the calcitonin only group, and no change was observed in all groups at the distal radius (96).

3.4.3. Effect of GH therapy v/s comparator on BMC measured by DXA

3.4.3.1 Lumbar spine

Two randomized placebo-controlled trials examined the effect of GH on BMC measured by DXA (86,73).

In the study by Erdtsieck, 23 patients were randomized to GH versus placebo, along with the bisphosphonate pamidronate given to both groups. At the end of the treatment period of 6 months, a significant increase from baseline in BMC was observed in the placebo/pamidronate group that was statistically different from the GH/pamidronate group (no increase) (86).

In the study by Landin-Wilhelmsen, a significant increase from baseline in BMC was observed in placebo and GH treatment groups after 18 months of controlled therapy, with concomitant hormone replacement. The increase from baseline remained significant after 3 years of uncontrolled GH therapy, and even at 1 and 2 years from stopping GH. Only at 1 year from GH cessation, the BMC in the high dose GH group (14% increased from baseline) became significantly different from the placebo group (73).

Results in BMC after controlled therapy with GH versus placebo were combined in meta-analysis. A total of 65 women were prescribed GH for a duration of 6 months to 18 months and compared to 36 women on placebo. A non-significant difference was observed between the two treatment groups (WMD = - 0.71, 95% CI [-1.63, 0.22], P = 0.13). There was no evidence of heterogeneity among included studies (Chi^2= 0.63, df = 1, P = 0.43, I^2 = 0) (figure 9).

3.4.3.2 Femoral neck

The same two studies reported on BMC at the femoral neck.

No significant change was observed with GH therapy versus placebo in femoral neck BMC in the Erdtsieck study (86).

In the Landin-Wilhelmsen trial, at 18 months of randomization, no significant difference in BMC was observed between study groups. Only the high dose GH group had a significant increase form baseline. A similar increase from baseline was observed at 3 years of therapy (partly uncontrolled) in both GH groups but not with placebo. This change from baseline remained significant after 1 and 2 years of stopping GH. At 1 year

of GH withdrawal, the BMC was statistically different between the two doses of GH favoring the high dose of 0.83 (73).

Results in femoral neck BMC after controlled therapy with GH versus placebo were combined in meta-analysis. A non-significant difference was observed between the two treatment groups (WMD = - 0.06, 95% CI [-0.28, 0.16], P = 0.58). There was substantial heterogeneity among included studies (Chi^2= 2.19, df = 1, P = 0.14, I^2 = 54%) (figure 10).

3.4.4 Effect of GH therapy v/s comparator on BMC measured by SPA at the distal forearm

The effect of GH therapy on BMC at the distal forearm, measured by SPA, was addressed by one randomized controlled trial (86) and 2 controlled studies (91,92).

In the study by Erdtsieck et al, a significant increase from baseline in BMC was observed in the placebo/pamidronate group (n = 11) but not in the GH/pamidronate group (n = 10), and no significant difference was found between groups after 6 months of randomization (86).

In the Aloia 1985 trial, 24 months treatment with alternating days of GH and calcitonin administered for women with osteoporosis and vertebral fractures (n = 13) resulted in a significant drop in BMC of 2.91% per year, as compared to a minimal change of 0.05% per year in the calcitonin only group (n = 12). BMC values at the distal forearm at the end of the treatment were significantly different between groups (91).

In the second trial by the same group (92), 24 months therapy with cycles of daily GH for 2 months followed by daily calcitonin for 3 months and then 3 months rest also given for women with vertebral fractures (n = 7) did not result in a significant

change in forearm BMC, as compared to cycles of rest and calcitonin without GH (n = 7), with 0.05% and -0.05% change per year, respectively (92).

Results from the two Aloia trials were combined in meta-analysis, resulting in 20 patients treated with GH compared to 19 patients receiving calcitonin. No significant difference was observed in BMC between the 2 groups (WMD = - 0.06, 95% CI [-0.23, 0.10], P = 0.44). There was considerable heterogeneity among included studies (Chi^2= 4.41, df = 1, P = 0.04, I^2 = 77%) (figure 11). The study by Erdtsieck et al could not be incorporated in meta-analysis, due to the BMC unit used that could not be converted (unit/cm).

The effect of GH therapy on bone densitometric endpoints is summarized in table 3.

3.4.5 Effect of GH therapy v/s comparator on bone biomarkers

The effect of GH therapy on bone turnover biomarkers was studied in 5 included publications, where markers were secondary outcomes (86,96,87,74,73). Different bone formation markers and bone resorption markers were reported in different studies (table 4). The effect of therapy on 4 markers could be combined in meta-analysis (table 5).

The influence of GH or a comparator on osteocalcin, a bone formation marker, was evaluated in the 5 studies, 4 of which could be combined for data synthesis. One study had outlier values (74). After contacting the concerned author and in the absence of a reply, the study was not included in meta-analysis. A total of 107 people received GH and 86 received a comparator. A non-significant difference was found between groups at the end of the study periods (figure 12). The effect of GH therapy on another bone formation marker, PICP, was the subject of 4 randomized, double blind placebo-

controlled trials, with 103 women on GH and 82 on a comparator therapy (86,87,74,73). We found a significant difference in PICP at the end of the study period, favoring GH as compared to placebo (WMD = 14.03, 95% CI 2.68, 25.38], P = 0.02) (figure 13).

As for bone resorption markers, the influence of therapy on total urinary pyridinolines was the subject of one trial with 2 pair-wise comparisons with a total of 32 subjects in GH and 40 on placebo (with or without calcitonin) (87). A trend to favor GH was observed but the difference in the results was non-significant (P = 0.06) (figure 14). Similarly, the effect on another resorption marker, CTX, was studied in the Holloway trial (87). Again a non-significant trend to favor GH was obtained (P = 0.09) (figure 15).

3.4.6 Effect of GH therapy v/s comparator on fracture risk

Fracture incidence during the study period or during later follow-up was reported in 4 studies and one follow-up trial (91,92,74,73,88). Three of the studies included women with prior history of vertebral fractures (91,92,74), while in the Landin-Wilhelmsen trial, 56% of study patients had a vertebral fracture (73). Fractures were detected in included studies either by symptoms that were confirmed radiographically, or by screening for silent vertebral fractures through spine X-Rays. A vertebral fracture was defined as a vertebral body height loss of > 15% anterior, middle or posterior in the study by Saaf et al (74), a drop > 20% in vertebral height in the study by Landin-Wilhelmsen et al and its 10-year follow-up report (73,88) and as a reduction in anterior vertebral body height of > 25% in the two reported trials by Aloia (91,92). In the study by Saaf et al, 3 symptomatic fractures (1 patellar and 2 vertebral) were detected in the GH group during the two-years treatment period (one year randomized

and one controlled) and 7 new or worsened vertebral fractures were found on spinal imaging (74). In the early trial by Aloia et al, evaluation of spine X-rays at the study end revealed further compression or new vertebral fracture in 5 patients in the combined GH/CT group and 6 in the placebo/CT group. The number of new vertebral fractures was equal in both groups (91). In the second study by Aloia, one further compression of a fracture was detected after 2 years of cyclic therapy with GH/CT, while patients on cyclic GH/Placebo sustained one new vertebral fracture and had further compression in two old fractures (92). In the study by Landin-Wilhelmsen et al, no fractures were observed during 3 years of GH therapy (73). In the recently published follow-up report, 7 years after stopping GH, the number of fractures in 80 women (initially recruited) dropped from 56% at baseline to 28% (P = 0.0003), with no significant difference between the 3 study groups (placebo, high dose and low dose GH) (88). The most common fracture site was at the radius, followed by the femoral neck, upper arm and then vertebrae, ribs and ankle. During the 10 years follow-up, 41% stopped hormone replacement therapy, 23% started bisphosphonates and 3% teriparatide (88). In parallel, in a random population sample of postmenopausal women (control group), a four-fold increase in the prevalence of fractures was observed, from 8% at baseline to 32% after 10 years (P = 0.0008). The most common fracture site was at the radius, followed by the ankle, the ribs, the upper arm, the vertebrae and finally the femoral neck. In this group, the use of hormone replacement therapy had declined throughout the follow-up period from 40% to 8%, while the use of bone-specific agents increased from 0% to 4% (88). There was no significant difference in the fracture incidence between the group initially treated by GH and the control group after 10 years of follow-up, even when ankle fractures were excluded (88).

Studies reporting the incidence of fractures after GH treatment as compared to a control group were combined in meta-analysis. We combined and compared in separate meta-analysis, the incidence of all fractures and of vertebral fractures at the end of the trial period and at the latest follow-up. The 2 time periods are similar for the Aloia trials and the study by Saaf et al, however for the Landin-Wilhelmsen trial, we have the initial 3-years reported results and the 10-years follow-up report by Krantz et al. A statistically significant 37% drop in the risk of all fractures was observed with GH treatment as compared to control at the latest follow-up period (RR = 0.63, 95% CI [0.46, 0.87], P = 0.004). There was no heterogeneity among included studies (Chi^2= 0.58, df = 2, P = 0.75, I^2 = 0%) (figure 16). No significant difference in all fracture risk was observed at the end of the original trial periods. Similarly, no significant difference was obtained in the occurrence of vertebral fractures both at the end of the initial trials and after including long-term follow-up (table 6).

3.4.7 Adverse events and mortality

Adverse events reported during growth hormone therapy include well-recognized side effects of GH related to fluid retention. They involve peripheral edema, musculoskeletal pain and carpal tunnel syndrome. They were reported in 6 out of 7 included studies (91,96,86,87,74,73) (table 7). One study reported no adverse events (92). The frequency and severity of events was poorly reported in the majority of included studies. When described, symptoms related to fluid retention were usually transient and mostly relieved by decreasing the GH dose. They were also reversible at the end of the treatment period (91,96,74). In the study by Holloway et al, a pre-planned strategy was adopted to relieve the expected side effects of fluid retention. It involved

using oral furosemide 20-40 mg once or twice weekly, and when not sufficient, it allowed a 50% reduction in GH dose. Change in dose was needed in less than 5% of the study population (87). 4 studies reported on treatment withdrawal due to adverse events. They were mainly related to fluid retention. The reported withdrawal rates were 1/23 in Erdtsieck et al, 9/12 in Holloway et al, 3/16 in Saaf et al and 1/80 in Landin-Wilhelmsen et al. (86,87,74,73). In the study by Saaf et al, one patient withdrew due to retinal vein thrombosis and in the Landin-Wilhelmsen trial, one patient revoked due to ichtyosis (74,73). In 5 out of 7 studies, changes in glucose metabolism were actively examined in the study population. No significant changes were reported on GH therapy (91,86,96,74,73).

No mortality was reported during the study period in included studies; however, most studies were of short duration. In the 10-year follow-up on the Landin-Wilhelmsen trial, 6/80 women (8%) died (3 in the placebo group, 2 in low dose GH group and 1 in the high dose group). No death could be directly attributed to GH itself. Causes of death included stroke, myocardial infarction, respiratory insufficiency, pulmonary and kidney cancer. In parallel, 28/223 (12%) women of the control population (a random population sample of postmenopausal women followed for 10 years) died (88).

3.5 Quality of evidence

The quality of evidence was assessed using GRADE for the primary outcome and the main outcomes addressed in meta-analysis (see table 8).

For the main outcome, the quality of the evidence addressing the effect of GH on BMD was low. Although included trials were randomized controlled, and as such of high quality, the quality of the evidence was downgraded secondary to the serious risk

of bias present and to the imprecision of the summary effect size. Indeed, the confidence interval for the mean difference in BMD obtained between GH therapy and a control varied between -0.03 and 0.04.

In studies addressing BMC, the quality of the evidence was moderate. It was only downgraded once secondary to the presence of heterogeneity among included studies.

The quality of the evidence was very low for the fracture risk outcome. The evidence derived from randomized trials and one prospective controlled group was downgraded in this case secondary to significant risk of bias and imprecision in the result. The summary risk ratio was 0.63 favoring GH with a wide confidence interval spanning from 0.46 to 0.87.

In studies combined for the bone formation marker PICP, the quality of evidence was low. It was downgraded secondary to serious risk of bias, serious heterogeneity in trials and the imprecision of the summary measure.

CHAPTER 4

DISCUSSION

4.1 Review and discussion of findings

In the present systematic review and meta-analysis, we found that growth hormone therapy may not be effective in improving bone density in women with age-related bone loss and without GHD. It may however decrease fracture risk.

We retrieved a limited number of studies addressing our outcomes of interest. Included studies only involved women, were characterized by severe osteoporosis (baseline fractures in 5/7 studies), variable GH dosing (including daily and cyclic regimens) and concomitant osteoporosis therapies in all except one study. They involved active treatment between 6 and 24 months. Combining the results of 3 randomized trials examining the effect of GH therapy on BMD at the lumbar spine and femoral neck, no significant difference was observed as compared to treatment with placebo with or without concomitant estrogen or calcitonin. A similar result was obtained for BMD at the total hip addressed in one trial. A non-significant change in BMC was also obtained when synthesizing results of 2 randomized trials comparing GH to placebo with concomitant pamidronate or estrogen. Two other trials addressed BMC at the forearm, yielding a similar negative result. Four trials examined the effect of GH therapy on fracture risk, one of which had a 10-years follow-up extension. A significant decrease in the risk of fracture was only observed when including in the result synthesis the long-term fracture risk results. A trend towards an increase in bone formation markers was observed when combining endpoints of 4 trials addressing this outcome. The trend reached statistical significance for the marker PICP. Bone resorption markers

were only addressed in one included trial with non-significant increase observed as compared to placebo with or without concomitant calcitonin.

The absence of a significant change in BMC and BMD in included studies may be the result of the short duration of GH therapy, whereby trials reporting on BMD extended between 12 and 24 months and the two trials including BMC were of 6 and 18 months duration. Growth hormone has been described to increase bone turnover and expand the bone remodeling space, a process that results in the first few months of therapy in a decrease or no change in BMC and BMD followed by a subsequent time-related increase in bone densitometric endpoints. This is illustrated in the observed non-significant increase in both bone formation and bone resorption markers in this review and in the literature, reflecting a state of increased bone turnover (53). In the study by Joseph et al., increase in the bone formation markers PINP and osteocalcin became higher than the increase in CTX (bone resorption marker) after 6 and 12 months of GH therapy, respectively (97). The effect of such changes in bone turnover with GH treatment are reflected in the results of our previous meta-analysis on GHR in GHD, whereby a significant increase in BMC and BMD is only observed after over one year of therapy (55). In the study by Saaf et al., no significant change in BMD was observed during the 12 months randomized phase of GH therapy versus placebo. It was only in patients who were maintained on a second additional year of GH that a significant increase from baseline was observed in lumbar spine BMD, along with a return to baseline after an initial drop in BMD at the femoral neck (74). Whether older patients treated with GH, such as our patient population, may require a longer duration of therapy to observe significant increases in bone mass remains unanswered but may be postulated. The reason being that remodeling efficiency presumably decreases with

aging, and as such, agents such as GH, that work through initiation of new remodeling spaces may require longer duration for observed efficacy (98,87).

The limited efficacy of GH therapy on bone observed in this review may be secondary to the fact that eligible studies included only women. A differential effect of growth hormone in men and women has been described in GHD. In our meta-analysis on GHR in GHD, in gender-related subgroup analysis, BMD results were only significant in men participating in randomized controlled trials and were consistently higher in men as compared to women (55). The reason being that women might require higher treatment doses than men, namely when on oral estrogen, because of an inhibitory effect of estrogen on GH-induced IGF-1 production (99). In the two largest included trials, GH was administered concomitantly with estrogen replacement therapy, either in the whole study population or in 41% of participants (87,73). The need for a higher dose of GH in that case was not particularly addressed and might have affected observed results. In an uncontrolled trial in 29 men with idiopathic osteoporosis, intermittent or continuous treatment with GH over a period of 24 months resulted in increased BMC and BMD (100).

The majority of included studies in this review (5/7) involved treatment with GH along with an anti-resorptive agent. The potential contribution of this additional therapy to GH effect on bone is unclear. It has been postulated that the use of an anti-resorptive agent such as bisphosphonates may counteract the initial decrease in bone density observed with GH secondary to an increased bone turnover state. In a small trial including six patients with GHD, combination of GH with the bisphosphonate pamidronate resulted in increased BMC in the first six months of therapy, as compared to a drop in BMC in the GH-only group (101). In a randomized controlled trial in 18

GHD patients with osteoporosis on stable GH therapy, the addition of the bisphosphonate alendronate also led to an increase in BMD at 12 months of therapy, as compared to GH alone, an effect obtained presumably through a decrease in bone turnover markers (102). Whether this effect of concomitant GH and bisphosphonate therapy truly holds true is not clear. In the included trial by Erdtsieck et al, treatment with a combination of GH and pamidronate was less effective on BMC than therapy with the bisphosphonate alone. This was accompanied by a decrease in bone turnover markers in the pamidronate group, as compared to no or minimal change in the combination group (86). The addition of an anti-resorptive agent could therefore counteract the mechanism of action of GH through blunting of the expected increase in bone turnover. An analogy to treatment with GH and anti-resorptives can be made regarding the combination therapy of PTH, another anabolic agent with bisphosphonates. In a randomized trial in men with osteoporosis, BMD increased significantly more in patients treated with parathyroid hormone alone than in those treated with alendronate alone or with a combination of both, an effect that was potentially attributable to the attenuation of PTH-induced increase in bone formation markers (103). In contrast, the addition of another anti-resorptive agent, denosumab, to PTH increased BMD and improved bone microarchitecture more than either medication alone (104,105). However, whether this combination reduces the risk of fracture remains to be demonstrated.

A concerning characteristic of included studies is the use of different doses and treatment regimens of GH. In two studies, GH was administered daily (74,73) and in the rest of the retained trials, GH was given either 3 times weekly or in cycles every 2-6 months. The presence of such heterogeneity in therapy might have affected the observed

results. In our prior meta-analysis on GHR in GHD, no significant association was obtained between the treatment dose and the change in BMD. However, a positive association was observed with on-therapy IGF-1 level, a reflection of treatment effect (55). In all included studies in the present meta-analysis, a significant increase from baseline in serum IGF-1 level was observed with GH therapy, in parallel to GH administration with a quick return to baseline with treatment cessation. Whether these reversible changes in IGF-1 levels affected bone densitometric endpoints in trials where GH was administered cyclically is plausible. In two trials, GH was given for 1 week every 49-82 days and as such the change in IGF-1 was limited to this short treatment period (96,87).

Despite a non-significant change in bone densitometric endpoints, the current meta-analysis showed a significant decrease in fracture risk with GH therapy as compared to a control treatment. This differential effect of GH on bone-related endpoints might reflect a beneficial effect on bone beyond the potential impact on bone density. In a prospective trial studying the link between osteoporotic fractures and IGF-1 in postmenopausal women, decreased serum IGF-1 level predicted the risk of fractures independently of BMD, and as such IGF-1 was suggested to carry an important role in maintaining bone strength and bone quality. The mechanism of such an effect remains, however, unclearly defined (72). The differential effect of GH on fracture risk may also reflect differences in studies evaluating this outcome, as compared to trials including bone densitometric endpoints. An important distinguishing trial is the 10-year follow up study by Krantz et al reporting long-term fracture risk with GH (88). The inclusion of this study with prolonged duration might have potentially resulted in beneficial fracture outcomes, reflecting the need for extended follow-up to

observe clinically significant results in osteoporosis trials. This trial, however, has significant caveats; the actual randomized trial comparing GH to placebo extended for only 18 months (73). The actual Krantz study compared the initial GH treatment group to a control group of age-matched postmenopausal women. The two groups differed in their BMI, level of physical activity and their use of osteoporosis therapies. In the GH group, at the end of follow-up, 59% of women were on HRT, 23% were started on bisphosphonates and 3% were on teriparatide. Meanwhile, in the control group, at the end of the study, 8% were on estrogen and the use of other osteoporosis therapies reached 4% (88). The groups were thus not matched with regards to their disease status and therapy, which might have affected and interacted with any potential effect of GH treatment on bone.

Any potential use of GH as a therapeutic agent has to take into account its potential safety. In the present systematic review, adverse events reported with GH were namely related to fluid retention, were often transient and reversible, and seldom resulted in treatment withdrawal. No mortality was reported in the included trials, except for the long-term follow-up study by Krantz et al. whereby a mortality rate of 8% was detected in patients initially treated with GH, without being directly attributable to treatment. The death rate was lower than that reported in the control population (12%) (88). The safety of GH replacement has been recently appraised by international professional societies involved with such treatment (106). In the consequent position paper, aggregate evidence did not support an association between GH therapy and all-cause mortality and data for cancer risk was reassuring. The reviewed literature again recognized that the most common side effects related to GH were musculoskeletal symptoms associated with fluid retention, potential for exacerbation of obstructive sleep

apnea and increased risk of glucose intolerance and diabetes mellitus in individuals at increased risk (106). The safety of GH treatment in the absence of deficiency, however, cannot be equated to that of GH replacement.

4.2 Limitations and strengths

The current review has several limitations, arising from the nature of available primary data. The retrieved number of articles was small with only 7 studies being included, with the majority of the studies having a small sample size. Except for the studies by Holloway and Landin-Wilhelmsen that involved 84 and 80 patients respectively, the rest of the trials had a sample size ranging between 14 and 30. The representativeness of this review was also limited by the fact that it only included women. Indeed, retained articles did not involve men and as such, the effect of GH therapy on bone in this patient population could not be examined. Significant heterogeneity was also noted among included studies in terms of modes of GH therapy. GH was given either continuously or cyclically, in a fixed dosing regimen or using a weight-based calculation, in the form of human growth hormone or recombinant hormone. Such variability in GH administration could have potentially affected the obtained results, and limited any conclusive remarks about the ideal mode of GH treatment. In addition, the quality of several included trials could be classified as low which resulted in downgrading of the quality of included evidence. Finally, similarly to any meta-analysis addressing osteoporosis outcomes, the present review suffered from the potential of variability that could appear with the use of different assays to measure bone turnover markers in the different studies, along with a variability in the DXA

machines used to measure bone densitometry and the different definitions of osteoporotic fractures.

However, this review and meta-analysis fills an important knowledge gap in GH therapeutics. It is the first systematic review that specifically addresses the effect of GH treatment on bone in postmenopausal women and men above the age of 50 with osteopenia and osteoporosis and without organic GH deficiency. Existing reviews addressed GH use in bone healing after fracture and effects of GH on multiple outcomes, including bone, in healthy elderly (not known to have osteoporosis at study entry). The literature search was extensive, including three important databases and relevant references. The review methodology was designed according to the evidence-based PRISMA suggestions and 2 independent reviewers worked on different steps in the work. Selection and eligibility criteria were carefully chosen to limit the study to the population and intervention of interest and to overcome limitations of prior reviews on the topic. Moreover, the effect of GH on different outcomes addressing bone was studied, including true clinical endpoints such as fractures and surrogate outcomes including bone densitometry and bone turnover markers. Finally, the present review revealed the different studies available addressing the effect of GH on bone along with the quality of the available evidence and permitted as such to identify knowledge gaps in the topic and areas requiring future research.

4.3 Conclusion and recommendations

Growth hormone treatment, according to the available evidence, may not be effective in improving bone density in women with age-related bone loss. It may however decrease fracture risk, without significant adverse events. Whether this

conclusion holds true universally remains unanswered. Existing literature suffers indeed from significant caveats. There is as such a pressing need for designing randomized controlled trials addressing the effect of a stable extended treatment with recombinant GH therapy in men and women with age-related decline in bone density on both bone density and bone quality and on long-term fracture risk.

Figure 1. Flow diagram of study selection

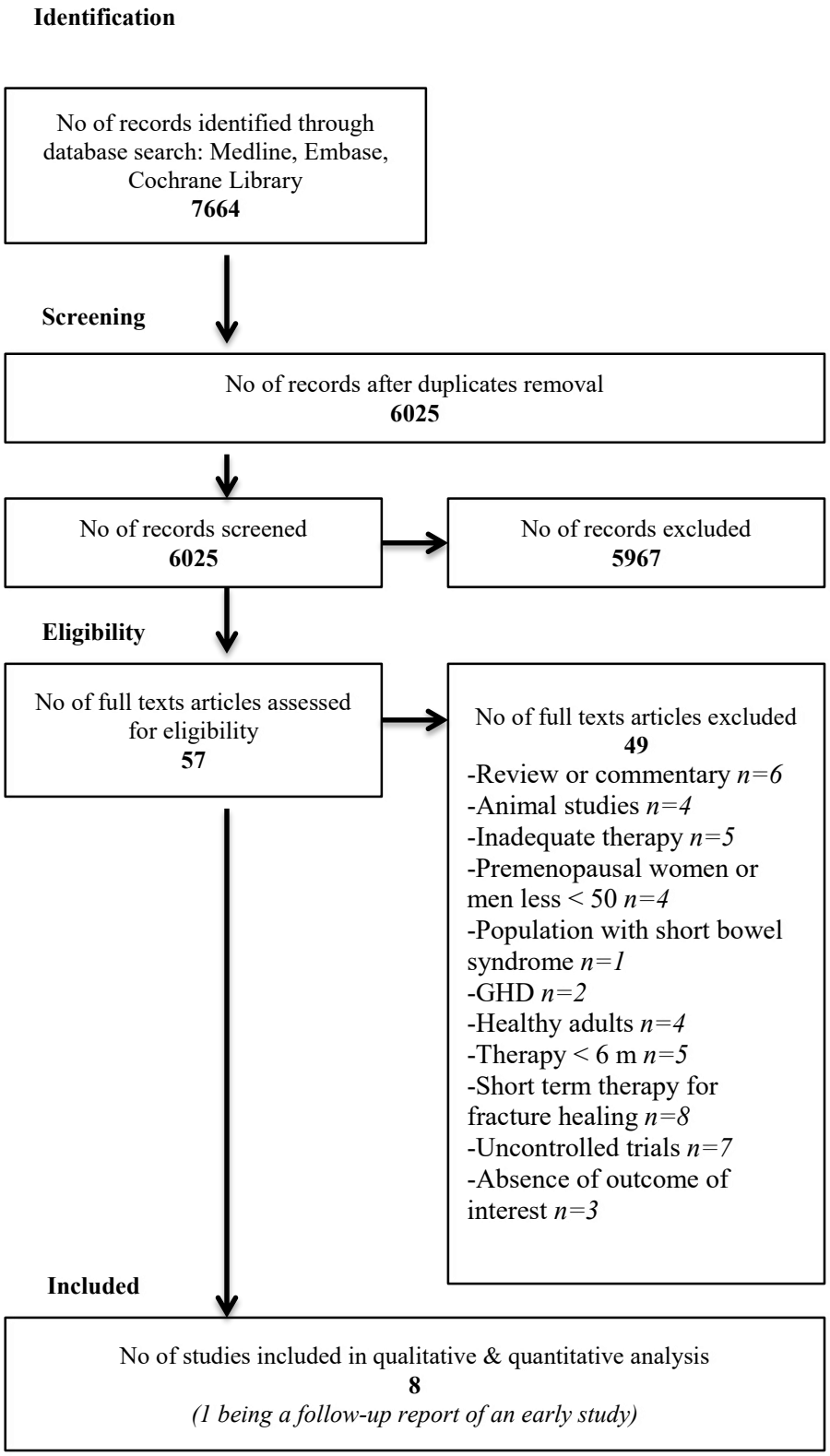
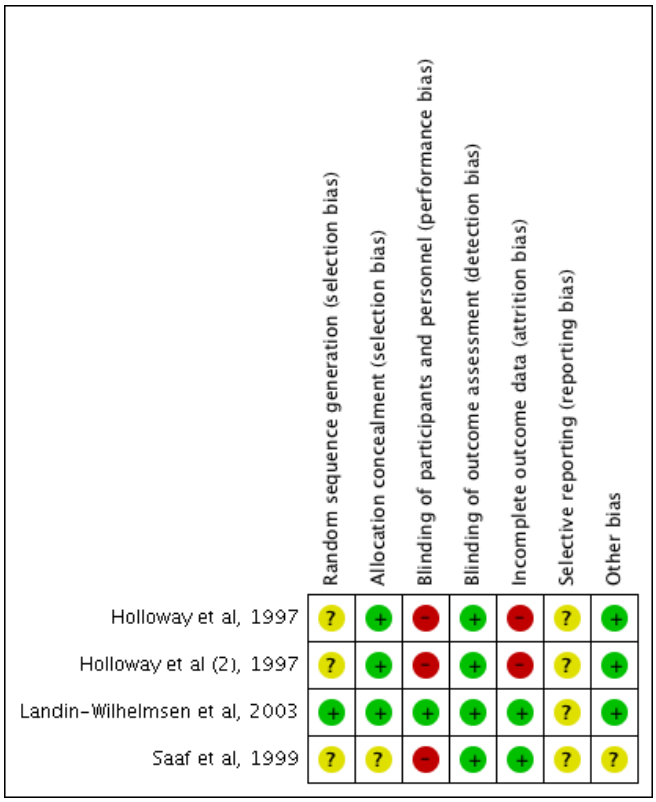


Figure 2: Risk of bias assessment across studies evaluating BMD

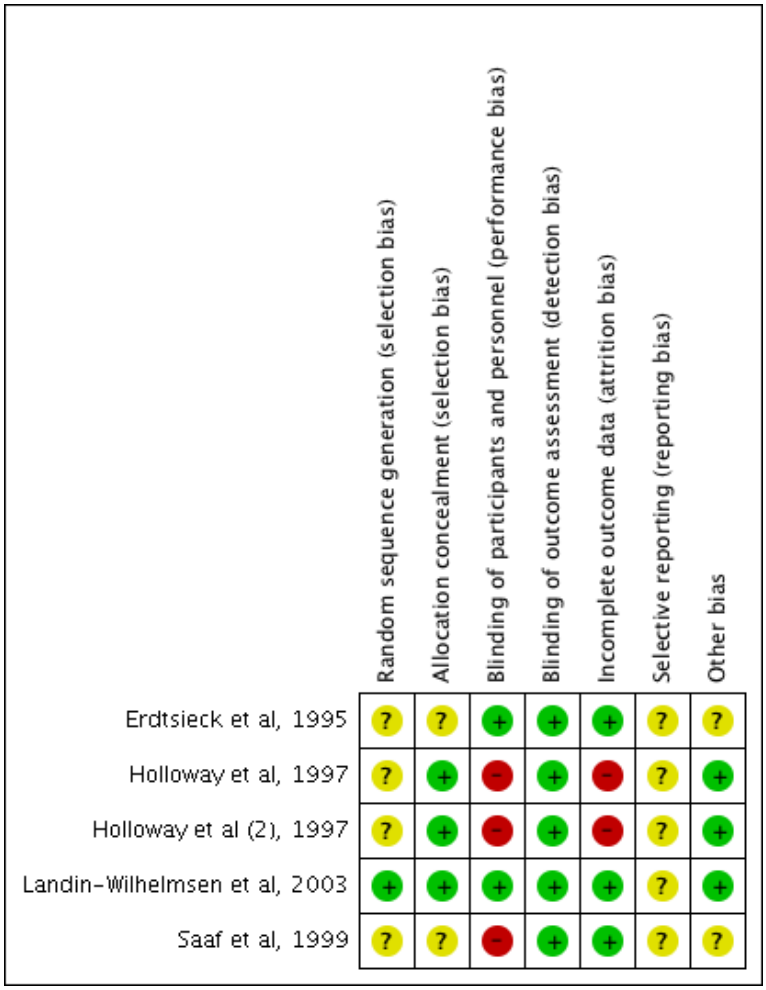


A green color and (+) sign represent a low risk of bias; A yellow color and (?) sign represent an unclear risk of bias; A red color and a (-) sign represent a high risk of bias

Figure 3: Risk of bias assessment across studies evaluating BMC

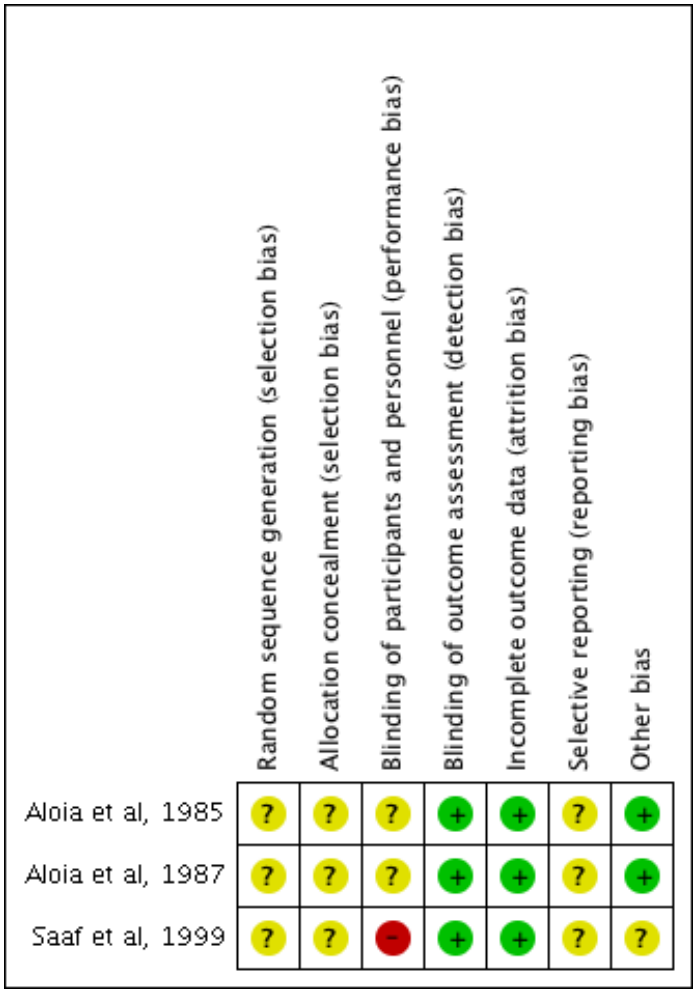
A green color and (+) sign represent a low risk of bias; A yellow color and (?) sign represent an unclear risk of bias; A red color and a (-) sign represent a high risk of bias.

Figure 4: Risk of bias assessment across studies evaluating PICP



A green color and (+) sign represent a low risk of bias; A yellow color and (?) sign represent an unclear risk of bias; A red color and a (-) sign represent a high risk of bias.

Figure 5: Risk of bias assessment across studies evaluating fracture risk



A green color and (+) sign represent a low risk of bias; A yellow color and (?) sign represent an unclear risk of bias; A red color and a (-) sign represent a high risk of bias.

Figure 6: Growth hormone effect versus control on BMD at the lumbar spine

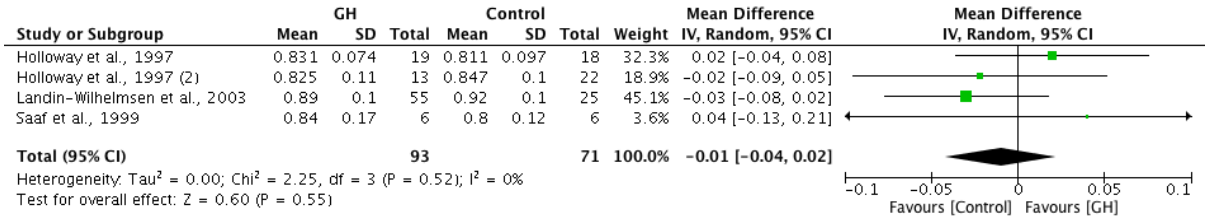


Figure 7: Growth hormone effect versus control on BMD at the total hip

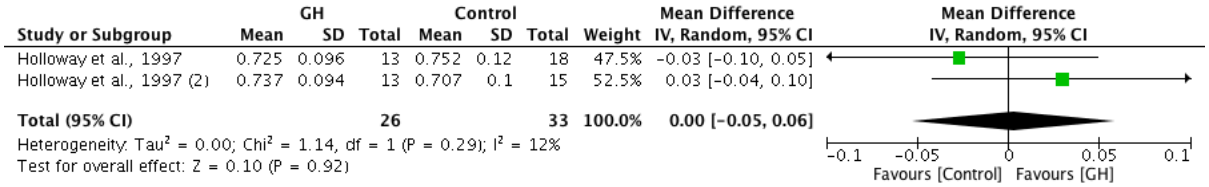


Figure 8: Growth hormone effect versus control on BMD at the femoral neck

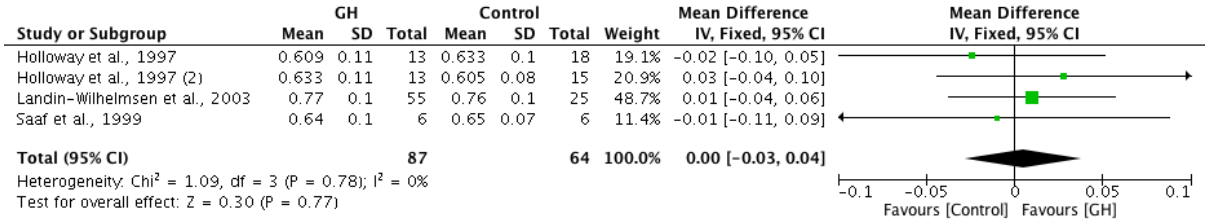


Figure 9: Growth hormone effect versus control on BMC at the lumbar spine

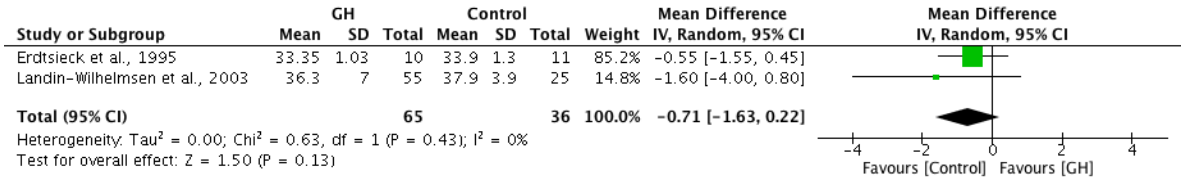


Figure 10: Growth hormone effect versus control on BMC at the femoral neck

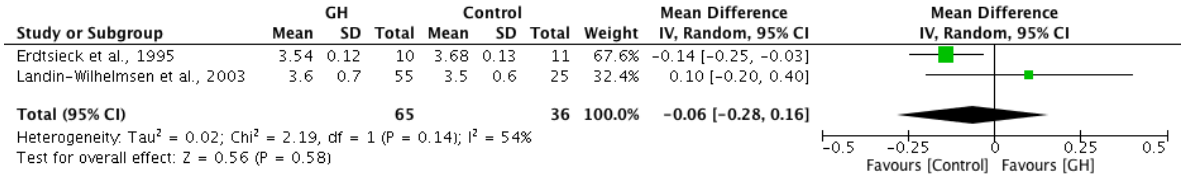


Figure 11: Growth hormone effect versus control on BMC at the forearm

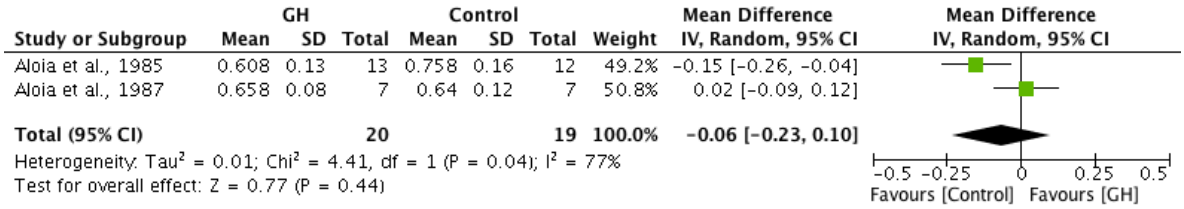


Figure 12: Growth hormone effect versus control on osteocalcin

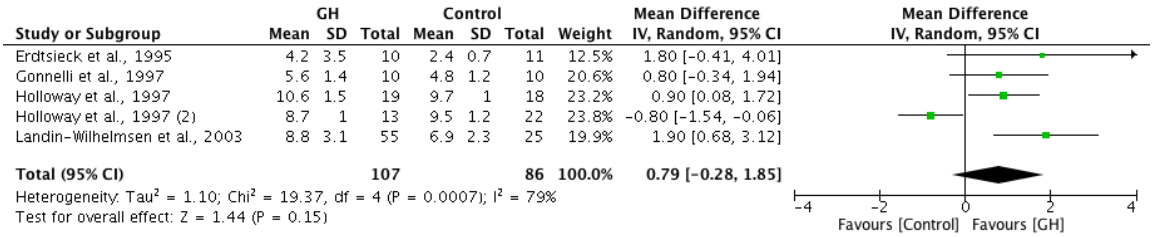


Figure 13: Growth hormone effect versus control on PICP

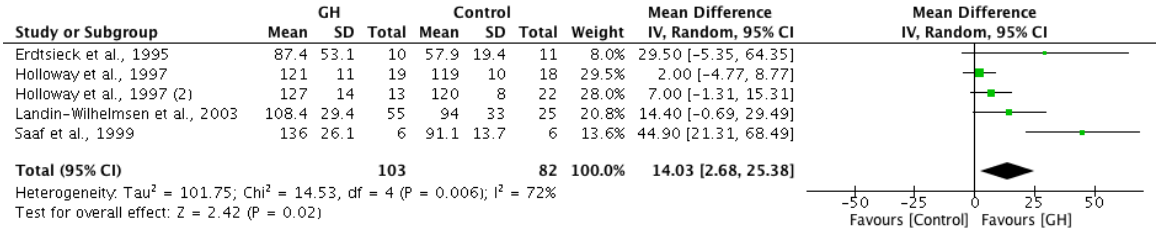


Figure 14: Growth hormone effect versus control on total urinary pyridinolines

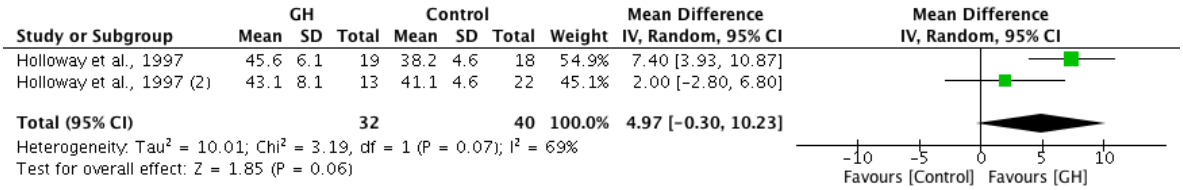


Figure 15. Growth hormone effect versus control on CTX

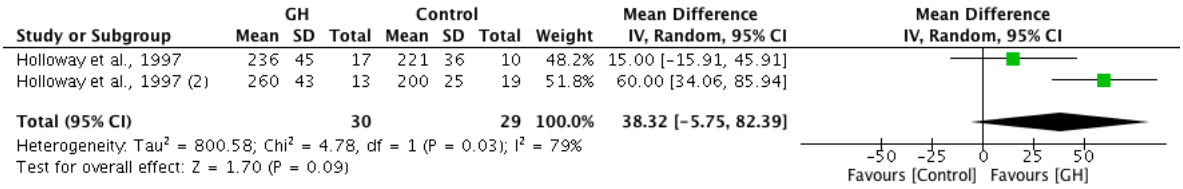


Figure 16: Growth hormone effect versus control on fracture risk

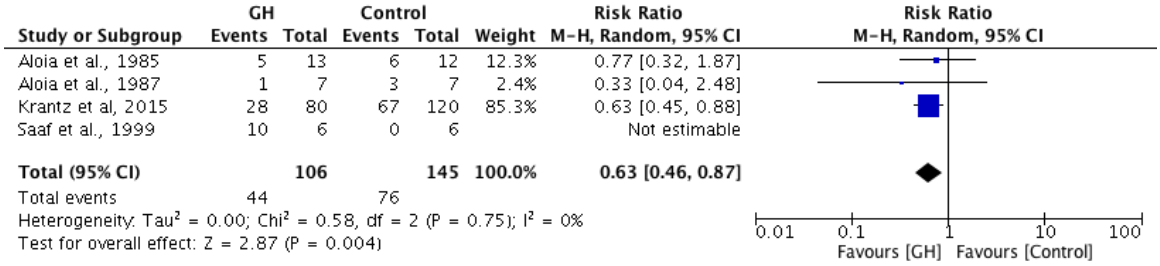


Figure 17: Growth hormone effect versus control on vertebral fracture risk

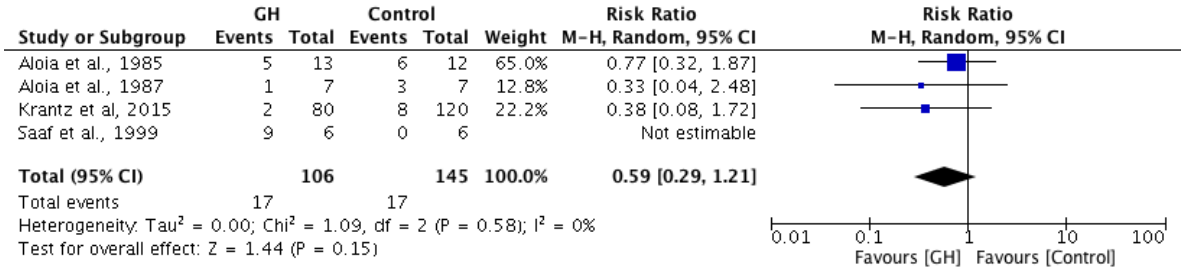


Table 1: Characteristics of included studies

Participants								Osteoporosis/Osteopenia					Intervention			Outcome		
First author, year	Study design	Age (yrs)	Yrs since menopause	Gender (% M/F)	BMI (kg/m2)	No. of subjects	No. withdrawn	Yrs since diagnosis	Prior therapy	Concurrent therapy	Number/site of fractures	Disease definition	Baseline T-score	Approximate mean target GH dose [IU/d (mg/d)], regimen	Control	Duration of therapy (m)	Outcome assessed	Funding	Conflict of interest
Aloia *et al*, 1985	R, C	64.3 ± 7.3	NA	0/100	25.2 ± 4.5	25	0	NA	NA	Ca 1000 mg/d, CT 100 MRC U 4*/w	All ≥ 1 VFx (3.4 ± 2.15)	VFx (Loss of height >25%)	NA	hGH 6/d (2/d), 3*/w	CT	24	BMC FA (SPA), Fractures	GH (National Pituitary Agency), CT (Armour Pharmaceuticals)	–
Aloia *et al*, 1987	R, C	62.4 ± 6.7	NA	0/100	NA	14	0	NA	NA	Ca 1000-1200 mg/d, CT 100 MRC U cyclic	All ≥ 1 VFx (3.05 ± 1.85)	VFx (Loss of height >25%)	NA	hGH 7/d (2.3/d), cyclic (hGH 2m, CT 3m, rest 3m for 3 cycles)	CT	24	BMC FA (SPA), Fractures	GH (National Pituitary Agency), CT (Armour Pharmaceuticals)	–
Erdtsieck *et al*, 1995	R, DB, PC	63.1 (55-74)	17.2 (5-30)	0/100	25.3 (19-32.5)	23	2	NA	NA	Ca elemental 500 mg/d (if dietary Ca < 1000	All ≥ 1 VFx	VFx (clinical assessment)	NA	rhGH 0.0625/kg, up to 4 (0.02/kg), 3*/w	Placebo	6	BMC LS,FN (Lunar), FA (SPA), Bone markers	Eli Lilly (partly)	Eli Lilly provided the rhGH and

									mg/d), Pamidronate 150 mg/day									financial support
Gonnelli et al, 1997	R, PC	61.2 ± 5.6	11.8 ± 7.5	0/100	24.4	30	4	0	none	CT 50 IU/d cyclic	All ≥ 1 VFx	VFx	NA	rhGH 12/d (4/d), cyclic (rhGH 7d, CT 21d, rest 61d for 8 cycles)	Placebo	24	BMD(DPA), Bone markers	–	–
Holloway *et al*, 1997	R, DB, PC	69.2 ± 6.6	NA	0/100	NA	84	12	0	estrogen 41%	Ca 500 mg/d, HRT 41%, ± CT 100 U/d	0	BMD LS T score < -1	NA	rhGH 0.06/kg/d (0.02 mg/kg/d), cyclic (rhGH 7 d, CT 5d, rest 44d Ca for 12 cycles)	Placebo	24	BMD LS, TF, FN (Hologic), Bone markers	Medications provided by suppliers	–
Saaf *et al*, 1999	R, DB, PC	67.8 ± 4.4 (60-74)	≥ 5	0/100	24.6 ± 3.2 (20.2-33.2)	16	4	NA	Ca/MV	Ca 500 mg/d	All ≥ 1 VFx, 2 hip Fx, 1 shoulder Fx, 7 wrist Fx	BMD < 1 g/cm2 in ultradistal radius & presence of 1-4	-2.8 ± 0.72 (TB)	rhGH 1.5 (±0.3)/d (0.5/d) Range 0.9-2.7/d (0.3-0.9/d)	Placebo	12	BMD LS,FN (Lunar), Bone markers, Fractures	Swedish Medical Research Council + foundations	–

											VFx (Loss of height >15%)							
Landin-Wilhelmsen *et al*, 2003	R, DB, PC	60.7 ± 5.7	NA	0/100	NA	80	0	1.9 ± 2.3	HRT	Ca 750 mg/d, D 400 U/d, HRT	56% with VFx	BMD LS T score < -2.5	-2.7 (LS)	rhGH 1/d (0.33/d) or 2.5/d (0.83/d)	Placebo	18	BMD/ BMC LS,FN (Lunar), Bone markers, Fractures	Pharmacia Upjohn & Goteborg University	–
Krantz *et al*, 2015	C	71 ± 6	NA	0/100	24.6 ± 3.2	80	6	11.9 ± 2.3	HRT (41% stopped), Bisphosphonates (23%), Teriparatide (3%)	Ca 750 mg/d, D 400 U/d, HRT	56% with VFx	BMD LS T score < -2.5	-2.7 (LS)	NA	Control Population	NA	Fractures	–	–

R, Randomized; DB, Double-blind; PC, Placebo-controlled; C, Controlled; NA, Not available, M, Male; F, Female; BMI, Body mass index; Ca, Calcium; MV, Multivitamin; HRT, Hormone replacement therapy; CT, Calcitonin; MRC, Medical Research Council; VFx, Vertebral fracture; BMD, Bone mineral density; LS, Lumbar spine; FN, Femoral neck; TB, Total body; GH, Growth hormone; hGH, Human growth hormone; rhGH, Recombinant human growht hormone; D, day; W, week; M, month; BMC, Bone mineral content; FA, Forearm; SPA, Single-photon absorptiometry; LS; lumbar spine, FN, Femoral neck; DPA, Dual-photon absorptiometry; TF, Total femur. Data are shown as mean ± SD or mean (range) if not otherwise specified

Table 2: Risk of bias assessment across studies by outcome

Studies evaluating BMD

Author, Year	Random sequence generation (Selection bias)	Allocation concealment (Selection bias)	Blinding of participants and personnel (Performance bias)	Blinding of outcome assessment (Detection bias)	Incomplete outcome data (Attrition bias)	Selective outcome reporting (Reporting bias)	Other sources of bias	Summary Assessment
Holloway *et al*, 1997	"Randomization performed by pharmacist" but no further details on randomization method ***Unclear risk***	The pharmacist who did the randomization "selected a sealed envelope containing the study drug assignment for each participant" ***Low risk***	"Study drugs were administered from masked multidose vials" the participants were thus blinded to their treatment. However personnel not clearly blinded as doses of GH or CT were adjusted according to side effects ***High risk***	Blinding of outcome assessment was not described. But outcome not affected by the assessor blinding ***Low risk***	"Twelve women withdrew entirely from the study. Of these, 9 had been assigned to GH. Missing data for BMD at total femur and femoral neck" ***High risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	"No significant differences were observed between the baseline characteristics of the original study cohort and those who competed the protocol", "intention-to-treat strategy" ***Low risk***	High risk
Saaf *et al*, 1999	No description provided ***Unclear risk***	No description provided ***Unclear risk***	"Double blind schedule" but no further description. Personnel not clearly blinded as dose of GH adjusted throughout the trial based on adverse events ***High risk***	Blinding of outcome assessment was not described. But outcome not affected by the assessor blinding ***Low risk***	"Six patients in the placebo/GH group and six patients in the GH/GH group completed one year of GH therapy", number of withdrawals is the same between groups ***Low risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	Use of intention-to-treat analysis is not mentioned ***Unclear risk***	High risk

Landin-Wilhelmsen *et al*, 2003	"The patients were randomized in blocks" ***Low risk***	"Computerized allocation to treatment groups" ***Low risk***	"Neither the investigator nor the patients were aware of the type of treatment" **Low *risk***	"Data was analyzed by an external statistician" ***Low risk***	"There were no drop-outs, and no code was broken before termination of the double-blind phase" ***Low risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	Use of intention-to-treat analysis ***Low risk***	Unclear risk

Studies evaluating BMC

Erdtsieck *et al*, 1995	No description provided ***Unclear risk***	No description provided ***Unclear risk***	"Treatment was double blind, placebo vials were indistinguishable from rhHG vials" ***Low risk***	Blinding of outcome assessment was not described. But outcome not affected by the assessor blinding ***Low risk***	Withdrawal: 1/12 patient in the control group and 1/11 in the GH group ***Low risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	Unclear how patient withdrawal or missing data was handled ***Unclear risk***	Unclear risk
Landin-Wilhelmsen *et al*, 2003	"The patients were randomized in blocks" ***Low risk***	"Computerized allocation to treatment groups" ***Low risk***	"Neither the investigator nor the patients were aware of the type of treatment" **Low *risk***	"Data was analyzed by an external statistician" ***Low risk***	"There were no drop-outs, and no code was broken before termination of the double-blind phase" ***Low risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	Use of intention-to-treat analysis ***Low risk***	Unclear risk

Studies evaluating PICP

Erdtsieck *et al*, 1995	No description provided ***Unclear risk***	No description provided ***Unclear risk***	"Treatment was double blind, placebo vials were indistinguishable from rhHG vials" ***Low risk***	Blinding of outcome assessment was not described. However the outcome bone marker is not affected by knowledge of the treatment assignment ***Low risk***	Withdrawal: 1/12 patient in the control group and 1/11 in the GH group ***Low risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	Unclear how patient withdrawal or missing data was handled ***Unclear risk***	Unclear risk
Saaf *et al*, 1999	No description provided ***Unclear risk***	No description provided ***Unclear risk***	"Double blind schedule" but no further description. Personnel not clearly blinded as dose of GH adjusted throughout the trial based on adverse events ***High risk***	Blinding of outcome assessment was not described. However the outcome bone marker is not affected by knowledge of the treatment assignment ***Low risk***	"Six patients in the placebo/GH group and six patients in the GH/GH group completed one year of GH therapy", number of withdrawals is the same between groups ***Low risk***	No published protocol, Outcomes included in methods section reported in results ***Unclear risk***	Use of intention-to-treat analysis is not mentioned ***Unclear risk***	High risk

Holloway *et al*, 1997	"Randomization performed by pharmacist" but no further details on randomization method ***Unclear risk***	The pharmacist who did the randomization "selected a sealed envolope containing the study drug assignment for each participant" ***Low risk***	"Study drugs were administered from masked multidose vials" the participants were thus blinded to their treatment. However personnel not clearly blinded as doses of GH or CT were adjusted according to side effects ***High risk***	Blinding of outcome assessment was not described. However the outcome bone marker is not affected by knowledge of the treatment assignment ***Low risk***	"Twelve women withdrew entirely from the study. Of these, 9 had been assigned to GH" ***High risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	"No significant differences were observed between the baseline characteristics of the original study cohort and those who competed the protocol", "intention-to-treat strategy" ***Low risk***	High risk
Landin-Wilhelmsen *et al*, 2003	"The patients were randomized in blocks" ***Low risk***	"Computerized allocation to treatment groups" ***Low risk***	"Neither the investigator nor the patients were aware of the type of treatment" **Low *risk***	"Data was analyzed by an external statistician" ***Low risk***	"There were no drop-outs, and no code was broken before termination of the double-blind phase" ***Low risk***	No published protocol, outcomes included in methods section reported in results ***Unclear risk***	Use of intention-to-treat analysis ***Low risk***	Unclear risk

Studies evaluating fracture risk

Aloia et al, 1985	No description provided ***Unclear risk***	No description provided ***Unclear risk***	It was not mentioned that blinding was done, study described only as randomized ***Unclear risk***	Blinding of outcome assessment was not described. Fractures evaluated by X-rays and are thus not affected by the knowledge of the treatment assignment **Low *risk***	No withdrawal from the trial ***Low risk***	No published protocol, fractures not mentioned as an outcome ***Unclear risk***	Study population same throughout the trial ***Low risk***	Unclear risk
Aloia et al, 1987	No description provided ***Unclear risk***	No description provided ***Unclear risk***	It was not mentioned that blinding was done, study described only as randomized ***Unclear risk***	Blinding of outcome assessment was not described. Fractures evaluated by X-rays and are thus not affected by the knowledge of the treatment assignment **Low *risk***	No withdrawal from the trial ***Low risk***	No published protocol, fractures not mentioned as an outcome ***Unclear risk***	Study population same throughout the trial ***Low risk***	Unclear risk

Saaf *et al*, 1999	No description provided ***Unclear risk***	No description provided ***Unclear risk***	"Double blind schedule" but no further description. Personnel not clearly blinded as dose of GH adjusted throughout the trial based on adverse events ***High risk***	Blinding of outcome assessment was not described. Fractures evaluated by X-rays and are thus not affected by the knowledge of the treatment assignment **Low *risk***	"Six patients in the placebo/GH group and six patients in the GH/GH group completed one year of GH therapy", number of withdrawals is the same between groups ***Low risk***	No published protocol, fractures not mentioned as an outcome ***Unclear risk***	Use of intention-to-treat analysis is not mentioned ***Unclear risk***	Unclear risk

Evaluation of risk of bias in the controlled observational study by Krantz

"Do inclusion/exclusion criteria vary across groups?"	*"Does the strategy for recruiting participants vary across groups?"*	*"Is the selection of the comparison group inappropriate, after taking into account feasibility and ethical considerations?"*	*"Does the study fail to account for important variations in the execution of the study from the proposed protocol?"*	*"Was the outcome assessor not blinded to the intervention or exposure status of participants?"*	*"Were valid and reliable measures, implemented consistently across all study participants…?"*	*"Was the length of follow-up different across study groups?"*

Yes, differs "Intervention group selected specifically patients with osteoporosis, control group included postmenopausal women"	**Yes, differs** "The intervention group was selected among patients actively coming for screening, while the control group was selected from the city census"	**Yes, inappropriate** "The control population did not have comparable initial characteristics as the intervention group"	**Cannot determine** "no protocol available"	**No, blinded** "Neither the investigator nor the patients were aware of the type of treatment"	**Yes, valid and reliable** measure used "All outcome measured were clearly decribed along with their precision, accuracy…"	**Yes, different** "Women with exposure followed for 10 yrs, the controls for a mean of 12 yrs"

"In cases of high loss to follow-up (or differential loss to follow-up), was the impact assessed ?"	"Are any important primary outcomes missing from the results?"	"Are any important harms or adverse events that may be a consequence of the intervention/exposure missing from the results?"	"Are results believable taking study limitations into consideration?"	"Any attempt to balance the allocation between the groups or match groups?"	"Were important confounding variables not taken into account in the design and/or analysis?"

No, impact not assessed "7.5% died in exposure grp and 38% lost in control grp"	**No important outcomes missing** "BMD, fractures, adverse events reported	**No adverse events missing** "adverse events reported for the GH group"	**No, not believable** "different inclusion criteria, length of follow-up, no matching"	**No matching was done** "A random control grp was selected from the populaiton"	**Yes, many potential confounders were not accounted for**, like intial disease status, HRT tx, presence of fractures

Table 3: Summary of meta-analysis results for BMD and BMC

Outcome	**Studies Included**	**Sample Size** GH/Control	**WMD** (95% CI)	**Test of null** Z value (*p*-value)	**Test of heterogeneity** Chi2 (p-value)	**Inconsistency** I^2
BMD LS	Holloway 1997 Saaf 1999 Landin-Wilhelmsen 2003	93/71	-0.01 (-0.04; 0.02)	0.60 (0.55)	2.25 (0.52)	0%
BMD TF	Holloway 1997	26/33	0 (-0.05; 0.06)	0.10 (0.92)	1.14 (0.29)	12%
BMD FN	Holloway 1997 Saaf 1999 Landin-Wilhelmsen 2003	87/64	0.00 (-0.03; 0.04)	0.30 (0.77)	1.09 (0.78)	0%
BMC LS	Erdtsieck 1995 Landin-Wilhelmsen 2003	65/36	-0.71 (-1.63; 0.22)	1.50 (0.13)	0.63 (0.43)	0%
BMC FN	Erdtsieck 1995 Landin-Wilhelmsen 2003	65/36	-0.06 (-0.28; 0.16)	0.56 (0.58)	2.19 (0.14)	54%
BMC Distal FA (SPA)	Aloia 1985 Aloia 1987	20/19	-0.06 (-0.23; 0.10)	0.77 (0.44)	4.41 (0.04)	77%

GH, Growth Hormone; WMD, Weighted Mean Difference; CI, Confidence Interval; Z value, Test statistic of the WMD; Chi2, Chi-squared test of heterogeneity; BMD, Bone Mineral Density; LS, Lumbar Spine; TF, Total Femur; FN, Femoral Neck; BMC, Bone Mineral Content; FA, Forearm; SPA, Single Photon Absorptiometry

Table 4: Detailed description of changes in bone markers in individual studies

	Studies Included	Study Design	Number of Participants	Assay	Treatment Regimen	Treatment Duration (months)	Results		
							Change from baseline	**Comparison between groups**	**Temporal change**
Bone Marker									
Bone Formation Markers									
Osteocalcin (µg/L)	Erdtsieck *et al* 1995	R, DB, PC	21	RIA	GH (3x/weekly)/ Pamidronate **v/s** Pl/Pamidronate	6	Significant decrease from baseline in Pl group	Significant difference favoring GH at end of therapy	Significant decrease in Pl and significant difference from GH as of 3m, maintained till 12 m (uncontrolled phase)
	Gonnelli et al 1997	R, PC	30	RIA	GH(7d)/CT(21d)/Rest(61d) **v/s** Pl/CT/Rest **v/s** GH/Pl/Rest	24	No significant change	No significant difference	Significant increase from baseline in GH groups and significant difference between GH groups & Pl/CT at week 1 of therapy, significant difference maintained at 1 month
	Holloway *et al* 1997	R, DB, PC	72	RIA	Cycles of 7d GH or Pl, 5d CT or Pl, 44d rest	24	Significant increase from baseline in GH groups	No significant difference	Significant increase as of week 1 of therapy

	Saaf *et al* 1999	R, DB, PC	12	IRMA	GH daily v/s Pl	12	Significant increase (88 ± 21%SE) from baseline in GH group	No significant difference	Significant increase as of 6 m, maximal at 9 m, still significant at 12 m
	Landin-Wilhelmsen *et al* 2003	R, DB, PC	80	International CIS	GH daily/HRT v/s Pl/HRT	18	Significant increase from baseline in GH groups	Significant difference favoring GH at end of therapy	–
PICP (μg/L)	Erdtsieck *et al* 1995	R, DB, PC	21	RIA	GH (3x/weekly)/ Pamidronate **v/s** Pl/Pamidronate	6	Significant decrease from baseline in Pl group	No significant difference	Significant decrease in Pl as of 3m and maintained at 12m (uncontrolled phase)
	Holloway *et al* 1997	R, DB, PC	72	Immunoassay	Cycles of 7d GH or Pl, 5d CT or Pl, 44d rest	24	No significant change	No significant difference	Significant increase as of week 1 of therapy, less increase at week 3 then non-significant decrease
	Saaf *et al*, 1999	R, DB, PC	12	RIA	GH daily v/s Pl	12	Significant increase (36 ± 11%SE) from baseline in GH group	No significant difference	Significant increase as of 6 m, maximal at 12 m
	Landin-Wilhelmsen *et al*, 2003	R, DB, PC	80	Orion Diagnostica	GH daily/HRT v/s Pl/HRT	18	Significant increase from baseline in GH groups	No significant difference	Significant increase from baseline at 18 m, maintained at 3 years (uncontrolled phase)

B-ALP (µg/L)	Landin-Wilhelmsen *et al*, 2003	R, DB, PC	80	ELISA	GH daily/HRT v/s Pl/HRT	18	No significant change	No significant difference	No significant change
PIIINP (U/L)	Saaf *et al*, 1999	R, DB, PC	12	RIA	GH daily v/s Pl	12	Significant increase (26 ± 8%SE) from baseline in GH group	No significant difference	Significant increase from baseline in GH group as of 6 m (maximal), still significant at 12 m
Bone Resorption Markers									
Urine fasting (2-hr) hydroxyproline (mmol/mmolCr)	Erdtsieck *et al*, 1995	R, DB, PC	21	Hypronosticon	GH (3x/weekly)/ Pamidronate **v/s** Pl/Pamidronate	6	–	No significant difference	–
24hr urine hydroxyproline (mg/gCr)	Gonnelli et al, 1997	R, PC	30	Hypronosticon	GH(7d)/CT(21d)/Rest(61d) **v/s** Pl/CT/Rest **v/s** GH/Pl/Rest	24	No significant change	No significant difference	Significant increase from baseline in GH/Pl group at 1 week
Urine fasting (2-hr) free deoxypyridinoline (nmol/molCr)	Erdtsieck *et al*, 1995	R, DB, PC	21	ELISA	GH (3x/weekly)/Pamidronate **v/s** Pl/Pamidronate	6	–	No significant difference	Significant decrease from baseline in Pl group as of 3 m, maintained at 6 m then non-significant increase
Total urinary pyridinolines (nmol/mMCr)	Holloway *et al*, 1997	R, DB, PC	72	Competitive Enzyme Immunoassay	Cycles of 7d GH or Pl, 5d CT or Pl, 44d rest	24	No significant change	No significant difference	Significant increase from baseline in GH/CT group at week 3 then non significant decrease

Free Pyridinoline Cross-Links (nmol/mmolCr)	Gonnelli et al, 1997	R, PC	30	ELISA	GH(7d)/CT(21d)/Rest(61d) **v/s** Pl/CT/Rest **v/s** GH/Pl/Rest	24	No significant change	No significant difference	Significant increase from baseline in GH groups and significant difference from CT/Pl at week 1 . **At 1 m**, CT significant drop from baseline, GH/Pl significantly higher than CT
CTX (µg/mMCr)	Holloway *et al*, 1997	R, DB, PC	72	ELISA	Cycles of 7d GH or Pl, 5d CT or Pl, 44d rest	24	No significant change	No significant difference	Significant increase from baseline in GH groups at week 1 of therapy, then non-significant
ICTP (µg/L)	Landin-Wilhelmsen *et al*, 2003	R, DB, PC	80	Orion Diagnostica	GH daily/HRT v/s Pl/HRT	18	Significant increase from baseline in GH groups	Significant difference favoring high dose GH at end of therapy	Significant increase from baseline at 18 m in GH groups and significant difference from Pl in high dose GH group, maintained at 3 years (uncontrolled phase)

PICP, Procollagen Type-I Carboxy-terminal Propeptide; B-ALP, Bone-specific Alkaline Phosphatase; PIIINP, N-terminal Propeptide of Type-III Collagen; PINP, Procollagen Type-I Amino-Terminal Propeptide; CTX, Carboxy-terminal Collagen Crosslinks; ICTP, Carboxy-terminal Telopeptide of Type-I Collagen; R, Randomized; DB, Double Blind; PC, Placebo-Controlled; C, Controlled; RIA, Radioimmunoassay; IRMA, Immunoradiometric Assay; ECLIA, Electro-Chemiluminescence Immunoassay; ELISA, Enzyme-linked Immunosorbent Assay; GH, Growth Hormone; Pl, Placebo; CT, Calcitonin; HRT, Hormone Replacement Therapy

Table 5: Summary of meta-analysis results for bone turnover markers

Bone Marker (unit)	**Studies Included**	**Sample Size** GH/Control	**WMD** (95% CI)	**Test of null** Z value (*p*-value)	**Test of heterogeneity** Chi2 (p-value)	**Inconsistency** I^2
Osteocalcin (µg/L)	Erdtsieck 1995 Gonnelli 1997 Holloway 1997 Landin-Wilhelmsen 2003	107/86	0.79 (-0.28; 1.85)	1.44 (0.15)	19.37 (0.0007)	79%
PICP (µg/L)	Erdtsieck 1995 Holloway 1997 Saaf 1999 Landin-Wilhelmsen 2003	103/82	14.03 (2.68; 25.38)	2.42 (0.02)	14.53 (0.006)	72%
Total urinary pyridinolines (nmol/mMCr)	Holloway 1997	32/40	4.97 (-0.30; 10.23)	1.85 (0.06)	3.19 (0.07)	69%
CTX (µg/mMCr)	Holloway 1997	30/29	38.32 (-5.75; 82.39)	1.70 (0.09)	4.78 (0.03)	79%

GH, Growth Hormone; WMD, Weighted Mean Difference; CI, Confidence Interval; Z value, Test statistic of the WMD; Chi2, Chi-squared test of heterogeneity; PICP, Procollagen Type-I Carboxy-terminal Propeptide; CTX, Carboxy-terminal Collagen Crosslinks

Table 6: Summary of meta-analysis results for fracture risk

Outcome	**Studies Included**	**Sample Size** GH/Control	**RR** (95% CI)	**Test of null** Z value (*p*-value)	**Test of heterogeneity** Chi2 (p-value)	**Inconsistency** I^2
Fractures at latest follow-up	Aloia 1985 Aloia 1987 Saaf 1999 Krantz 2015	106/145	0.63 (0.46;0.87)	2.87 (0.004)	0.58 (0.75)	0%
Fractures at end of treatment period	Aloia 1985 Aloia 1987 Saaf 1999 Landin-Wilhelmsen 2003	81/50	0.63 (0.30;1.51)	0.96 (0.34)	0.58 (0.45)	0%
Vertebral Fractures at latest follow-up	Aloia 1985 Aloia 1987 Saaf 1999 Krantz 2015	106/145	0.59 (0.29;1.21)	1.44 (0.15)	1.09 (0.58)	0%
Vertebral Fractures at end of treatment period	Aloia 1985 Aloia 1987 Saaf 1999 Landin-Wilhelmsen 2003	81/50	0.63 (0.30;1.51)	0.96 (0.34)	0.58 (0.45)	0%

GH, Growth Hormone; RR, Relative Risk; CI, Confidence Interval; Z value, Test statistic of the RR; Chi2, Chi-squared test of heterogeneity

Table 7: Adverse events and mortality reported in included studies

Study	Approximate mean target GH dose [IU/d (mg/d)]	Number of subjects	Number withdrawn	Reason for stopping protocol in GH group	Adverse Events from GH	Mortality
Aloia *et al*, 1985	hgH 6 IU (2 mg/d) 3x/w	25	0	NA	2 trigger-fingers, 1 carpal tunnel	None reported
Aloia *et al*, 1987	hGH 7 IU/d (2.3 mg/d) for 2m, repeated every 6m for 3 cycles	14	0	NA	none reported	None reported
Erdtsieck *et al*, 1995	rhGH 0.0625 IU/kg (max 4 IU) 3x/w (0.02 mg/kg)	23	2 (1 in GH, 1 in Pl)	Fluid retention	Fluid retention in 1 patient	None reported
Gonnelli *et al*, 1997	rhGH 12IU/d (4 mg/d) for 7 d, repeated every 82 days for 8 cycles	30	4 (1 in GH/CT, 1 in Pl/CT, 2 in GH/Pl)	2 for personal reasons, 1 for non-compliance	Transient arthralgia, muscle pain, ankle swelling	None reported
Holloway *et al*, 1997	rhGH 0.02 mg/kg/d for 7 d, repeated every 49 d for 12 cycles	84	12 (5 in GH/CT, 4 in GH/Pl and 3 in Pl)	Peripheral edema that was not controlled by diuresis	Peripheral edema	None reported
Saaf *et al*, 1999	rhGH 1.5 ± 0.3 U/d (0.5 mg/d) Range 0.9-2.7 U/d (0.3-0.9 mg/d)	16	4 (in GH group)	2 from carpal tunnel syndrome, 1 from retinal vein thrombosis, 1 from thigh ache	5 mild and reversible stiffness, 2 carpal tunnel, 1 trigger finger, 4 peripheral edema, 4 swollen aching knees, 3 trochanteritis, 3 mild muscular tenderness	None reported
Landin-Wilhelmsen *et al*, 2003	rhGH 1 U/d (0.33 mg/d) or 2.5 U/d (0.83 mg/d)	80	2 (in high dose GH group)	1 from arthralgias and 1 from ichthyosis	**High dose GH:** 1 DVT, 1 breast CA, 1 diverticulitis, 1 accidental patellar fx, 1 influenza **Low dose GH:** bronchitis, DM, bradycardia, angina, 1 DVT	None reported

GH, Growth Hormone; Pl, Placebo; CT, Calcitonin; DVT, Deep vein thrombosis

Table 8:. Evaluation of the quality of the evidence using GRADE

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Growth Hormone	Control	Relative (95% CI)	Absolute (95% CI)		
BMD												
3	randomised trials	serious [1]	not serious	not serious	serious [2]		87	64		MD 0 (0.03 fewer to 0.04 more)	Low	Important
BMC												
2	randomised trials	not serious	serious [3]	not serious	not serious		65	36		MD 0.71 fewer (1.63 fewer to 0.22 more)	Moderate	Important
Fracture risk												
	Randomised trials + 1 observational study	serious [1]	not serious	not serious	serious [2]		44/106 (41.5%)	76/145 (52.4%)	RR 0.63 (0.46 to 0.87)	194 fewer per 1000 (from 68 fewer to 253 fewer)	Very Low	Critical
PICP												
4	randomised trials	serious [1]	serious [3]	not serious	not serious		103	82		MD 14.03 more (2.68 more to 25.38 more)	Low	IMPORTANT

Appendix 1: PRISMA Checklist of items to include when reporting a systematic review or meta-analysis

Section/topic	Item number	Checklist item
Title		
Title	1	Identify the report as a systematic review, meta-analysis, or both
Abstract		
Structured summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number
Introduction		
Rationale	3	Describe the rationale for the review in the context of what is already known
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)
Methods		
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (such as web address), and, if available, provide registration information including registration number
Eligibility criteria	6	Specify study characteristics (such as PICOS, length of follow-up) and report characteristics (such as years considered, language, publication status) used as criteria for eligibility, giving rationale
Information sources	7	Describe all information sources (such as databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated
Study selection	9	State the process for selecting studies (that is, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)
Data collection process	10	Describe method of data extraction from reports (such as piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators

Section/topic	Item number	Checklist item
Data items	11	List and define all variables for which data were sought (such as PICOS, funding sources) and any assumptions and simplifications made
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis
Summary measures	13	State the principal summary measures (such as risk ratio, difference in means).
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (such as I2) for each meta-analysis
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (such as publication bias, selective reporting within studies)
Additional analyses	16	Describe methods of additional analyses (such as sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified
Results		
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram
Study characteristics	18	For each study, present characteristics for which data were extracted (such as study size, PICOS, follow-up period) and provide the citations
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome-level assessment (see item 12).
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15)
Additional analysis	23	Give results of additional analyses, if done (such as sensitivity or subgroup analyses, meta-regression [see item 16])
Discussion		
Summary of evidence	24	Summarise the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (such as health care providers, users, and policy makers)
Limitations	25	Discuss limitations at study and outcome level (such as risk of bias), and at review level (such as incomplete retrieval of identified research, reporting bias)

Section/topic	Item number	Checklist item
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research
Funding		
Funding	27	Describe sources of funding for the systematic review and other support (such as supply of data) and role of funders for the systematic review

Appendix 2: Search strategy

Medline Search Strategy

Database: Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R) <1946 to Present>
Search Strategy:
---
1 exp Osteoporosis, Postmenopausal/ or exp Osteoporosis/
2 exp Bone Density/
3 osteop*.ti,ab.
4 bone loss*.ti,ab.
5 (low adj2 (density* or content*) adj3 bone*).ti,ab.
6 demineralis*.ti,ab.
7 exp fractures, bone/
8 (bone* fracture* or bone* broken*).ti,ab.
9 or/1-8
10 exp Growth Hormone/
11 exp Human Growth Hormone/
12 growth hormon*.ti,ab.
13 (gh or hgh or h-gh or rhgh or r-hgh or rh-gh).ti,ab.
14 (somatrop* or somatotrop*).ti,ab.
15 (genotropin or humatrope or norditropin or nutropin or omnitrope or saizen or serostim or tev-tropin or zorbtive).ti,ab.
16 or/10-15
17 9 and 16

****

Embase Search Strategy

#2. 'osteoporosis'/exp OR 'bone density'/exp OR 'fracture'/exp OR osteop*:ab,ti OR (bone NEAR/2 density):ab,ti OR (bone:ab,ti AND mineral:ab,ti AND content:ab,ti) OR (bone NEAR/1 loss):ab,ti OR demineral*:ab,ti OR (bone*:ab,ti AND fracture*:ab,ti) OR (bone*:ab,ti AND broken*:ab,ti) AND ('growth hormone'/exp OR (growth:ab,ti AND hormone*:ab,ti) OR gh:ab,ti OR hgh:ab,ti OR 'h gh':ab,ti OR rhgh:ab,ti OR 'r hgh':ab,ti OR 'rh gh':ab,ti OR genotropin:ab,ti OR humatrope:ab,ti OR norditropin:ab,ti OR nutropin:ab,ti OR omnitrope:ab,ti OR saizen:ab,ti OR serostim:ab,ti OR 'tev tropin':ab,ti OR zorbtive:ab,ti OR somatrop*:ab,ti OR somatotrop*:ab,ti OR somacton:ab,ti OR somantin:ab,ti OR somatotrofin:ab,ti)

#1. 'osteoporosis'/exp OR 'bone density'/exp OR 'fracture'/exp OR osteop*:ab,ti OR (bone NEAR/2 density):ab,ti OR (bone:ab,ti AND mineral:ab,ti AND content:ab,ti) OR (bone NEAR/1 loss):ab,ti OR demineral*:ab,ti OR (bone*:ab,ti AND fracture*:ab,ti) OR (bone*:ab,ti AND broken*:ab,ti) AND ('growth hormone'/exp OR

(growth:ab,ti AND hormone*:ab,ti) OR gh:ab,ti OR

hgh:ab,ti OR 'h gh':ab,ti OR rhgh:ab,ti OR 'r

hgh':ab,ti OR 'rh gh':ab,ti OR genotropin:ab,ti

OR humatrope:ab,ti OR norditropin:ab,ti OR

nutropin:ab,ti OR omnitrope:ab,ti OR saizen:ab,ti

OR serostim:ab,ti OR 'tev tropin':ab,ti OR

zorbtive:ab,ti OR somatrop*:ab,ti OR

somatotrop*:ab,ti OR somacton:ab,ti OR

somantin:ab,ti OR somatotrofin:ab,ti)

Cochrane Search Strategy

ID	Search
#1	MeSH descriptor: [Osteoporosis] explode all trees
#2	MeSH descriptor: [Bone Density] explode all trees
#3	MeSH descriptor: [Fractures, Bone] explode all trees
#4	osteop*
#5	bone loss*
#6	low near/2 bone near/2 density
#7	low near/2 bone near/2 content
#8	demineralis*
#9	bone* fracture*
#10	bone* broken*
#11	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10
#12	MeSH descriptor: [Growth Hormone] explode all trees
#13	MeSH descriptor: [Human Growth Hormone] explode all trees
#14	growth hormon*
#15	gh or hgh or h-gh or rhgh or r-hgh or rh-gh
#16	somatrop* or somatotrop*
#17	genotropin or humatrope or norditropin or nutropin or omnitrope or saizen or serostim or tev-tropin or zorbtive
#18	#12 or #13 or #14 or #15 or #16 or #17
#19	#11 and #18

Appendix 3: Full-text screening form

Study ID: First author: Year: Screener initials:

1- Is study design a prospective controlled trial or not?

No --------- → Exclude
Yes -------- → Go to the next question

2- Is the study population: Postmenopausal women or men above the age of 50 years with osteoporosis or osteopenia (defined as BMD T-score less than -1 or presence of fragility fracture), without secondary causes of osteoporosis and without growth hormone deficiency?

No --------- → Exclude
Yes -------- → Go to the next question

3- Is the intervention growth hormone therapy?

No --------- → Exclude
Yes -------- → Go to the next question

4- Is comparison to a comparator group?

No --------- → Exclude
Yes -------- → Go to the next question

5- Are outcomes needed present?

No --------- → Exclude
Yes --------

Final decision:

Reason for exclusion (Please select):

1. Lack of adequate design
2. Lack of adequate population
3. Lack of adequate intervention
4. Lack of adequate comparison
5. Lack of adequate outcome assessment
6. Other:

Appendix 4: Data extraction form

		Participants						**Osteoporosis/Osteopenia**					**Intervention**				**Outcome**		
First author, year	Study design	Age (yrs)	Yrs since menopause	Gender (%M/F)	BMI (kg/m2)	Number of subjects	Number withdrawn	Yrs since diagnosis	Prior tx	Concurrent tx	Number/ site of fractures	Baseline T-score	Approximate mean target GH dose [IU/d (mg/d)]	Duration of therapy (m)	IGF-1 baseline	IGF-1 after therapy	Outcome assessed	Funding	Conflict of interest

REFERENCES

1. "Who's at risk?," The International Osteoporosis Foundation, accessed August 30, 2015, http://www.iofbonehealth.org/whos-risk.
2. Kanis JA. 2002 Diagnosis of osteoporosis and assessment of fracture risk. *The Lancet* 359(9321): 1929-1936
3. World Health Organization. 1994 Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. *World Health Organ Tech Rep Ser* 843: 1-129
4. Stone KL, Seeley DG, Lui LY et al. 2003 BMD at multiple sites and risk of fracture of multiple types: long-term results from the Study of Osteoporotic Fractures. *J Bone Miner Res* 18(11): 1947-1954
5. Melton LJ, Atkinson EJ, O'Connor MK, et al. 1998 Bone density and fracture risk in men. *J Bone Miner Res* 13:1915
6. Melton LJ, Chrischilles EA, Cooper C, et al. 1992 Perspective. How many women have osteoporosis? *J Bone Miner Res* 7:1005
7. Kanis JA, Johnell O, Oden A, et al. 2000 Long-term risk of osteoporotic fracture in Malmo. *Osteoporos Int* 11:669
8. National Osteoporosis Foundation. 2014 Clinician's guide to prevention and treatment of osteoporosis. Washington, DC: National Osteoporosis Foundation
9. Lippuner K 2012 The future of osteoporosis treatment – A research update. *Swiss Med Wkly* 142:13624
10. Suresh E, Pazianas M, Abrahamsen B 2014 Safety issues with bisphosphonate therapy for osteoporosis. *Rheumatology* 53:19-31
11. Van Cauter E, Plat L, Copinschi G 1998 Interrelations between sleep and the somatotropic axis. *Sleep* 21: 553-566
12. Giustina A, Veldhuis JD 1998 Pathophysiology of the neuroregulation of growth hormone secretion in experimental animals and the human. *Endocrine Reviews* 19/6: 717-797
13. Baumann G, Stalor MW, Armburn K et al. 1986 A specific growth hormone-binding protein in human plasma: initial characterization. *J Clin Endocrinol Metab* 62: 134-141
14. Baumann G, Shaw MA 1990 A second, lower affinity growth hormone-binding protein in human plasma. *J Clin Endocrinol Metab* 70: 680-686
15. Fisker S, Vahl N, Jorgensen JOL et al. 1997 Abdominal fat determines growth hormone binding protein levels in healthy nonobese adults. *J Clin Endocrinol Metab* 82: 123-128
16. Drake WM, Howell SJ, Monson JP et al. 2001 Optimizing GH therapy in adults and children. *End Rev* 22(4): 425-450
17. D'Ercole AJ, Stiles AD, Underwood LE 1984 Tissue concentrations of somatomedin C: further evidence for multiple sites of synthesis and paracrine or autocrine mechanisms of action. *Proc Natl Acad Sci USA* 81: 935-939
18. Van den Berg G, Veldhuis JD, Frolich M et al. 1996 An amplitude-specific divergence in the pulsatile mode of growth hormone (GH) secretion underlies the gender difference in mean GH concentrations in men and premenopausal women. *J Clin Endocrinol Metab* 81: 2460-2467

19. Goodman-Gruen D, Barrett-Connor E 1996 Effect of replacement estrogen on insulin-like growth factor-I in postmenopausal women: the Rancho Bernardo Study. *J Clin Endocrinol Metab* 81: 4268-4271
20. Copeland KC, Colletti RB, Devlin JT et al. 1990 The relationship between insulin-like growth factor-I, adiposity, and aging. *Metab Clin Exp* 39: 584-587
21. Zadik Z, Chalew SA, McCarter Jr RJ et al. 1985 The influence of age on the 24-hour integrated concentration of growth hormone in normal individuals. *J Clin Endocrinol Metab* 60: 513–516
22. Iranmanesh A, Lizarralde G, Veldhuis JD 1991 Age and relative adiposity are specific negative determinants of the frequency and amplitude of growth hormone (GH) secretory bursts and the half-life of endogenous GH in healthy men. *J Clin Endocrinol Metab* 73: 1081–1088
23. Clemmons DR, Van Wyk JJ 1984 Factors controlling blood concentration of somatomedin C. *Clin Endocrinol Metab* 13: 113–143
24. Corpas E, Harman SM, Pineyro MA et al. 1992 Growth hormone (GH)-releasing hormone (1–29) twice daily reverses the decreased GH and insulin-like growth factor-I levels in old men. *J Clin Endocrinol Metab* 75: 530–53
25. Ohlsson C, Bengtsson BA, Isaksson OGP, et al. 1998 Growth hormone and bone. *Endocrine Reviews* 19(1): 55-79
26. Parker EA, Hegde A, Buckley M et al 2007 Spatial and temporal regulation of GHIGF- related gene expression in growth plate cartilage. *Journal of Endocrinology* 194: 31-40
27. Barnard R, Ng KW, Martin TJ, et al. 1991 Growth hormone (GH) receptors in clonal osteoblast-like cells mediates a mitogenic response to GH. *Endocrinology* 128:1459-1464
28. Chihara K, Sugimoto T 1997 The action of GH/IGF-1/IGFBP in osteoblasts and osteoclasts. *Horm Res* 48:45-49
29. Krishnan V, Bryant HU, MacDougald A 2006 Regulation of bone mass by Wnt signaling. *J Clin Invest* 116(5): 1202-1209
30. Lindahl A, Isgaard J, Nilsson A et al. 1986 Growth hormone potentiates colony formation of epiphyseal chondrocytes in suspension culture. *Endocrinology* 118(5): 1843-1848
31. Ernst M, Rodan GA 1990 Increased activity of insulin-like growth factor (IGF) in osteoblastic cells in the presence of growth hormone (GH): positive correlation with the presence of the GH-induced IGF-binding protein BP-3. *Endocrinology* 127(2): 807-814
32. Uronen-Hansson H, Allen ML, Lichtarowicz-Krynska E et al. 2003 Growth hormone enhances proinflammatory cytokine production by monocytes in whole blood. *GH & IGF Res* 13(5): 282-286
33. Renier G, Clement I, Desfaits AC et al. 1996 Direct stimulatory effect of insulin-like growth factor-I on monocyte and macrophage tumor necrosis factor-α production. *Endocrinology* 137(11): 4611-4618
34. Niu T, Rosen CJ 2005 The insulin-like growth factor-I gene and osteoporosis: a critical appraisal. *Gene* 361(1-2): 38-56
35. Harris WH, Heaney RP 1969 Effect of growth hormone on skeletal mass in adult digs. *Nature* 223: 403-404

36. Sims NA, Clement-Lacroix P, Da Ponte F et al. 2000 Bone homeostasis in growth hormone receptor-null mice is restored by IGF-I but independent of Stat-5. *J Clin Invest* 106(9): 1095-1103
37. Yakar S, Canalis E, Sun H et al. 2009 Serum IGF-I determines skeletal strength by regulating subperiosteal expansion and trait interactions. *J Bone Miner Res* 24(8): 1481-1492
38. Zhang M, Xuan S, Bouxsein ML, et al. 2002 Osteoblast-specific knockout of the insulin-like growth factor (IGF) receptor gene reveals an essential role of IGF signaling in bone matrix mineralization. *J Biol Chem* 277: 44005-44012
39. Lupu F, Terwilliger JD, Lee K et al. 2001 Roles of growth hormone and insulin-like growth factor I in mouse postanatal growth. *Dev Biol* 229:141-162
40. Baroncelli GI, Bertelloni S, Sodini F, et al. 2003 Acquisition of bone mass in normal individuals and in patients with growth hormone deficiency. *J Pediatr Endocrinol Metab* 16 Suppl 2:327-35
41. Carroll PV, Christ ER, Bengtsson BA, et al. 1998 Growth hormone deficiency in adulthood and the effects of growth hormone replacement: a review. Growth Hormone Research Society Scientific Committee. *J Clin Endocrinol Metab* 83:382-95
42. Koranyi J, Svensson J, Gotherstrom G et al. 2001 Baseline characteristics and the effects of five years of GH replacement therapy in adults with GH deficiency of childhood or adulthood onset: a comparative, prospective study. *J Clin Endcocrinol Metab* 86: 4693-4699
43. Colao A, Di Somma C, Pivonello R, et al. 1999 Bone loss is correlated to the severity of growth hormone deficiency in adult patients with hypopituitarism. *J Clin Endocrinol Metab* 84: 1919-1924
44. Mazziotti G, Bianchi A, Cimino V et al. 2008 Effect of gonadal status on bone mineral density and radiological spinal deformities in adult patients with growth hormone deficiency. *Pituitary* 11: 55-61
45. Hitz MF, Jensen JE, Eskildsen PC 2006 Bone mineral density in patients with growth hormone deficiency: does a gender difference exist? *Clin Endocrinol (Oxf)* 65: 783-791
46. Rosen T, Wilhelmsen L, Landin-Wilhelmsen K, et al. 1997 Increased fracture frequency in adult patients with hypopituitarism and GH deficiency. *Eur J Endocrinol* 137: 240-5
47. Wuster C, Abs R, Bengtsson BA et al. 2001 The influence of growth hormone deficiency, growth hormone replacement therapy, and other aspects of hypopituitarism on fracture rates and bone mineral density. *J Bone Miner Res* 16: 398-405
48. Mazziotti G, Bianchi A, Bonadonna S et al. 2006 Increased prevalence of radiological spinal deformities in adult patients with GH deficiency: influence of GH replacement therapy. *J Bone Miner Res* 21: 520-528
49. Mazziotti G, Doga M, Frara S, et al. 2015 Incidence of morphometric vertebral fractures in adult patients with growth hormone deficiency. *Endocrine* [epub ahead of print] DOI 10.1007/s12020-015-0738-z
50. Gharib H, Cook DM, Saenger PH, et al. 2003 American Association of Clinical Endocrinologists medical guidelines for clinical practice for growth hormone use in adults and children – 2003 update. *Endocr Pract* 9:64-76

51. Ho KK 2007 Consensus guidelines for the diagnosis and treatment of adults with GH deficiency II: a statement of the GH Research Society in association with the European Society for Pediatric Endocrinology, Lawson Wilkins Society, European Society of Endocrinology, Japan Endocrine Societ, and Endocrine Society of Australia. *Eur J Endocrinol* 157(6): 695-700
52. Conway G, Szarras-Czapnik M, Racz K, et al. 2009 Treatment for 24 months with recombinant human growth hormone has a beneficial effect on bone mineral density in young adults with childhood-onset growth hormone deficiency. *Eur J Endocrinol* 160(6): 899-907
53. Hansen TB, Brixen K, Vahl N, et al. 1996 Effects of 12 months of growth hormone (GH) treatment on calciotropic hormones, calcium homeostasis, and bone metabolism in adults with acquired GH deficiency: a double blind, randomized, placebo-controlled study. *J Clin Endocrinol Metab* 81(9): 3352–3359
54. Marcus R. 1998 Recombinant human growth hormone as potential therapy for osteoporosis. *Bailliere's Clin Endocrinol Metab* 12(2): 251-259
55. Barake M, Klibanski A, Tritos NA 2014 Effects of recombinant human growth hormone therapy on bone mineral density in adults with growth hormone deficiency: a meta-analysis. *J Clin Endocrinol Metab* 99(3): 852-860
56. Gotherstrom G, Bengtsson BA, Bossaeus I, et al. 2007 Ten-year GH replacement increase bone mineral density in hypopituitary patients with adult onset GH deficiency. *Eur J Endocrinol* 156(1): 55-64
57. Elbornsson M, Gotherstrom G, Bossaeus I, et al. 2012 Fifteen years of GH replacement increases bone mineral density in hypopituitary patients with adult-onset GH deficiency. *Eur J Endocrinol* 166(5): 787-795
58. Appelman-Dijkstra NM, Claessen KM, Hamdy NA, et al. 2014 Effects of up to 15 years of recombinant human GH (rhGH) replacement on bone metabolism in adults with growth hormone deficiency (GHD): the Leiden Cohort Study. *Clin Endocrinol (Oxf)* 81(5): 727-735
59. Biller BM, Sesmilo G, Baum HB, et al. 2000 Withdrawal of long-term physiological growth hormone (GH) administration: differential effects on bone density and body composition in men with adult-onset GH deficiency. *J Clin Endocrinol Metab* 85(3): 970-976
60. Snyder PJ, Biller BM, Zagar A, et al. Effect of growth hormone replacement on BMD in adult-onset growth hormone deficiency. *J Bone Miner Res* 22(5): 762-770
61. Mo D, Fleseriu M, Qi R, et al. 2015 Fracture risk in adult patients treated with growth hormone replacement therapy for growth hormone deficiency: a prospective observational cohort study. *Lancet Diabetes Endocrinol* 3(5):331-338
62. Giustina A, Mazziotti G. 2015 Growth hormone replacement therapy and fracture risk. *Lancet Diabetes Endocrinol* 3(5): 307-308
63. Romagnoli E, Minisola S, Carnevale V et al. 1994 Circulating levels of insulin-like growth factor binding protein 3 (IGFBP-3) and insulin-like growth factor I (IGF-I) in perimenopausal women. *Osteop Int* 4(6): 305-308
64. Nasu M, Sugimoto T, Chihara M, et al. 1997 Effect of natural menopause on serum levels of IGF-I and IGF-binding proteins: Relationship with bone

mineral density and lipid metabolism in perimenopausal women. *Eur J Endocrinol* 136(6): 608-616
65. Boonen S, Lesaffre E, Aerssens J, et al. 1996 Deficiency of the growth hormone, insulin-like growth factor-1 axis potentially involved in age-related alterations in body composition. *Gerontology* 42(6): 330-338
66. Munoz-Torres M, Mezquita-Raya P, Lopez-Rodriguez F, et al. 2001 The contribution of IGF-I to skeletal integrity in postmenopausal women. *Clin Endocrinol* 55(6): 759-766
67. Ljunghall S, Johansson AG, Burman P, et al. 1992 Low plasma levels of insulin-like growth factor-1 (IGF-1) in male patients with idiopathic osteoporosis. *J Int Med* 232(1): 59-64
68. Gillberg P, Olofsson H, Mallmin H, et al. 2002 Bone mineral density in femoral neck is positively correlated to circulating insulin-like growth factor (IGF)-I and IGF-binding protein (IGFBP)-3 in Swedish men. *Cal Tissue Int* 70(1): 22-29
69. Wuster C, Blum WF, Schlemilch S, et al. 1993 Decreased serum levels of insulin-like growth factors and IGF binding protein 3 in osteoporosis. *J Int Med* 234(3): 249-255
70. Landin-Wilhelmsen K, Wilhelmsen L, Bengtsson BA 1999 Postmenopausal osteoporosis is more related to hormonal aberrations than to lifestyle factors. *Clin Endocrinol (Oxf)* 51:387-394
71. Ohlsson C, Mellstrom D, Carlzon D, et al. 2011 Older men with low serum IGF-1 have an increased risk of incident fractures: the MrOS Sweden study. *J Bone Min Res* 26(4): 865-872
72. Garnero P, Sornay-Rendu E, Delmas PD. 2000 Low serum IGF-1 and occurrence of osteoporotic fractures in postmenopausal women. *Lancet* 355(9207): 898-899
73. Landin-Wilhelmsen K, Nilsson A, Bosaeus I, et al. 2003 Growth hormone increases bone mineral content in postmenopausal osteoporosis: a randomized placebo-controlled trial. *J Bone Miner Res* 18:393-405
74. Saaf M, Hilding A, Thoren M, et al. 1999 Growth hormone treatment of osteoporotic postmenopausal women – a one-year placebo-controlled study. *Eur J Endocrinol* 140: 390-399
75. Yang S, Cao L, Cai S, et al. 2012 A systematic review of growth hormone for hip fractures. *Growth Hormone & IgF Research* 22:97-101
76. Liu H, Bravata DM, Olkin I, et al. 2007 Systematic review: the safety and efficacy of growth hormone in the healthy elderly. *Ann Intern Med* 146: 104-115
77. PRISMA statement. Available from www.prisma-statement.org/. Accessed in January 2016
78. COMET initiative. Available from www.comet-initiative.org/. Accessed in January 2016
79. Cranney A, Tugwell P, Cummings S, et al. 1997 Osteoporosis clinical trials endpoints: Candidate variables and clinimetric properties. *J Rheumatol* 24(6): 1222-1229
80. Larcos G, Wahner HW. 1991 An evaluation of forearm bone mineral measurement with dual-energy x-ray absorptiometry. *J Nucl Med* 32(1): 2101-2106

81. Neer R. 1992 The utility of single-photon absoptiometry and dual-energy x-ray absorptiometry. *J Nucl Med* 33(1): 170-171
82. Lees B, Stevenson JC. 1992 An evaluation of dual-energy x-ray absorptiometry and comparison with dual-photon absorptiometry. *Osteoporosis Int* 2(3): 146-152
83. Burch J, Rice S, Yang H, et al. 2014 Systematic review of the use of bone turnover markers for monitoring the response to osteoporosis treatment: the secondary prevention of fractures, and primary prevention of fractures in high-risk groups. *Health Technol Assess* 18(11): 1-180
84. Bezza A, Ouzzif Z, Naji H, et al. 2008 Prevalence and risk factors of osteoporosis in patients with Parkinson's disease. *Rheumatol Int* 28(12): 1205-1209
85. Choi HJ, Im JA, Kim SH, et al. 2008 Changes in bone markers after once-weekly low-dose alendronate in postmenopausal women with moderate bone loss. *Maturitas* 60(2): 170-176
86. Erdtsieck RJ, Pols HAP, Valk NK, et al. 1995 Treatment of post-menopausal osteoporosis wuth a combination of growth hormone and pamidronate: a placebo controlled trial. *Clin Endcocrinol* 43(5): 557-565
87. Holloway L, Kohlmeier L, Kent K, et al. 1997 Skeletal effects of cyclic recombinant human growth hormone and salmon calcitonin in osteopenic postmenopausal women. *J Clin Endocrinol Metab* 82(4): 1111-1117
88. Krantz E, Trimpou P, Landin-Wilhelmsen K. 2015 Effect of growth hormone treatment on fractures and quality of life in postmenopausal osteoporosis: a 10-year follow-up study *J Clin Endocrinol Metab* 100(9): 3251-3259
89. Higgins JPT, Green S. 2011 Cochrane handbook for systematic reviews of interventions. Version 5.1.0. The Cochrane Collaboration
90. Viswanathan M, Berkman ND, Dryden DM, et al. 2013 Assessing Risk of Bias and Confounding in Observational Studies of Interventions or Exposures: Further Development of the RTI Item Bank. *AHRQ Publication No. 13-EHC106-EF*
91. Aloia JF, Vaswani A, Kapoor A, et al. 1985 Treatment of osteoporosis with calcitonin, with and without groth hormone. *Metabolism* 34(2): 124-129
92. Aloia JF, Vaswani A, Meunier PJ, et al. 1987 Coherence treatment of postmenopausal osteoporosis with growth hormone and calcitonin. *Calcif Tissue Int* 40(5): 253-259
93. Statistics How To. Available from http://www.statisticshotto.com/weighted_mean/. Accessed in February 2016
94. Stat Trek. Available from: http://stattrek.com/statistics/formulas.aspx/. Accessed in February 2016
95. Guyatt GH, Oxman AD, Vist GE, et al. 2008 GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336(7650): 924-926
96. Gonnelli S, Cepollaro C, Montomoli M, et al. 1997 Treatment of post-menopausal osteoporosis with recombinant human growth hormone and salmon calcitonin: a placebo controlled study. *Clin Endocrinol (Oxf)* 46(1): 55-61
97. Joseph F, Ahmad AM, Ul-Haq M, et al. 2008 Effects of growth hormone administration on bone mineral metabolism, PTH sensitivity and PTH

secretory rhythm in postmenopausal women with established osteoporosis. *J Bone Miner Res* 23(5): 721-729
98. Marcus R. 1987 Normal and abnormal bone remodeling in man. *Ann Rev Med* 38: 129-141
99. Molitch ME, Clemmons DR, Malozowski S, et al. 2011 Evaluation and treatment of adult growth hormone deficiency: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 96(6): 1587-160
100. Gillberg P, Mallmin H, Petren-Mallmin M, et al. 2002 Two years of treatment with recombinant human growth hormone increases bone mineral density in men with idiopathic osteoporosis. *J Clin Endocrinol Metab* 87(11): 4900-4906
101. Valk NK, Erdtsieck RJ, Algra D, et al. 1995 Combined treatment of growth hormone and the bisphosphonate pamidronate, versus treatment with GH alone, in GH-deficient adults: the effects on renal phosphate handling, bone turnover and bone mineral mass. *Clin Endocrinol (Oxf)* 43(3): 317-324
102. Biermasz NR, Hamdy NAT, Janssen YJH, et al. 2001 Additional beneficial effects of alendronate in growth hormone (GH)-deficient adults with osteoporosis receiving long-term recombinant human GH replacement therapy: A randomized controlled trial. *J Clin Endocrinol Metab* 86(7): 3079-3085
103. Finkelstein JS, Hayes A, Hunzelman JL, et al. 2003 The effects of parathyroid hormone, alendronate, or both in men with osteoporosis. *N Engl J Med* 349(13): 1216-1226
104. Tsai JN, Uihlein AV, Lee H, et al. 2013 Teriparatide and denosumab, alone or combined, in women with postmenopausal osteoporosis: the DATA study randomized trial. *Lancet* 382(9886): 50-56
105. Tsai JN, Uihlein AV, Burnett-Bowie SM, et al. 2016 Effects of two years of teriparatide, denosumab, or both on bone microarchitecture and strength (DATA-HRpQCT study). *J Clin Endocrinol Metab* jc20161160 [Epub ahead of print]
106. Allen DB, Backeljauw P, Bidlingmaier M, et al. 2016 GH safety workshop position paper: a critical appraisal of recombinant human GH therapy in children and adults. *Eur J Endocrinol* 174(2): 1-9